Development of a child-centred preference-based measure of dental caries, suitable for use in economic evaluation

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy

University of Sheffield
School of Clinical Dentistry

Helen Jessica Rogers

September 2020

Acknowledgements

First and foremost, I would like to thank all the participants in this study without whom this research would not have been possible. I am particularly grateful to the Young People Panel members and parent representatives for their invaluable contribution in steering this study.

I owe a special thank you to my supervisors, Professor Zoe Marshman, Dr Donna Rowen and Professor Helen Rodd, for their support, advice and encouragement throughout this study. I am especially grateful to Donna for her patience and kindness during my attempts to get to grips with Stata!

I am indebted to the NIHR for funding this study and supporting me to access the necessary training to ensure its success. A further thanks to the Research Design Service for awarding me a Patient and Public Involvement grant, and to the Clinical Research Network for providing additional funding and support.

There have been many individuals both within the School of Clinical Dentistry and the School of Health and Related Research who have supported this research; some of whom have since moved to pastures new. I would like to thank Dr Katherine Stevens for her advice and support in the early stages of this study; Dr Kathryn Hurrell-Gillingham for guiding me through the complexities of IRAS; Professor John Brazier for his invaluable feedback and suggestions; Dr Tracey Young for sharing her expertise in Rasch analysis; Professor Paul Speight for his encouragement throughout my fellowship applications and beyond; Dr Mario Vettore and Dr Chu-Chang Ku for kindly translating papers for my systematic review; Sarab El Yousfi and Becky Knapp for their help in extracting data for the systematic review and Laura Timms for her assistance with qualitative interviews. A particular thanks to Dr Fiona Gilchrist and Professor Chris Deery for their mentorship and advice over the years.

Outside of the University of Sheffield, I must thank Erik Vermaire for his constant support of my academic endeavours, his assistance with the systematic review and deliveries of stroopwafels and cheese! A further thanks to Julian Sagabiel for his expert advice on the analysis of best-worst scaling data, and to Alex Parkman from Surveyengine for guiding me through the survey design process. I am also especially grateful to staff members at the secondary school that participated in this study for facilitating this research.

Lastly, I cannot express enough thanks to my family and friends for their unending encouragement throughout this process. The biggest thanks goes to Abs and the boys for supporting me every step of the way, keeping me focused, calm and motivated, and for figuring out the formula for Cohen’s *d*!

Abstract

Background: Dental caries is a prevalent oral disease with significant negative impacts on the lives of children and young people, their families and wider society. Caries is a largely preventable disease, thus there are a range of different programmes available to reduce the prevalence in children. However, there have been few economic evaluations to determine the cost-effectiveness of such programmes, hence commissioning of interventions to improve children's oral health is poorly informed. Within child oral health research, this paucity of economic evaluations could be attributed to the lack of a suitable instrument to measure Quality Adjusted Life Years (QALYs). Development of a validated utility measure, specifically for children is both recommended and potentially achievable through the adaptation of a novel child-centred caries-specific oral health-related quality of life measure, CARIES-QC (Caries Impacts and Experiences Questionnaire for Children). Ordinal tasks, such as discrete choice experiments (DCE) and best-worst scaling (BWS), have shown promise for use in preference elicitation with children, yet there is little evidence to identify the most appropriate task to use.

Aim: To develop a preference-based measure (PBM) of dental caries based on the preferences of children and young people.

Design: Four interlinked studies were undertaken to meet this aim. An initial systematic review of the scope and quality of economic evaluations in child oral health research was conducted to inform the following stages. A preliminary classification system derived from CARIES-QC was identified, informed by Rasch analysis, classical psychometric testing, the views of children and young people, and the views of the developers of CARIES-QC. The preliminary classification system was then validated using a qualitative approach with children and young people aged 5-16 years with caries. Next, a qualitative evaluation of DCE and BWS tasks was undertaken with adolescents aged 11-16 years from a local secondary school to identify the most appropriate task to use with this population. The findings of this stage were used to inform the final study, which involved a UK-wide online valuation survey with children and young people. A concurrent valuation survey was conducted with adults to facilitate the anchoring of adolescent preferences onto the QALY scale.

Results: The systematic review identified 46 papers, highlighting a lack of high quality economic evaluations in child oral health research, with few cost-utility analyses and minimal involvement of children and young people as active participants, confirming the need for a child-centred PBM. These findings were used to inform the next stage, which identified five items from CARIES-QC relating to children experiencing pain (*‘hurt’*), feeling annoyed (*‘annoy’*), eating more carefully (*‘carefully’*), being kept awake (*‘kept awake’*) and crying (*‘cried’*) because of their teeth, which formed the preliminary classification system for a PBM specific to caries, to be named CARIES-QC-Utility (CARIES-QC-U). The qualitative validation study with 20 children revealed a preference for a different item relating to difficulty eating (*‘hard to eat’*), which was reinserted into the classification system in place of eating more *‘carefully’*. DCE and BWS tasks based upon this validated classification system were incorporated into a computer-based survey. This formed the basis for the qualitative interviews, which were conducted with 33 adolescents from a local secondary school. Adolescents had a greater understanding of, and a preference for BWS tasks compared to DCE. An online valuation survey comprising BWS tasks based upon the validated classification system was completed by 723 adolescents across the UK. A similar survey comprising DCE with duration tasks was completed by 626 adults concurrently. The values were analysed using the conditional logit model. Finally, the adolescent preferences were anchored onto the QALY scale using the adult values. The final algorithm was then able to estimate a preference weight for each health state defined by the classification system, enabling QALYs to be generated.

Conclusion: This child-centred, caries-specific preference-based measure, known as CARIES-QC-U (Caries Impacts and Experiences Questionnaire for Children – Utility) is suitable for use in economic evaluations to inform the commissioning of interventions to improve children’s oral health.

Frequently used abbreviations

BWS	Best-worst scaling
CARIES-QC	Caries Impacts and Experiences Questionnaire for Children
CARIES-QC-NL	Dutch version of Caries Impacts and Experiences Questionnaire for Children
CARIES-QC-U	Caries Impacts and Experiences Questionnaire for Children-Utility
CHEERS	Consolidated Health Economic Evaluation Reporting Standards statement
CHU9D	Child Health Utility 9 Dimension
CHU9D-NL	Dutch version of Child Health Utility 9 Dimension
CRN	Clinical Research Network
CSPBM	Condition-specific preference-based measure
DCE	Discrete choice experiment
DCE_{TTO}	Hybrid of discrete choice experiment and time trade off, also referred to as discrete choice experiment with duration
DMFT	Decayed, missing and filled permanent teeth
dmft	Decayed, missing and filled primary teeth
DR	Donna Rowen, Senior Research Fellow and Health Economist, The University of Sheffield
EQ-5D	Euroqol 5 Dimension
EQ-5D-5L	Euroqol 5 Dimension - 5 Level
EQ-5D-Y	Euroqol 5 Dimension - Youth
EV	Jan Hendrik (Erik) Vermaire, Senior Research Associate, Netherlands Organisation for Applied Scientific Research
GPBM	Generic preference-based measure
HDR	Helen Rodd, Professor in Paediatric Dentistry, The University of Sheffield
HJR	Helen Rogers, PhD student
HRQoL	Heath-related quality of life
JS	Julian Sagabiel, Research Associate, Berlin Institute of Technology
KS	Katherine Stevens, Honorary Reader in Health Economics, The University of Sheffield

NHS	National Health Service
NIHR	National Institute for Health Research
OHRQoL	Oral health-related quality of life
PBM	Preference-based measure
PPI	Patient and public involvement
QALYs	Quality Adjusted Life Years
QATYs	Quality Adjusted Tooth Years
RDS	Research Design Service
SF-36	Short Form-36
SG	Standard gamble
TTO	Time trade-off
UK	United Kingdom
ZM	Zoe Marshman, Professor in Dental Public Health, The University of Sheffield

Contents

List of appendices

List of tables

List of figures

1 Introduction

Dental caries is a prevalent oral disease amongst children and can cause chronic pain, local infection and in some cases may lead to emergency hospitalisation due to spread of the infection and systemic illness. In the UK, approximately 41,558 children aged up to 16-years were admitted to hospital in 2018-2019 with a diagnosis of dental caries. As a result, dental caries remains the most common reason for children to require an admission with an estimated annual cost of £39 million to the NHS (HSCIC, 2016).

Dental caries is a largely preventable disease, thus there are a range of different programmes available to reduce the prevalence in children. However, there have been few high quality economic evaluations to determine the cost effectiveness of such programmes. Within child oral health research, this scarcity of economic evaluations could be attributed to the lack of a suitable instrument to measure Quality Adjusted Life Years (QALYs). It can therefore be argued that there is a need for the development of a validated preference-based measure (PBM), specifically for children, that is appropriate for measuring the benefits of interventions to prevent and treat dental caries. This is achievable through the adaptation of a novel child-centred caries-specific oral health-related quality of life (OHRQoL) measure, known as CARIES-QC (Caries Impacts and Experiences Questionnaire for Children). Preference elicitation tasks such as discrete choice experiments (DCE) and best-worst scaling (BWS) have shown promise for use in the valuation of utility measures with children, yet there is little evidence to identify the most appropriate method to use.

This study aimed to develop a PBM of dental caries based upon the preferences of children and young people.

The following specific objectives were identified in order to meet this aim:

1. To identify the quality and scope of published economic evaluations in child oral health research and the measures of benefit currently used
2. To develop and validate a classification system for child dental caries, based upon the CARIES-QC caries-specific measure of OHRQoL, that is amenable to health state valuation

3. To determine what age range of children and young people can complete ordinal health state valuation tasks, whether they prefer DCE or BWS and at what age they can use a computerised format independently
4. To undertake a population-based valuation survey with children and young people to determine their preferences
5. To model their preferences to produce a valuation algorithm that provides preference weights for each health state defined by the classification system

This thesis describes a body of research undertaken in order to meet these objectives, which culminates in the development of CARIES-QC-Utility (CARIES-QC-U), a child-centred PBM specific to dental caries. An outline of each chapter is detailed below.

Chapter 2 provides a narrative review of the literature surrounding the significance of dental caries and its impacts on children, families and society. It introduces the concept of oral health-related quality of life and how it can be measured. The review highlights the lack of a suitable tool to evaluate the benefits of interventions to improve children's oral health that could be used in economic evaluations, and considers ways in which children could be involved in the development of such an instrument.

Chapter 3 clarifies the aims and objectives of this study in greater detail, as well as providing a summary of the rationale for this research.

Chapter 4 reports a systematic review of the quality and scope of economic evaluations in the field of child oral health research. This focuses on the limitations in quantity and quality (related to methodology and reporting) of published economic evaluations in this field, as well as the lack of involvement of children in these studies, considering ways in which this could be addressed.

Chapter 5 details the approaches taken to identifying a preliminary classification system for the PBM, including Rasch analysis, classical psychometric testing, involvement of children and young people, parent representatives and the views of the team involved in the development of the original CARIES-QC instrument. The qualitative, child-centred approaches to validate this preliminary classification system are also described.

Chapter 6 describes a school-based qualitative study designed to investigate the ability of adolescents to complete DCE and BWS tasks and which type of task they prefer. This centres on adolescents' self-completion of a survey incorporating both types of task, alongside the use of think-aloud and semi-structured interviews to explore their views and level of comprehension.

Chapter 7 details the development and administration of a valuation survey to elicit preferences from adolescents, based upon the findings from the previous chapter. It also provides details of a concurrent adult survey, which was undertaken in order to allow the adolescent values to be anchored onto the 1-0 full health to death QALY scale.

Chapter 8 provides a discussion of the wider considerations surrounding the use of a child-centred caries-specific PBM and the implications for research, policy and commissioning.

Chapter 9 demonstrates how the aims and objectives have been met and summarises the key conclusions of this study.

2 Background

This chapter explores the impacts of dental caries on children, young people and their families, and current approaches to prevent and manage the condition. An introduction to the concepts surrounding economic evaluation is provided, with consideration to how children and young people could be involved.

2.1 The oral health of children

There are a number of conditions affecting children's oral health, with various aetiologies, including developmental, environmental and traumatic. Of all the conditions, the most prevalent is a largely preventable disease, namely dental caries.

Dental caries, otherwise known as tooth decay, is a multifactorial condition, characterised by bacterial destruction of tooth tissue in the presence of dietary substrate. Ultimately, the disease can progress to cause chronic pain, infection and in some cases may lead to emergency hospitalisation (Majewski et al., 1988, Benjamin, 2010).

A recent systematic review reported 9% of children worldwide have untreated dental caries in their primary teeth, highlighting it as a major public health problem internationally (Kassebaum et al., 2015). Whilst the introduction of fluoridated toothpaste, alongside wider public health measures, has successfully reduced the prevalence of caries in developed countries since the 1970's, this is not the case in less developed countries, where caries experience has remained unchanged (Do, 2012). Nonetheless, the prevalence of caries in developed countries such as the United Kingdom (UK) remains significant, with marked inequalities in the distribution and impact of the condition, disproportionately affecting poorer and marginalised groups in society (Peres et al., 2019). The 2013 Child Dental Health Survey in England, Wales and Northern Ireland reported caries experience of 31%, 46%, 34% and 46% of 5-, 8-, 12- and 15-year-olds, with children from lower income families (and therefore being eligible for free school meals) being more likely to have dental disease than their peers (Pitts et al., 2015). Similarly, a more recent oral health survey of 5-year-olds in England identified a higher prevalence of caries experience in children from more deprived areas (34.3%) than those from less deprived areas (13.7%) (Public Health England, 2020). This is in line with previous studies that have shown the association between caries incidence and higher levels of socioeconomic deprivation (Schwendicke et al., 2015a, Slade and Sanders, 2017). Importantly, childhood deprivation may have a lasting influence on oral health. A birth cohort study conducted in Brazil found that experience of poverty in at least one stage of early life had an effect on caries experience later on in life, which was only partially mitigated by upward social mobility between childhood and adolescence (Peres et al., 2007).

Whilst the aforementioned decennial Child Dental Health Survey has been monitoring the trends in dental health amongst 5-, 8-, 12- and 15-year-olds in England, Wales and Northern Ireland since 1973, the oral health of 3-year-olds was investigated for the first time in 2013. The study identified that 12% of 3-year-olds in England had already experienced dental caries at this young age (Public Health England, 2014b). Similarly, a more recent oral health survey of 5-year-olds in England reported almost a quarter (23.4%) of those sampled had experience of dental caries (Public Health England, 2020). Interestingly, this figure had not decreased since the previous survey of 5-year-olds that was undertaken in 2017 (Public Health England, 2018). As a result, the continuing poor oral health of children has been the subject of intense media, public and professional attention (RCS England, 2015, BSPD, 2015).

2.1.1 Impacts of dental caries on children

Dental caries can significantly affect the lives of children and young people. Most of the impacts of caries originate from dental pain.

Dental pain

Toothache was experienced by 18% of 12-year-olds and 15% of 15-year-olds in England, Wales and Northern Ireland in the most recent Child Dental Health survey, with a disproportionately higher prevalence of dental pain reported by those from deprived backgrounds (Tsakos et al., 2015). These figures are lower than that of a previous UK study, which found 48% of children reported pain (Shepherd et al., 2012). Nonetheless, it should be noted that in this latter study, a ‘wobbly tooth’ was felt to be the source of pain for almost a third of the participants rather than a carious tooth, hence the true experience of dental pain may have been less than the results imply. An even higher prevalence of dental pain (69%) was reported by Goodwin and coworkers (2015b), in children awaiting a general anaesthetic (GA) for dental treatment, particularly during the fortnight prior to the appointment. It should be noted however, that this study used parental reports of pain, rather than gaining self-reports from the children themselves.

Parent proxies have been used widely in healthcare as a whole, yet they may not accurately reflect the child's own experiences (Zhou et al., 2008). Chambers and colleagues (1998) found inconsistencies in the reporting of pain by parents compared to children's self-reported pain, raising concerns that parental underestimation of pain could lead to inadequate pain control, or even under-treatment. Conversely, other studies have reported strong associations between children and parents (West et al., 1994, Miller, 1996). Nonetheless, as pain is a subjective phenomenon, self-report measures are now considered to be the gold standard, though it is acknowledged that these may not be suitable for use in pre- or non-verbal, pre-cooperative or non-English speaking children (Gaffney et al., 2003). The involvement of children in research in this way is discussed further in section 2.4.1.

Further afield, a range of self-report measures have been used to determine the prevalence of dental pain in children. In Thailand, Krisdapong and coworkers (2012) found 39% of 12-year-olds and 34% of 15-year-olds reported symptoms attributable to dental pain. Similarly, a study of 12-year-old Greek children reported a 37% prevalence of oral pain, which was attributable to a tooth in 60% of cases (Pau et al., 2007). Overall, it can be seen that pain is experienced by a significant proportion of children with caries.

Impacts relating to pain

Whilst pain is the most common feature of caries, there is a growing body of evidence on the further impacts relating to pain on children's daily lives (Schuch et al., 2015). These include time off school, difficulty sleeping, speaking eating and interference with everyday activities (Pau et al., 2007, Krisdapong et al., 2009).

A number of studies have highlighted links between dental caries and general health, with higher levels of untreated dental caries reported to be associated with reduced weight, poorer growth (Miller et al., 1982, Acs et al., 1999). More recently, Alkarimi and colleagues (2014) reported an inverse linear relationship between caries status and height, weight and body mass index in 6- to 8-year old children in Saudi Arabia. These findings relating to growth and development have led to concerns that dental caries could contribute to failure to thrive in children under the age of three years (Elice and Fields, 1990, Sheiham, 2006).

The impacts of caries on growth and development are complex, yet they are likely to be related in part to difficulty in eating (Sheiham, 2006). A number of the aforementioned studies identified difficulty eating to be an important impact on children (Shepherd et al., 1999, Ratnayake and Ekanayake, 2005, Pau et al., 2007, Krisdapong et al., 2012). This was investigated further through a qualitative study by Gilchrist et al. (2015), who concluded that some children may have restricted diets for lengthy periods of time, relating to difficulty eating hard foods, and getting food stuck in their teeth. Interestingly, the authors of this study reported that children adapt to these restrictions by trying to chew on one side of their mouth to minimize the discomfort.

The effects of these dietary limitations may extend further than just weight. A number of studies have investigated the association between iron-deficiency anaemia, a common form of malnutrition, and dental caries (World Health Oganisation, 2011). Rodd and Blankenstein (1995) identified a statistically significant increase in the number of teeth which required extraction amongst UK children with anaemia, compared to those without, indicating that caries severity may be greater in anaemic children.

Similar findings were reported from studies conducted in Canada and Israel, with positive associations between iron-deficiency anaemia and caries in children with an average age of 41 months, and 6 years respectively (Shaoul et al., 2012, Schroth et al., 2013). Interestingly, another similar study reported that most of their sample had a normal height, weight and body mass index despite the nutritional deficiency (Clarke et al., 2006). Nonetheless, there is some disagreement on whether routine haematological screening of children with caries is necessary or feasible (Rodd and Blankenstein, 1995, Clarke et al., 2006). Moreover, there are no published studies which have investigated whether this association in present in adolescents with caries, nor whether the deficiency is corrected after dental treatment.

Whilst the aforementioned studies suggest that dietary limitation and potential subsequent malnutrition may present as impacts due to dental pain from caries, some studies have associated caries with obesity in children. Hayden and coworkers (2013) conducted a systematic review and meta-analysis to investigate this relationship further, reporting a strong significant relationship with obesity according to body mass index (BMI) criteria, and children with caries in the permanent dentition. Moreover, a significant positive relationship was identified between caries and obesity in children from industrialised, but not newly-industrialised countries. The authors proposed that this finding may be related to socioeconomic status, where by higher Gross Domestic Product (GDP) nations may be more likely to have excessive consumption of products such as soft drinks (Hayden et al., 2013). Nonetheless, the authors identified that a range of non-standardised definitions of obesity had been used in the included studies, which led to some difficulties in analysis. A slightly earlier systematic review of 48 studies by Hooley and colleagues (2012) explained the converse findings relating to weight and dental caries by describing a non-linear relationship, whereby caries can be both positively and negatively associated with weight with variations across age groups and countries of different levels of socioeconomic development. For example, an inverse relationship between dental caries and BMI may be seen in developing countries with a high prevalence and severity of dental caries, whereby the caries may cause difficulty eating and affect weight gain. Alternatively, a positive relationship between caries and BMI may be found in more affluent countries, whereby there may be greater consumption of cariogenic and obesogenic food and drinks (Hooley et al., 2012).

External to the health impacts relating to dental pain, there is a growing body of evidence relating to both school performance and absenteeism amongst children with dental caries. A recent cross-sectional study of Thai primary school children demonstrated that poor oral health had a negative effect on school performance. Nonetheless, the authors acknowledge that socio-demographic economic status and social capital variables may be important confounding factors, and that deprivation may be the primary cause of low school performance (Kaewkamnerdpong and Krisdapong, 2018).

Jackson et al. (2011) used data collected from telephone interviews with parents about their various aspects of their child's health to determine the relationship between oral health status and school performance and attendance. The results suggest that children with poor oral health were nearly three times as likely to miss school because of dental pain than their counterparts with good oral health. School absences resulting from dental pain were also associated with poorer school performance, as was poor oral health overall. Whilst one cannot ignore the methodological weaknesses of this study, in particular the reliance on parental recall, and parent's perceptions of oral health and school performance, it would appear that there is a relationship between these factors.

2.1.2 Impacts on families

Due to the very nature of dental caries and the impacts on children detailed above, the condition is likely to have a substantial effect on family life.

A Brazilian cross-sectional study compared clinical data from 837 preschool children with data from questionnaires completed by their parents or carers relating to family functioning in relation to oral health conditions (Ribeiro et al., 2015). Although the prevalence of caries was high in this sample, only 9% of parents and carers reported absenteeism from work. A possible explanation for this could be that one of the parents was unemployed, though the study did not collect data on this. Higher figures of parental absenteeism from work were reported in a study of Sri Lankan 8-year-old children, where almost one third of parents reported having taken time off work to take their child to dental appointments (Ratnayake and Ekanayake, 2005).

Reported parental absenteeism from work was particularly high in a UK study. BaniHani and coworkers (2018) found that 46% of parents of children with caries had taken time off work due to problems with their child's teeth, mouth or jaw. Importantly, the same study reported that 20% of parents had been financially impacted by their child's dental problems or treatments. Whilst dental care for children in the UK is free at the point of delivery via the National Health Service (NHS), parents of children with caries may still be affected financially by lost working hours, the cost of transportation to appointments, parking fees and even childcare costs for other children. Furthermore, the cost of toothpaste and toothbrushes to prevent caries should not be ignored.

A further study exploring the impacts of oral and oro-facial conditions, including caries, malocclusion, cleft lip and/or palate on families, identified that 56% of a sample of 266 parents had to take time off work due to their child's condition (Locker et al., 2002). Of these, 5.6% reporting this occurring often or every day. Similarly, over 44% of the sample reported financial impacts due to their child's condition, with almost 14% of these parents reporting this occurring often or every day. Nonetheless, it is important to acknowledge that these may have been parents of children with more complex conditions than caries alone.

Further to the more practical impacts of childhood caries on parents, there is also evidence to suggest there may be psychological implications. Carvalho and colleagues (2018) collected data from 1313 child-and-parent dyads, utilizing children's clinical data pertaining to caries status, dental trauma and malocclusion alongside questionnaires relating to socio-demographic information and psychological variables completed by parents. They conducted a hierarchical regression analysis on the data, and reported that almost 24% of parents expressed feelings of guilt for their child's oral health problems. This feeling of guilt was significantly associated with the presence of caries, as well as the thought that the problems could have been avoided. This figure is lower than that reported from a previous study by the same research group, where 36% of parents reported feelings of guilt (Carvalho et al., 2012). However, this previous study was conducted with parents who were seeking dental care for their children, however, and these feelings of guilt are likely to have prompted this action.

Furthermore, in the aforementioned study of 8-year-old children in Sri Lanka, 52% of parents reported 'mental suffering' and 39% had disturbed sleep patterns as a result of their child's oral pain (Ratnayake and Ekanayake, 2005). Similarly, a wide range of negative emotions were experienced by parents of young children requiring a GA for treatment of dental caries in a qualitative study conducted in Canada (Amin et al., 2006). Parents felt troubled by the GA experience, expressing feelings of 'guilt', 'worry' and 'fear'.

Whilst the impacts of dental caries on non-affected siblings within the family have not been directly explored within the literature, this aspect has been indirectly investigated through use of the Family Impacts Scale (FIS). Developed by Locker et al. (2002), this parent-completed instrument aims to evaluate the impact of a child's oral condition on the family. Whilst the FIS asks whether the child with caries has required more attention, and whether the parent has had less time for other family members, it relies on parent's perceptions of these factors and cannot capture experiences from non-affected siblings themselves. Nonetheless, research using this tool has suggested that children with caries required more attention from parents, leaving them with less time for other family members (Abanto et al., 2012). This is an aspect which may require further investigation.

The FIS tool was employed in a more recent secondary analysis of data from the previously mentioned 2013 Child Oral Health Survey; a cross-sectional study of children in England, Wales and Northern Ireland (Abed et al., 2019). This identified high levels of guilt and stress amongst parents of children with more severe caries involving pain and infection. Further impacts on families included a restriction of normal activities, a requirement for parents to take time off work, and disturbed sleep for parents as they were caring for a child in pain (Abed et al., 2019).

The impacts on children, young people and their families vary within the literature, depending on the country in which the study was conducted. This is likely to be due to a number of factors, including access to healthcare services and the associated costs.

2.1.3 Impacts on society

The impacts of caries on children and their families have been outlined above, but it is also important to consider the wider impacts of childhood caries on society as a whole.

Regardless of whether healthcare services are funded publically, privately or through an insurance-based system, the use of resources to treat caries has a financial impact. Unfortunately, the costs and provision of routine treatment provided by general dental practitioners is not readily available in the UK. Similarly, treatment provided by specialist services is also difficult to obtain. Information relating to hospitalisation and emergency attendances is available from some countries, however, and is provided below to highlight the extent of the financial impacts of caries on society.

In England, approximately 41,558 children aged up to 16-years were admitted to hospital in 2018-2019 with a diagnosis of dental caries, making it the most common reason for children to require an admission with an estimated annual cost of £39 million to the NHS (NHS Digital, 2019). Similar findings are reflected further afield, whereby 215,073 children and young people under the age of 21 years attended emergency departments across the US in 2008 (Allareddy et al., 2014). Of these, the majority of attendees were diagnosed with dental caries, pulpal and periapical conditions. The mean charge to attendees each visit was $564, with the total emergency department charges across the US at $104.2 million. For those who were admitted to hospital following the emergency visit, the total hospitalisation charges across the US were $162 million. A total of 43% of the emergency visits were covered by Medicaid insurance, yet 32% of these children and young people were uninsured.

A ten-year study of dental admission patterns from 2000-2009 in Western Australian children aged 14 years and younger identified 43,937 children who had been hospitalised for an oral health-related condition (Alsharif et al., 2015). Dental caries was the most common diagnosis in non-indigenous patients, whilst pulpal and periapical conditions were most common in indigenous children. Overall hospitalisation costs, both public and private, were found to exceed AUS $92 million over the decade.

It is important to acknowledge that the impacts on society extend beyond the obvious financial aspects. As mentioned in the previous section, a significant proportion of parents have reported taking time off work due to their child's dental disease. Whilst this can present direct financial implications to the parent, particularly if they are paid hourly, or self-employed, it may also affect employers (Drummond et al., 1997). Small businesses in particular could suffer from parents lost working days, and may need to pay out for temporary replacement workers. The effects of lost productivity could also impact the wider economy. Unfortunately, data relating to these aspects specifically are not available.

2.1.4 The physiological process of caries

There are at least two major groups of bacteria that have been implicated in the caries disease process, namely the mutans streptococci and lactobacilli species (Featherstone, 2008). These bacteria reside in the oral cavity, specifically the plaque biofilm on the tooth surface, and are known to produce organic acids during their metabolism of fermentable carbohydrates. These acids diffuse into the enamel surface and begin to dissolve the tooth tissue; a process called *demineralisation*. If sustained over a period of months, or even years depending on a number of factors, the tooth surface will eventually become undermined and a cavity will form (Featherstone, 2008).

Symptoms may arise once the caries progresses beyond the outer enamel layer. Pain is the most common feature, which is understandable as the tooth has a complex neurovascular supply, known as the pulp, which extends into the dentine layer of the tooth. Pain can present as anything from a mild transient sensitivity to cold stimuli, through to a prolonged response to warm stimuli with spontaneous pain; the latter indicating the bacteria have infiltrated beyond the dentinal tubules to significantly infect the pulpal tissue (irreversible pulpitis). This inflammatory response from the pulpal tissue cannot be sustained and typically culminates in pulpal necrosis, whereby the vital tissue dies (Renton, 2011). The infection then progresses to the apical tissues and periodontium, which results in a dental abscess. An acute spreading infection can cause lymphadenopathy, and swelling around the jaw, or orbit, depending on the tooth involved (Renton, 2011). In children in particular, infection can spread rapidly. If left untreated, death can result, typically from a swelling sizeable enough to impede the airway, or from infection entering the brain to produce sepsis (Otto, 2017).

Dietary free sugars (sugars added to foods by manufacturers, plus those naturally present in honey, syrups, fruit juices and concentrates) are the most important risk factor for dental caries (Moynihan and Kelly, 2014, Moynihan, 2016). There is a wealth of evidence from many years of research to clarify the role of sugar in caries (Gustafsson et al., 1954, Fisher, 1968, Curzon and Curzon, 1979, Blinkhorn, 1982, Rugg-Gunn et al., 1984, Rugg-Gunn et al., 1986). Unfortunately these sugars are an all-too-common feature of modern diets, with the consumption of sugars worldwide having tripled over the past 50 years (FDI World Dental Federation, 2015).

Further to the presence of fermentable carbohydrates, bacterial colonisation of the oral cavity is necessary for the caries process to initiate. Unfortunately, the transmission of cariogenic bacteria from parents and caregivers to babies is near impossible to avoid, with cariogenic bacteria having been identified in babies mouths prior to the eruption of any teeth (Berkowitz, 2006, Featherstone, 2008). Whilst researchers have been trying to develop a vaccine against dental caries for many years, bacterial colonisation remains an important factor in the caries process that we are currently unable to control (Michalek et al., 2001). Notably, there are no specific ‘pathogens’ that fulfill Koch’s postulates, which were designed to establish a causative relationship between a microorganism and a disease (Neville et al., 2018, Twetman, 2018). For example, whilst Streptococcus mutans has been heavily implicated as a causative pathogen, it has also been found in individuals without evidence of dental caries (Phattarataratip et al., 2011). It is for these reasons that caries is not referred to as an infectious disease, and instead considered to be a non-communicable disease (Twetman, 2018).

Whilst cariogenic bacteria and fermentable carbohydrates are fundamental to this disease, there are also many other factors that play a role in modifying the caries process, such as saliva properties, genetic factors, as well as features of the tooth structure itself. Further factors that can be employed to control the caries process are discussed in section 2.1.5.

Within the dental profession, our knowledge of the disease has increased significantly over the years. A key development is the understanding that dental caries is a dynamic disease process, and hence *remineralisation* and arrest of the carious lesion is possible, providing certain environmental conditions are met (Pitts et al., 2017). As such, many dental interventions (section 2.1.5) are now focused on shifting the caries continuum in the direction of health through promoting remineralisation (Pitts et al., 2017). Crucially, we recognise that the placement of restorations (fillings) does not ‘fix’ the caries; it simply reduces the bacterial load by removing heavily infected tooth tissue, and repairs the cavity (Featherstone, 2008). Overall, the disease process will still continue in that individual, unless changes are made to the oral environment. As our knowledge advances, the role of the dental profession is gradually shifting towards a more preventive approach, hence reducing the necessity for reparative interventions (Birch et al., 2015).

2.1.5 Controlling and managing dental caries in children and young people

As mentioned in section 2.1.4, dental caries is a largely preventable condition, hence it is no surprise that the overwhelming majority of dental interventions for children and young people are centered around *caries control*. This term highlights that caries is a dynamic disease process that our interventions seek to control through promoting remineralisation and inhibiting demineralisation, though it is used interchangeably with *prevention*.

Caries control

The preventive effect of most of these interventions is derived from fluoride. Fluoride acts in a wide range of ways to control the caries process, but its ability to reduce enamel solubility, and promote remineralisation of dental tissues is considered to be most important (Featherstone, 1999). In previous years, ingested dietary fluoride supplements were recommended for caries prevention purposes, and indeed in some countries this practice continues (Holt et al., 1996, Rozier et al., 2010). However, concerns were raised regarding the risks of dental fluorosis and poor compliance with supplement regimes by families (Fayle et al., 2001). As such, current guidance for caries prevention in the United Kingdom no longer recommends supplements, and instead focuses on topical applications (SDCEP, 2010, Public Health England, 2017).

There are a multitude of different topical applications of fluoride, some of which are briefly described below. Many are delivered 'at-home' in the form of toothpastes, gels or mouth rinses, whilst others require involvement of dental professionals, such as fluoride varnishes.

'At home' fluoride delivery

As our knowledge of how to optimise fluoride delivery increases, the dental profession has adapted its messages to the public for oral care at home.

Toothbrushing with fluoridated toothpaste is the mainstay of fluoride delivery in the home. There is high quality evidence to show that increased frequency of fluoride exposure can increase the caries preventive effect, hence toothbrushing twice daily can reduce caries by a further 14% when compared to brushing once per day (Marinho et al., 2003). Furthermore, the addition of an alcohol-free mouthrinse at a separate time of day to toothbrushing can provide an additional opportunity for fluoride exposure, though current guidance recommends that these are not used by children aged 7 years and younger, as they may not be able to spit effectively (Marinho et al., 2004, SDCEP, 2010, Marinho et al., 2016, Public Health England, 2017).

Although secondary to the frequency of fluoride delivery, the concentration of fluoride is also important. The literature demonstrates that for every increase in 1000ppm fluoride in toothpaste, there is a further 8% reduction in caries (Marinho et al., 2003, Walsh et al., 2010). This has been adopted by current guidance, which advocates use of a toothpaste containing an appropriate concentration of fluoride; at least 1,000ppm for children under 3 years of age, and 1,450ppm for those over 3 years of age (Public Health England, 2017).

To investigate whether the amount of toothpaste is important in caries prevention, Ashley and coworkers (1999) studied 2,888 adolescents aged 15-16 years who had used 1,000ppm fluoridated toothpaste during previous three years. The authors found no significant difference in caries experience of those children that covered less than half brush head with toothpaste (547 children) compared to those who used more than half a brush head of toothpaste (2,341 children), suggesting that the amount of toothpaste used does not have an effect on efficacy (Ashley et al., 1999). Acknowledging these findings, alongside the potential adverse effects from fluoride ingestion, which are discussed further in the next section, national guidelines currently recommend the use of a smear of toothpaste only for under 3-year olds, and a pea-sized amount for young children aged 3 to 7 years (Public Health England, 2017).

Furthermore, Chesters and colleagues (1992) provided evidence to suggest that rinsing the mouth with large quantities of water following toothbrushing can reduce the caries-preventive effect of toothpaste. In a clinical trial involving 3005 Scottish adolescents, they found that those who rinsed their mouths with a beaker of water had more new carious lesions than those who used their hand to transfer water to their mouth. Dental professionals now advise their patients to spit excess toothpaste out after brushing, but not to rinse out their mouths (SDCEP, 2010, Public Health England, 2017).

Professionally-delivered interventions

Professionally-applied fluoride varnish containing 22,600ppm fluoride has a strong evidence base to support its use in caries prevention. A Cochrane systematic review conducted by Marinho and colleagues (2013) reported a reduction in carious tooth surfaces by over a third in both the primary and permanent dentitions, as evidenced by 10 and 13 studies respectively. Fluoride varnish application requires minimal cooperation from the child, and can be delivered by dentists, dental therapists and dental nurses with advanced skills training.

Another commonly used preventive approach, the professionally-applied fissure sealant, aims to isolate the most ‘at-risk’ sites of the tooth from bacterial substrate, acting as a physical barrier to caries formation (Ahovuo-Saloranta et al., 2017). A resin-based sealant is most commonly used, yet as its placement involves a few stages and hence requires greater co-operation from the patient than application of fluoride varnish. For children who are less comfortable in the dental setting, a less retentive, but easier-to-place glass ionomer sealant can be placed, with no significant increase in caries increment (Ahovuo-Saloranta et al., 2017). Fissure sealants can only be placed by dentists and dental therapists.

Since caries cannot progress once the biofilm (a community of microorganisms within an extracellular matrix) has been physically isolated from dietary substrate, a resin fissure sealant can also be placed over a non-cavitated carious lesion (Going et al., 1978, Mertz-Fairhurst et al., 1998, Deery, 2013). This approach has been shown to slow or arrest caries extending up to a third of the way into the dentine layer radiographically (Handelman et al., 1972, Deery, 2013). Nonetheless, there appears to be a reluctance to the widespread adoption of this approach potentially due to a lack of awareness amongst the dental community (Deery, 2017). Furthermore, practitioners may be concerned that this novel approach could generate litigation claims for negligence relating to a lack of caries removal and provision of a conventional 'filling'. It should also be considered that the current primary dental care contract provides little financial incentive for practitioners to provide non-restorative treatments such as this.

Community-based preventive programmes

Community-based preventive interventions are also available in some areas, predominantly to increase access to fluoride. Nursery and school-based toothbrushing initiatives are provided in Scotland and Wales, as part of the established Childsmile and Designed to Smile schemes respectively (Macpherson et al., 2010, Welsh Government, 2013). These comprehensive schemes also incorporate professionally-applied fluoride varnish, and coordinated care with health visiting services to support children to access a general dental practitioner (D2S, 2018, Childsmile, 2018). With robust monitoring and evaluation systems in place since inception, Childsmile in particular has been praised for reducing both the caries increment and inequalities in children accessing dental care in Scotland (Macpherson et al., 2013). Whilst plans are in place to commence a similar scheme within England, any current initiatives here are organised locally.

A recent community-based study, known as the Seal or Varnish trial, had two parallel arms providing interventions using mobile dental clinics in schools located in deprived areas of Wales. One arm applied resin fissure sealants to first permanent molars, which were maintained at six-monthly intervals, whilst the other arm applied fluoride varnish at baseline and six-monthly intervals. Interestingly, the findings indicated a similar caries preventive effect from the two interventions at 36 months (Chestnutt et al., 2017). Importantly, the Seal or Varnish trial also found the fluoride arm of the study to be more cost-effective than the fissure sealant arm within their study population (Chestnutt et al., 2017). Nonetheless, it should be noted that the longer-term effectiveness of fluoride varnish and fissure sealants as community-based interventions has not yet been investigated.

Dietary advice

The role of dietary free sugars as a modifiable risk factor for caries was highlighted in section 2.1.4. As children themselves have limited control over their diet, general dental practitioners are advised to provide dietary advice tailored to both the child and parents, often following some form of analysis of what the child consumed over the previous day(s) (Public Health England, 2017). Current recommendations arising from a systematic review suggest that dietary free sugars should form less than 5% of total energy intake to reduce caries incidence (Moynihan and Kelly, 2014, Moynihan, 2016). However, quantity of sugar intake is only one of many dietary factors involved in caries development; frequency, duration in the mouth and the timing of consumption also have key roles. Interestingly, in contrast to the emphasis of the literature, a qualitative study by Arheiam and colleagues found that dentists paid little attention to the amount of sugar consumed by patients when deciding what advice to provide, and instead focused on the other diet-related factors (Arheiam et al., 2016).

Threlfall and coworkers (2007) identified that dentists typically provided dietary advice in the form of a short educative talk, rarely using props or visual aids. Furthermore, they identified that dentists were more likely to provide advice if the child had caries, or if the parents seemed motivated. This is likely to be a logical approach, acknowledging that most parents are not immediately receptive to new information and prepared to change well-established behavior patterns (Weinstein et al., 2004). Motivational interviewing, a technique focused on moving people from inaction to action, is increasingly used within healthcare, particularly in difficult-to-change, problem behaviours (Prochaska et al., 2008). Having shown promise in improving children’s oral health in the United States, the recent Dental RECUR trial in the UK aimed to assess whether dental nurse-delivered motivational interviewing on both dietary and non-dietary factors could be used to avoid recurrence of caries in children who previously had a primary tooth extracted (Weinstein et al., 2004, Pine et al., 2015). The results indicated that this single low-cost intervention was successful in significantly reducing the risk of new caries experience in high caries-risk children (Pine et al., 2020). Nonetheless, further research is required to determine whether this approach would have similar success in preventing initial caries experience, or caries progression.

Upstream preventive approaches

As previously established, there are a wide range of disparities in oral health, with children from deprived backgrounds experiencing disproportionately more caries. These inequalities encompass more than simply the differences between the rich and the poor; they are caused by the broad conditions in which people are born, grow, live, work and age (Watt, 2012). These conditions are commonly referred to as the social determinants of health. Oral health inequalities cannot be tackled through so-called 'downstream' approaches of treating caries and providing clinical prevention alone (ICOHIRP, 2015). In fact, as many families are unable to access healthcare services, this approach could serve to widen oral health inequalities. Whilst downstream approaches are important and necessary, they should be combined with upstream actions, such as healthy public policies and legislation.

Water fluoridation is a key upstream preventive intervention to optimise fluoride availability. A systematic review, commonly referred to as the York review, conducted by McDonagh and colleagues (2000) demonstrated the effect of water fluoridation in reducing caries incidence. The authors surmised that a median of six people would need to drink fluoridation water for one person to be rendered caries-free. A subsequent Cochrane systematic review reported that water fluoridation can increase the number children without caries by 15% (Iheozor-Ejiofor et al., 2015).

Despite the impressive outcomes, only 10% of the UK currently has optimally fluoridated water, and there remains a great deal of opposition to water fluoridation. Concerns exist regarding the potential for adverse effects from ingesting fluoride. Whilst there is no clear evidence to suggest that 0.7 to 1.0ppm water fluoridation could pose a health hazard, the aforementioned Cochrane review estimated that fluoridation at 0.7ppm could result in dental fluorosis of aesthetic concern in 12% of participants (Iheozor-Ejiofor et al., 2015).

Nonetheless, a high risk of bias was observed in over 97% of the studies involved in this assessment. Importantly, many studies of fluorosis involve a clinical examination of the anterior dentition only, without taking a thorough history and assessing the posterior dentition. As a result, potential alternative diagnoses that can produce a similar pattern of enamel opacities, such as amelogenesis imperfecta and molar-incisor hypomineralisation, are not considered.

One recent upstream approach in the UK was the introduction of a sugar tax, for both dental and general health purposes, following in the footsteps of countries including Mexico, France, Hungary and Finland (BDJ, 2018). A two-tier levy on soft drinks manufacturers has been implemented for drinks with a sugar content over 5g per 100ml, and over 8g per 100ml respectively (BDJ, 2018). Further to its role in dental caries, sugar is a major culprit behind the rise in obesity, and obesity-related health problems. Research from Cancer Research UK suggested that 20% tax of sugar-sweetened beverages could prevent 3.7 million people from becoming obese by the year 2025 (UK Health Forum and Cancer Research UK, 2016). The authors estimated that this could reduce obesity prevalence by 5% overall, saving the NHS £10 million in the year 2025 alone. This is in concurrence with other evidence which suggests that taxes on sugar-sweetened foods and drinks can improve health in people from lower socio-economic groups in particular, as they are most affected by the price increases (Eyles et al., 2012). Schwendicke and colleagues (2016) used a model-based approach to investigate the impact of a 20% sugary drinks tax on oral and general health in the context of Germany. They reported the effect to be a reduction in consumption of carbonated drinks and a reduction in net caries increments at the population level.

Whilst it is no surprise that soft drinks manufacturers have criticised this approach, some healthcare bodies feel the soft drinks tax does not go far enough (BDJ, 2018). Others are concerned that the levy does not apply to fruit juices, some of which have a greater sugar content than carbonated drinks (Boulton et al., 2016). Nonetheless, many campaigners welcome the new tax and have turned their attention towards gaining a ban of sales of high-caffeine and high-sugar energy drinks to children and young people (Coghlan, 2018).

Further to the effects of marketing on influencing both children and adult's food choices, Public Health England have called for a significant reduction in advertising of 'less healthy' foods to children across all media platforms (Public Health England, 2015). Whilst the UK media regulatory body Ofcom has restricted advertising of these foods during children's programmes since 2007, with a reported reduction in children's exposure to these adverts by 37%, research has suggested that the thresholds still enable products with relatively high levels of sugar, fat and salt to be advertised to children (Ofcom, 2010, Adams et al., 2012, POST, 2016). As such, the National Institute for Health and Care Excellence (NICE) has advocated for a 9pm watershed for advertising of products with a high sugar, fat and salt content (NICE, 2010).

This section has outlined a range of evidence-based interventions to control or prevent dental caries in children. Nonetheless, the cost-effectiveness of many of these interventions has not been widely investigated. This is discussed further in Chapter 4.

Managing established caries

Once the caries process has taken effect, there are a variety of treatment options available to restore the tooth, such as fillings and crowns. Nonetheless, the evidence base remains somewhat conflicted regarding which material and technique to use, particularly when considering the primary dentition.

Following introduction of the 2017 European Union legislation on mercury, the Scottish Dental Clinical Effectiveness Programme (SDCEP) released guidance for practitioners in limiting their use of amalgam in children, as part of a wider amalgam phase-down (SDCEP, 2018). Adhesive restorations have long been the material of choice for restoring the anterior dentition, yet reliance on these materials for restoring the posterior teeth is now set to increase.

Glass ionomer cements (GIC) are commonly used in general dental practice to restore carious primary molars, predominantly due to their ease of placement (Milsom et al., 2002b). There are two main types; the conventional GIC, which sets via acid-base reaction, and the resin-modified GIC (RMGIC) which comprises an initial polymerisation reaction to start the setting process (Chadwick and Evans, 2007). A systematic review comparing the use of the two variants advised that conventional GIC should not be used to restore two-surface cavities, reporting failure rates of up to 60% (Chadwick and Evans, 2007). Better outcomes were reported for RMGIC, particularly when placed under careful moisture control using a rubber dam (Chadwick and Evans, 2007, Dias et al., 2018). Nonetheless, the very nature of the rubber dam requires the use of local anaesthetic, which is not universally provided for the restoration in itself, despite evidence to suggest that its use alone could improve a restoration's survival rate (Kilpatrick, 1993). This is likely to be due to the apprehension of dental practitioners in administering local anaesthetic for children (Rasmussen et al., 2005). Nonetheless, there is evidence to suggest that once a rubber dam is in place, both children and the treating dentist find treatment less stressful (Ammann et al., 2013).

Preformed metal crowns (PMCs) are considered to be the gold standard for restoring primary teeth with multi-surface caries (Fayle, 1999, Fayle et al., 2001, Kindelan et al., 2008). Traditionally PMCs were placed under local anaesthetic, following caries removal and preparation of the tooth. However, more recently a biological approach to PMC placement was introduced, known as the Hall Technique. Based upon the aforementioned principal that caries cannot progress once the biofilm (a community of microorganisms within an extracellular matrix) has been physically isolated from dietary substrate, the Hall Technique does not involve caries removal or tooth preparation, and hence does not require local anaesthetic (Innes et al., 2017). Over the past decade, the evidence-base for the Hall Technique has rapidly grown. A number of randomised controlled trials have reported particularly strong results, culminating in a 97% success rate at 5 years; significantly outperforming standard restorations (Innes et al., 2011, Santamaria et al., 2014). Whilst PMCs seem to be well-accepted by children, it should be noted that parents may have differing views. In a service evaluation by Bell and colleagues (2010), the majority of parents had no concerns about the appearance of the PMC, yet a small number had very strong objections to the aesthetics.

Prior to the introduction of the Hall Technique, preformed metal crowns were infrequently placed by general dental practitioners (GDPs) despite the recommendations from national guidelines (Threlfall et al., 2005). However, placement of preformed metal crowns is now increasing amongst GDPs, particularly in Scotland, where almost half of GDPs surveyed reported that they were routinely using the Hall Technique in practice (Dean et al., 2011).

This is likely to increase further as the Hall Technique is incorporated in the undergraduate dental curriculum and taught to dental students in all UK dental schools (Innes et al., 2017).

Interestingly, a number of retrospective studies of general dental services have suggested that the majority of carious primary teeth may exfoliate naturally without symptoms arising, irrespective of whether they were filled or not (Tickle et al., 1999, Tickle et al., 2002, Levine et al., 2002). Yet a similar study found pain and its sequelae to be common findings, with almost half of a sample of 677 children receiving extractions or antibiotics due to symptoms (Milsom et al., 2002a). Furthermore, Stephenson and colleagues (2010) analysed cohort data from 5,168 carious primary molars and reported that provision of treatment doubled the survival rates for these teeth. Understandably, these conflicting findings have raised questions surrounding the need to restore carious primary teeth, and if appropriate, how this should be achieved (Milsom et al., 2003). The FiCTION (Filling Children's Teeth: Indicated Or Not?) trial was established to investigate the clinical and cost-effectiveness of the provision of conventional restorations when compared to both the aforementioned biological approach, and a non-restorative best practice prevention arm in primary care (Innes et al., 2013b). The three-year trial concluded that there was no overall difference in terms of experience of or episodes of dental pain and/or sepsis, with best practice prevention being the least costly intervention (Maguire et al., 2020). Nonetheless, it must be acknowledged that less than half of the participants in the study had radiographs taken at baseline, hence it is not possible to determine how extensive the caries was when the teeth were restored. It is possible that the caries had already reached the pulp, in which case pain and subsequent sepsis may have been inevitable, regardless of which intervention was provided. The authors acknowledged that this may explain the high levels of pain and sepsis experienced in all three arms of the trial (Maguire et al., 2020). Unfortunately the lack of any overall differences between the three arms suggests that the debate over the restoration of carious primary molars is likely to persist.

More recently, there has been rapidly growing interest in the use of more minimally invasive approaches to managing caries. One intervention that has attracted much attention is silver diamine fluoride (SDF). This solution of silver, ammonia and fluoride has a cariostatic action, which appears to be predominantly due to the ability of silver to inhibit development of the cariogenic biofilm (Chibinski et al., 2017, Burgess and Vaghela, 2018). The evidence base supports the use of SDF to arrest caries in the primary dentition, with a recent systematic review reporting arrest in 81% of teeth treated with the solution (Gao et al., 2016). There is also a growing evidence base to demonstrate the ability of SDF to prevent caries, reduce dentine hypersensitivity and arrest caries in the permanent dentition, though further research is necessary to determine how SDF compares to alternative interventions such as fluoride varnish and fissure sealants (Braga et al., 2009, Castillo et al., 2011, Monse et al., 2012, Llodra et al., 2016, Oliveira et al., 2019). SDF has only recently started to be used in the UK, with treatment predominantly being provided within secondary care. This is likely to be due to a number of factors, including the limited availability of inexpensive SDF formulations, regulatory issues pertaining to its use 'off-label', and concerns regarding a key side effect; namely the production of black staining (Timms et al., 2020).

The options for restoring the tooth are more invasive and more limited once the caries progresses to the pulp. At this stage, pulp therapy or endodontic (root canal) treatment can be carried out in primary and permanent teeth respectively, yet if the caries has caused extensive destruction to the tooth, extraction may be the only viable option.

There is a growing evidence base for a number of interventions to restore the carious dentition, yet there is limited evidence to establish the cost-effectiveness of these interventions. This is discussed further in Chapter 4.

Implications of dental anxiety

Dental anxiety is particularly prevalent amongst children and young people, with 76% of 12-year-olds reporting moderate dental anxiety (HSCIC, 2015). Dental anxiety alone has significant impacts on the lives of children and young people (Luoto et al., 2009). Many dentally anxious children attend the dentist irregularly, or only when symptoms arise, and as a result, tend to suffer from increased caries morbidity (Nuttall et al., 2008, Menezes Abreu et al., 2011, Soares et al., 2015). Treatment for these children can be difficult for general dental practitioners to provide, and hence many are referred to sedation clinics or secondary care services for treatment with pharmacological adjuncts (Ryan et al., 2010). Typically, this takes the form of inhalation sedation with nitrous oxide and oxygen, or intravenous sedation with either midazolam or propofol as the primary sedative agent (Hosey, 2002). However, for many of these children, their age, level of co-operation, the severity of their anxiety or the extent of the treatment required may preclude the use of conscious sedation. These children often require treatment to be provided under general anaesthetic. Further research is required to compare the morbidity and cost associated with both sedation and general anaesthetic in children and young people (Ashley et al., 2015).

The evidence base for more novel approaches aimed at tackling the underlying causes of anxiety is also becoming established, with emerging research into different types of Cognitive Behavioural Therapy, including psychologist-delivered, internet-based and self-help formats (Shahnavaz et al., 2016, Marshman et al., 2016, Shahnavaz et al., 2018, Bux et al., 2019). Nonetheless, the economic aspects of these interventions have not yet been investigated.

2.1.6 The evidence base for interventions for dental caries

Rigorous and pertinent scientific research is a necessity for evidence-based dentistry, which has been defined as an approach to oral health care that requires the judicious integration of systematic assessments of clinically relevant scientific evidence relating to the patients oral and medical condition and history, together with the dentist's clinical expertise and the patient's treatment needs and preferences (American Dental Association, 2001). The evidence base for many of the interventions detailed above is well-developed, with multiple Cochrane reviews and high quality trials reporting their efficacy and effectiveness.

A wide range of outcome measures are used in these studies, yet the criteria described in the World Health Organisation (WHO) manual 'Oral Health Surveys: Basic Methods' are the most frequently used standard for caries detection (World Health Organisation, 2013). These criteria incorporate the use of the DMFT index as a means of reporting outcomes relating to caries experience. This index has been used for decades in epidemiological and interventional research, combining the number of decayed (D), missing (M), and filled (F) teeth (T) in an individuals' mouth into an objective measure (Klein and Palmer, 1940). The lower case format (dmft) refers to the primary dentition whilst the upper case (DMFT) denotes the permanent dentition. A range of variations on this are also widely used, such as the shortened DFT version, or DMFS, where the S refers to surfaces (Levey et al., 2017). Whilst measuring caries experience in this way is of great importance for research purposes, alone this approach cannot measure the wider impacts of caries on children's health and quality of life.

2.2 Caries and quality of life

2.2.1 Health, quality of life and health-related quality of life

Health

The attainment of health was previously considered by many to be achieved purely through the diagnosis and treatment of disease. Whilst this approach is still regarded by a number of societies and organisations today, its limitations are now widely acknowledged. As a result of these limitations, the World Health Organisation (WHO) adopted the biopsychosocial approach to healthcare in preference to the former model (World Health Organisation, 2002), with multiple other prominent bodies following suit. First proposed in 1977, the biopsychosocial approach considers the wider consequences of a health condition, integrating the biological and physiological features with the impacts they may have on participation within society (Engel, 1977). This is reflected in the first principle of the WHO constitution, defining health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (World Health Organisation, 2005). Nonetheless, some would disagree with the inclusion of social well-being as part of 'health', and instead feel it should contribute to the bigger picture of 'quality of life' (Torrance, 1987).

Quality of Life

Quality of life (QoL) is considered a broad and multi-dimensional concept, encompassing health and social well-being, alongside more diverse non-health domains, such as the economic, political, cultural, environmental, aesthetic, and even spiritual aspects of a person's life (Torrance, 1987).

Numerous and varied definitions of QoL exist, and hence researchers have endeavored to categorise the different approaches. Ferrans arranged a large selection of QoL definitions into five main themes: (1) normal life, (2) happiness/satisfaction, (3) achievement of personal goals, (4) social utility, and (5) natural capacity (Ferrans, 1990). The headings of these categories alone clearly highlight the subjective nature of the definitions, as many are based upon an individual’s own perception of aspects of their life (Karimi and Brazier, 2016). The subjectivity of these definitions could pose some difficulties for researchers aiming to quantify or measure QoL (Karimi and Brazier, 2016).

Whilst a definition cannot be agreed upon, it should be acknowledged that the concept of QoL may not be wholly applicable to some purely clinical research. Most clinical researchers are typically interested in the aspects of quality of life that relate specifically to health, thus introducing the notion of health-related quality of life (Wilson and Cleary, 1995).

Health-related quality of life

As for QoL, multiple definitions of health-related quality of life (HRQoL) exist. Many describe HRQoL as relating to all the aspects of QoL affected by health. However, critics would argue that all aspects of QoL, including education and politics, are in some way, even indirectly, related to health, making it difficult to distinguish HRQoL from QoL (Karimi and Brazier, 2016). Similarly, some definitions of HRQoL are thought to closely reflect the definition of health (Guyatt et al., 1993). However, whilst determining an individual’s health status allows the assessment of the physical and psychological symptoms, disability and social dysfunction, it does not enable information relating to the impact of these factors on the individual’s life to be captured. As seen in Wilson and Cleary’s commonly accepted model (**Error! Reference source not found.**), there are many components to HRQoL and hence its measurement can be complex.

2.2.2 Measuring health-related quality of life

There are a multitude of different tools designed to measure HRQoL, varying from a single global question, to an instrument multiple questions covering numerous domains, or dimensions. These dimensions contain questions or items relating to a similar experience of behavior being measured, such as mobility, or self-care (Guyatt et al., 1993).

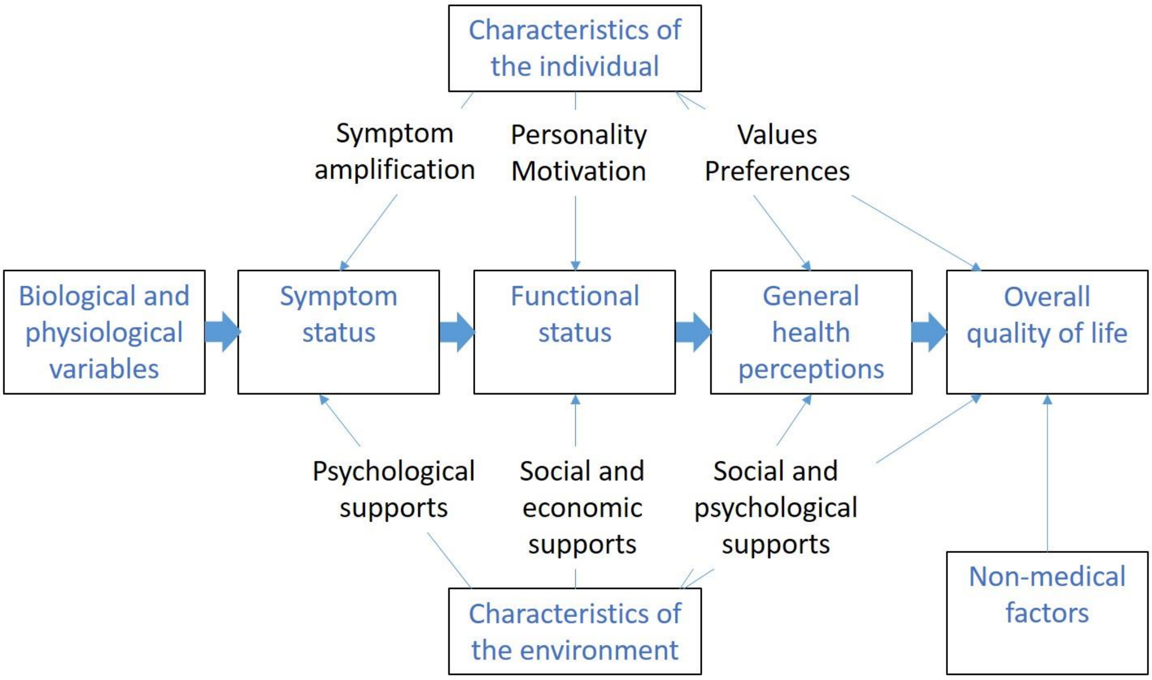


Figure 2-1: A conceptual model of health-related quality of life, proposed by Wilson and Cleary (1995)

Measures of HRQoL can be generic and applicable to a wide range of healthcare conditions. Nonetheless, some generic measures of HRQoL can be considered to be inappropriate or lacking in sensitivity to adequately describe the impacts of a certain condition. As a result, many measures of HRQoL have been designed specifically for use in particular healthcare conditions, such as the Overactive Bladder Questionnaire (Coyne et al., 2002). Measures can also be developed specifically to assess HRQoL within a certain population, such as children, or carers. These can also be either generic or condition-specific.

Most measures of HRQoL use a summative scoring approach. The summative scoring system gives equal weighting to all items and each response option of the same severity or frequency level, yet within a measure of HRQoL, there are likely to be items which are considered to have a greater impact on HRQoL than others. Furthermore, the response options within a measure may not be equal in terms of their impact on the overall HRQoL. Moreover, in summative scoring systems, the total score can be greatly influenced if there are multiple items that cover the same domain.

One way to address this issue is to assign a weighting to each item according to their relative impact on HRQoL, or utility. When a HRQoL has preference weights such as these, it is known as a preference-based measure (Drummond et al., 2015). One such example is the widely used generic measure known as EQ-5D (Rabin and de Charro, 2001), which is discussed in further detail amongst other preference-based measures in section 2.6. An alternative type of scoring system provides scores based upon psychometric properties such as item-response theory. Some measures have available multiple different scoring systems, for example the SF-36 has summative scores, domain scores and utility values (Brazier et al., 2002).

2.2.3 Oral health-related quality of life

As detailed above, dental caries can have impacts extending beyond the mouth, to affect a child's general health. There are a number of ways in which the impacts and benefits of dental disease can be determined, though the concept of oral health-related quality of life (OHRQoL) provides a holistic and standardised option.

Locker and Allen (2007) proposed the following definition of OHRQoL, which is widely used:

"the impact of oral disease and disorders on aspects of everyday life that a patient or person values, that are of sufficient magnitude, in terms of frequency, severity or duration to affect their experience and perception of their life overall."

This notion has been used widely in children's oral health research, primarily measured through use of generic non-preference-based measures.

2.2.4 Measuring oral health-related quality of life in children

There is a growing body of evidence to show the effect of dental interventions in improving children's 'oral health-related quality of life' (Yawary et al., 2016, Abanto et al., 2016). As a result, there are an increasing number of instruments that have been developed to evaluate these changes.

Currently, the most widely used measure of paediatric OHRQoL is the Child Perceptions Questionnaire (CPQ), of which multiple versions exist which are intended for different age groups (Gilchrist et al., 2014). The Child Oral Impacts on Daily Performances Index (C-OIDP) and Child Oral Health Impact Profile (COHIP) are also increasing in popularity amongst researchers in this field.

Despite the number of measures of OHRQoL available, a systematic review by Gilchrist and colleagues (2014) identified that many have inherent limitations. Critically, some measures have sought minimal input from children and lay people in their development, and hence may not provide a true representation of what is important to them. As such they may be more accurately termed ‘measures of oral health status’ (Locker and Allen, 2007). A further concern is that these measures are not condition-specific and fail to capture impacts of caries. Furthermore, they were not developed to assess treatment-related changes and may lack the psychometric properties to do so (Gilchrist et al., 2014).

In view of the perceived need for a caries-specific measure of OHRQoL, which is sensitive to change, a condition- and population-specific measure was developed and validated for clinical use with children (Gilchrist, 2015, Gilchrist et al., 2015, Gilchrist et al., 2018). The 12-item Caries Impacts and Experiences Questionnaire for Children, known as CARIES-QC, seeks children’s assessment of the severity of their caries-related impacts, and is appropriate for use in 5-16 year-olds. This child-centred measure has been shown to have good face, content and construct validity, responsiveness and reliability. This instrument is intended to be employed in randomised controlled trials (RCT) as a patient-reported outcome measure (PROM) for the prevention and management of caries in children. Further consideration of CARIES-QC and alternative measures is provided in section 5.2.1.

To summarise, there are a range of approaches to measuring OHRQoL in children. Specifically, CARIES-QC may address some of the limitations of the other measures. As an improvement in OHRQoL is one of the main benefits of treatment, it should be incorporated into economic evaluation. However, it is important to note that paediatric oral health-specific measures such as CARIES-QC have limited use in economic evaluations of oral healthcare interventions as they are not preference-based.

2.3 Economic evaluations in healthcare

There are a number of resources involved in healthcare, including people, time, equipment, facilities and knowledge (Drummond et al., 2015). These resources are not limitless, regardless of whether the healthcare system is publicly funded or based on social or private insurance. It is not practicable for every intervention to be provided, and hence a decision must be made to determine which interventions are the best value for money. The organisation tasked with making these decisions must be provided with details surrounding the cost and consequences of all relevant alternative interventions to ensure their judgement is well-informed. As summarised by Morris and colleagues (2014), economic evaluations can provide decision-makers with the information required to enable them to:

- maximise the benefits from health care spending;
- overcome regional variations in access;
- contain costs and manage demand;
- provide bargaining power with suppliers of health care products

An economic evaluation is typically defined as *"the comparative analysis of alternative courses of action in terms of both their costs and consequences"* (Drummond et al., 2015). As shown in Table 2-1, any studies which fail to compare two or more alternative, such as a cost analysis, or an outcome description, cannot be considered a full economic evaluation. Likewise, if a study does not attempt to compare the cost and consequences of an intervention with an alternative, it does not meet the above definition, and can only be considered a partial evaluation.

Whilst partial evaluations can still provide an important contribution to knowledge in this field, they cannot be used to answer questions relating to efficiency; that is, the achievement of acceptable efficacy and efficiency with the most appropriate use of resources (Mackenzie and Dixon, 1995, Drummond et al., 2015). As seen in Table 2-1, there are a number of different types of full economic evaluation, the selection of which is often dependent on the research question, the viewpoint of the analyst, the context of the study and the nature of the consequences.

2.3.1 Types of full economic evaluation

As shown in Table 2-1, there are three accepted forms of full economic evaluation. Nonetheless, some would argue that a fourth type of full economic evaluation exists, in the form of the cost-minimisation analysis (CMA). For completeness, this is also discussed below.

Cost-benefit analysis

CBA is the broadest form of analysis, characterised by the measurement of all costs and benefits in monetary terms. Grounded in welfare economic theory, the CBA considers the wider social implications of an intervention on the economy as a whole, including the costs and consequences external to the healthcare sector (Parkin et al., 2015). Through comparing the incremental programme benefits with the incremental programme costs, the net social benefit of the programme can be identified (Drummond et al., 2015).

Table 2-1: Characteristics of different types of healthcare evaluations, from Drummond *et al* (2015)

Are both costs and consequences of the alternatives examined?

Is there comparison of two or more alternatives?	**No**		**Yes**
No	Examines only consequences	Examines only costs	
	Partial evaluation		**Partial evaluation**
	Outcome description	*Cost description*	*Cost outcome description*
Yes	**Partial evaluation**		**Full economic evaluation**
	Efficacy or effectiveness evaluation	*Cost analysis*	*Cost-effectiveness analysis (CEA)* *Cost-benefit analysis (CBA)* *Cost-utility analysis (CUA)*

A key strength of this approach is that a direct comparison of the costs and benefits is possible when they are measured in the same unit (money) (Morris et al., 2014). However, problems can arise due to the difficulty in measuring those consequences which are not easily expressed in monetary terms, which may result in a somewhat limited analysis (Drummond et al., 2015). Nonetheless, the willingness to pay (WTP) method is gaining in popularity amongst researchers, whereby individuals state the maximum amount of money they would be prepared to give up to secure a proposed programme (Johannesson and Jönsson, 1991).

Whilst the types of analyses described below seek to inform the decision-maker as to whether an intervention is cost-effective enough to fund within the constraints of a set financial budget, the CBA is unique in that it can determine whether the budget should be expanded to accommodate the new programme. This is due to the CBA considering the *opportunity costs,* that is, the value of the benefits achievable from a programme which have been forgone by selecting another programme, external to the health care sector (Drummond et al., 2015). As such, CBA can also be used to inform resource allocation across different sectors of the economy, not just within healthcare (Ratcliffe, 2000).

Cost-effectiveness analysis

Cost-effectiveness analysis, henceforth referred to as CEA, is the most common form of full economic evaluation. The consequences are measured in natural units specific to the intervention or programme under analysis, such as a reduction in dmft/DMFT or an increase in the number of teeth free of obvious caries, though sometimes a range of outcome measures are presented.

Whilst these physiologic units provide important information for clinicians, they often correlate poorly with daily functioning and wellbeing; the aspects that patients are most interested in (Guyatt et al., 1993). Furthermore, the measure of effectiveness can vary between, or even within fields of healthcare, and so CEA is unable to make comparisons in these situations (Drummond et al., 2015). Moreover, whilst CEA is popular amongst clinical researchers, many healthcare programmes have numerous objectives or outcomes. As such, it becomes increasingly important to consider assigning preference weighting or values to the outcomes (Drummond et al., 2015).

Cost-utility analysis

Whilst there are many similarities between the two approaches, CUA was developed to address some of the disadvantages of the CEA. The CUA compares the incremental cost of a programme, from a particular viewpoint, with the incremental health improvement attributable to the programme; the latter which is most commonly measured in terms of Quality Adjusted Life Years (QALYs) gained (Drummond et al., 2015). The QALY combines both the quality of life and length into a common unit of measure. Quality of life is indexed (anchored) a 1-0 scale where 1 is full health and 0 is dead and is multiplied by the number of years the health state is experienced for. Typically, the values for QoL are generated using preference-based measures such as EQ-5D. Preference-based measures (PBMs) are discussed further in section 2.5.

The results of a CUA are typically presented in terms of the cost per QALY gained, and hence can be compared both within and across different fields of healthcare (Drummond et al., 2015). Due to its wide applicability, CUA can be more useful to decision-makers than the CEA approach (NICE, 2012).

Whilst QALYs are the unit of measurement recommended for health technology assessment, there are some alternatives, namely the Disability Adjusted Life Year and the Healthy Years Equivalent, though these are less frequently used (NICE, 2018).

A dental alternative to QALYs was developed in the form of Quality Adjusted Tooth Years (QATYs) (Fyffe and Kay, 1992). The QATY measures the production of additional years of life (tooth-year) of each tooth adjusted for the quality of the tooth. An unrestored tooth has a utility of 1, whilst an extracted tooth has a utility of 0 in that year and subsequent years. Where a tooth has been restored in any way, it would have a utility of less than 1.

Whilst QATYs were generated alongside QALYs in the aforementioned Seal or Varnish trial, the authors acknowledged a number of limitations with the QATY (Chestnutt et al., 2017). Firstly, since Fyffe and Kay (1992) published utilities for four different tooth states, our knowledge of caries has increased substantially. Importantly, it is now understood that the disease is a dynamic process, moving between demineralisation and remineralisation (Featherstone, 2008). These utilities are unlikely to reflect the full range of dental states represented on the caries continuum. Furthermore, its applicability to the primary dentition has not yet been explored, nor has its comparability across different interventions

(Chestnutt et al., 2017). Moreover, the QATY may fail to acknowledge the impacts of caries beyond the tooth itself. These limitations currently preclude the wider use of the QATY as the primary means of measuring the benefits of dental interventions.

Cost-minimisation analysis

This is a specific type of analysis historically undertaken in a situation where there is no statistically significant difference between the consequences, or effectiveness, of the alternative interventions. As such, equivalence is assumed and the costs only are compared. Whilst this type of analysis is considered easier to conduct and interpret, it is now widely agreed amongst health economists that this approach should not be considered a full economic evaluation. Briggs and O'Brien (2001) criticised the CMA approach as the assumption of equivalence overlooks any uncertainty surrounding costs and consequences. Furthermore, the CMA approach fails to capture wider benefits of healthcare interventions that CUA would capture in relation to overall quality of life (Briggs and O'Brien, 2001, Dakin and Wordsworth, 2013). Claxton also argued that the intervention with the greatest net benefit should be adopted, irrespective of uncertainty or the statistical significance of the differences between consequences (Claxton, 1999, Dakin and Wordsworth, 2013). The only justifiable applications of CMA are limited to situations where near-identical technologies, such as drugs of the same pharmacologic class, are compared, or non-inferiority studies (Briggs and O'Brien, 2001, Drummond et al., 2015). Nonetheless, CMA studies continue to be published, with a high risk of bias identified, even in non-inferiority trials (Dakin and Wordsworth, 2013). Further reference to CMA can be found in Chapter 4.

Cost-consequence analysis

A cost-consequence analysis (CCA) is descriptive partial economic evaluation, that provides disaggregated costs and a range of outcomes that are considered separately (Drummond et al., 2015). These are most useful in complex interventions involving multiple outcomes and perspectives, or public health interventions with both health and non-health outcomes (Brazier, 2017). CCAs provide an easily interpretable overview for decision-makers, and are not restricted to any particular viewpoint. Nonetheless, CCAs have limited generalisability, and economic rationality is not guaranteed through this approach (Gage et al., 2006).

2.3.2 Economic evaluations in child oral health research

As described in section 2.1.5, there are a range of evidence-based interventions available to prevent and manage dental caries in children. Nonetheless, the cost-effectiveness of such interventions is known to be under-researched. An analysis of oral health promotion approaches, commissioned by NICE, confirms that there is a high level of uncertainty about the cost-effectiveness of preventive programmes due to a paucity of published economic evaluations with sound methodology (Lord et al., 2015). This deficiency is further acknowledged by commissioners of dental services and hence Local Authorities have been advised to consider economic evaluation of caries-prevention programmes (PHE, 2014). Moreover, Public Health England reported a lack of evidence regarding the cost-effectiveness of interventions to improve the oral health of children aged 0-5 years in their recently commissioned rapid review (York Health Economics Consortium, 2016). Indeed, of the few studies that did contribute to their review, many were conducted in Scotland and hence the authors acknowledged the difficulty in generalising these findings to England.

Chapter 4 details a systematic review of the quantity, quality and scope of existing published economic evaluations of oral healthcare interventions for children, and considers the reasons behind the deficiency in research in this field.

2.4 Healthcare services for children and young people

Children and young people comprise a significant proportion of healthcare service users, as evidenced by the findings of the General Household Survey, an inter-departmental multi-purpose survey of people living in private households in Great Britain (Office for National Statistics, 2009). Whilst administration of the survey ceased in 2007, it highlights the trends in outpatient and casualty attendances and consultations with doctors amongst children and young people under the age of 14 years, with uptake of these services second only to the age groups over 65 years. Whilst this survey omitted specific enquiry into the use of dental services, the findings could be easily applicable to this field of healthcare, as demonstrated by the aforementioned prevalence and impacts of dental disease.

Despite being such a large group of service-users, healthcare services for children and young people may not be considered ‘age-appropriate’. The National Health and Hospitals Reform Commission in Australia reported that failure to provide ‘youth-friendly’ services has led to reduced or delayed uptake of healthcare services amongst adolescents (National Health and Hospitals Reform Commission, 2009). This raised concerns that young people would not benefit from available healthcare interventions, placing them at greater risk of poor health. Furthermore, when adolescents do access healthcare services, their compliance with medical treatment regimens has been reported to be significantly poorer than that of adults or younger children (Dolgin et al., 1986, Shaw, 2001).

Further to the direct risks to the health of young people, concerns surround the inappropriate allocation of resources to unsuitable healthcare interventions (Ratcliffe et al., 2016b). Wasted expenditure of resources in this way can take a substantial toll on the limited budget of a healthcare system overall.

The need to understand the emotional, educational and social needs of children and young people was highlighted by the charitable body Action for Sick Children, formerly the National Association for the Welfare of Children in Hospital (NAWCH) back in 1990 (NAWCH, 1990). Their report advised healthcare professionals to view adolescents as a ‘distinct consumer group’, with unique healthcare requirements as a result of the social pressures, psychological impacts and physiological changes occurring during their transition to adulthood. Since this time, understanding and respecting the needs of children and young people of different ages has been the premise of multiple publications since from various government groups and charitable bodies over the years (Hogg, 1996, Southall et al., 2000, NHS, 2004, Santos Pais and Bissell, 2006, Branclazzi, 2008). This has also translated to dental care, with plans for the development of ‘child-friendly’ services in Scotland (Scottish Executive, 2005).

One way to ensure paediatric healthcare services are ‘age-appropriate’, ‘youth-friendly’ and fit-for-purpose is to involve children and young people as key stakeholders. Further to their participation in the design and delivery of healthcare services and facilities, the involvement of children and young people in decisions relating to their own healthcare was emphasised at the UN Convention on the Rights of the Child (UNCRC) (United Nations, 2009). Moreover, involvement of children and young people in healthcare research provides an opportunity for them to provide their unique perspectives, which in turn may improve the quality and uptake of healthcare services for this group.

2.4.1 Involving children in research

The rights of children and young people to both participate in, and benefit from research were outlined at the UNCRC (United Nations, 2009). Yet up until the 1970s, the majority of research had little involvement of children. Marshman and Hall (2008) clarify the distinction between research *on* children and research *with* children; the former considering children as objects of research working on the assumption that the knowledge of adults is superior to that of children, whilst the latter considers children as competent and active participants. The move towards 'child-centred research' is a result of the changing position of children in healthcare and society overall (Marshman and Hall, 2008). This has been reflected in child oral health research, with a reported increase in the involvement of children as active participants over the past decade (Marshman et al., 2015).

There is now persuasive evidence that children are able to report on their own health, which and so health researchers are encouraged to consider children as active participants in economic evaluation (United Nations, 2009, Stevens, 2010). The involvement of children and young people in economic evaluations is considered to be both feasible and reliable, from the development of a classification system for a preference-based measure, to completion of a valuation survey (Stevens, 2010, Stevens, 2015). Chapter 4 further explores the involvement of children and young people in existing economic evaluations of child oral health research.

2.5 Generating QALYs

As mentioned earlier, QALYs combine both the quality and length of life into a common unit of measurement. The quality of life component represents the value that is placed on different states of health (Brazier and Longworth, 2011). This is anchored on a scale of 0 to 1, whereby 0 represents death, or a health state considered to be as bad as being dead, whilst 1 represents full health. Therefore 1 QALY represents a year in full health. These are the recommended measure of health outcomes for health technology assessments submitted to NICE in England and Wales (NICE, 2013). Whilst PBMs are not the only method of generating the QALY, they are certainly the most widely used approach.

2.5.1 Preference-based measures

PBMs are pre-validated questionnaires that incorporate a multi-attribute classification system and a set of preference weights or scores. PBMs are typically designed so that they can generate utility values for health states experienced by people over time, thus showing both treatment impact and disease progression. PBMs are used to determine the quality adjustment component of the QALY, which is multiplied by the number of years the health state is experienced for, to generate the QALY. These differ from non-preference-based measures, such as CARIES-QC, which have a simple summative scoring algorithm that assumes that equal intervals between response levels and attributes are of equal importance (Ratcliffe et al., 2011).

The classification system for a PBM contains dimensions pertinent to the population (and condition if necessary) and the attributes within each dimension (Goodwin and Green, 2016). In order to calculate the QALY, a preference-weight or utility value must then be assigned to every health state described by the classification system (Goodwin and Green, 2016). The preference weights are elicited in studies where tasks such as time trade-off (TTO) or standard gamble (SG) are presented to respondents to facilitate them to express their preference for a particular health state over another (Tolley, 2009, Versteegh et al., 2012). Preference elicitation techniques are discussed in greater detail in section 2.6.

There are a wide range of PBMs available, some of which are specific to one condition, and others which are generic.

Generic preference-based measures

Generic preference-based measures of health, or GPBMs, are suitable for use across a range of different healthcare conditions. A GPBM, known as the EQ-5D, is currently recommended for use in health technology assessment in adult populations in the UK (Rabin and de Charro, 2001, NICE, 2013). The classification system for the EQ-5D is comprised of five dimensions of health, namely mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression. The EQ-5D has three levels of severity for each dimension, however a more recent adaptation of the instrument, the EQ-5D-5L, introduced five levels (Norman et al., 2013, NICE, 2017). The EQ-5D-5L was designed to be more discriminative than the three-level version (NICE, 2017).

Nonetheless, it is acknowledged that the EQ-5D and other adult measures are not suitable for self-completion by children (NICE, 2013). The dimensions used in adult measures may not be pertinent to children, and the response scales, wording and format may be inappropriate (Stevens, 2010). Instead, a validated measure, specifically developed for children is recommended when administering a measure of HRQoL to children.

There are a growing number of PBMs for use in paediatric populations. The Health Utilities Index mark 2 (HUI2) was initially designed for use in paediatric oncology, but it subsequently evolved to become the first paediatric GPBM (Torrance et al., 1996). The classification system for this measure was based upon the findings of a previous study whereby potential attributes were identified from the literature and ranked by lay adults in order of perceived importance (Torrance et al., 1996). The six most important attributes were selected for the HUI2, and a seventh attribute relating to fertility was added, as this was felt to be of particular relevance for children with cancer. The HUI2 developers then obtained preference weights from adults using two different cardinal preference elicitation techniques. It has since been translated into a number of languages and validated with populations in various countries (HUInc, 2018).

Whilst it continues to be used widely, there are a number of limitations to the HUI2. Firstly, the classification system was not developed with children, and hence the attributes selected may not hold particular importance with children. The very fact that the attributes were selected to reflect those of importance to paediatric oncology patients limits its ability to present itself as a generic measure. Furthermore, no qualitative work was undertaken to validate the classification system with children with other healthcare conditions. The language used within the PBM may also not be familiar or easy for children to understand. Importantly, whilst preference-elicitation by children has been the subject of debate, as the preferences for this measure were elicited by adults, they may not reflect the values of children. This is important as children and young people are likely to value their health differently to adults, given their social values, support structures, lifestyles, and experience (Moodie et al., 2010). Preference elicitation is discussed further in section 2.6.1.

The Health Utilities Index Mark 3 (HUI3) was developed to address some of the concerns surrounding the HUI2. It was intended to be applicable to both clinical and general populations and to have structural independence amongst the attributes, producing a more efficient descriptive system (Horsman et al., 2003). The HUI2 and HUI3 are independent but complementary descriptive systems (Chen and Ratcliffe, 2015), with both being deemed suitable for children aged five years and above, though with more response levels the HUI3 could be considered more complex. Interestingly, the developers do not recommend the combined 15 item HUI questionnaire for self-administration until a child is over 12 years of age, and instead suggest proxy assessment by a parent or carer (Horsman et al., 2003). Whilst the HUI3 is widely used in paediatric populations, the scoring algorithm was also derived from the general adult population (aged 16 years or above), hence it suffers the same disadvantages as the aforementioned HUI2 (Feeny et al., 2002).

Subsequently, the widely-used EQ-5D instrument was adapted to create a ‘youth’ version, the EQ-5D-Y (Wille et al., 2010). The classification system for this instrument was drawn from the same five dimensions as the adult version of the EQ-5D, but with rewording suitable for administration in children and adolescents, hence drawing similar criticism as the HUI2. Nonetheless, the developers have acknowledged that preference weights for this classification system should be drawn from an adolescent population (Kind et al., 2015, Dalziel et al., 2020).

A third paediatric GPBM is available in the form of the Assessment of Quality of Life – 6 Dimension (AQoL-6D) adolescent version (Moodie et al., 2010). Derived from the original adult measure (the AQoL-6D), the adolescent version uses preference weights elicited by adolescents. Nonetheless, the choice of preference-elicitation technique used with adolescents in this study (time trade-off) may not have been well-informed (see section 2.6.2). Regardless, there is little available literature reporting the use of the AQoL-6D adolescent version.

Of the four main GPBMs available for children, only one has involved children in the development of the classification system, namely the Child Health Utility – 9 Dimensions (CHU9D) (Stevens, 2010). Despite using adults to elicit preference-weights, the involvement of children in forming the classification system has led this measure to have greater content validity and relevance than alternatives. It has been shown to be a valid measure in a range of medical conditions, for those aged 7-17 years (Ratcliffe et al., 2012).

Application of generic measures in child oral health research

Of the aforementioned generic paediatric instruments, only the CHU9D has been applied to oral health. Foster-Page and colleagues were the first to use the CHU9D in dental research in an RCT investigating the efficacy of a caries-preventive intervention in 6 to 9-year olds in New Zealand. Whilst the CHU9D showed a difference in the hypothesised direction, this study provided evidence that the CHU9D is not responsive enough to detect changes in caries over time (Foster-Page et al., 2015).

Whilst there is no available evidence to show that the HUI2, HUI3, EQ-5D-Y or AQoL-6D would suffer the same limitations if applied to oral health research, this could be a realistic possibility given that a generic measure may not contain all important dimensions for oral health. In the situation where psychometric evidence shows that the EQ-5D is inappropriate, a condition-specific preference-based measure may be used (NICE, 2012).

Condition-specific PBMs

Condition-specific PBMs (CSPBMs) are developed to generate the quality adjustment component of the QALY specific to a healthcare condition, such as asthma. Whilst CSPBMs are targeted specifically at a population with a certain condition, some may also contain more generic dimensions (Rowen et al., 2017). Others can be unidimensional, containing a number of items which all relate to the same overarching dimension. In recent years there has been a surge in the development of CSPBMs in response to the increased need for economic evaluations in healthcare, and the relative drawbacks of GPBMs (Versteegh et al., 2012).

Because of the very nature of GPBMs in covering a broad range of impacts on HRQoL, they may in turn fail to capture more specific aspects of HRQoL that are important for certain patient groups (Rowen et al., 2017). This limitation can be addressed by CSPBMs, as the attributes within their classification system are likely to be more relevant for these patient groups. As a result, CSPBMs have been reported to have greater sensitivity and responsiveness than GPBMs, particularly in more milder states of a condition (Brazier et al., 2012b). However, it should be noted that even the attributes within a CSPBM may not be considered to be important and relevant across all patient groups with that condition. Rowen and coworkers (2017) provide the example of a CSPBM in cancer, where features may not be considered important to those with different types of cancer, for example solid tumours and blood cancers alike. Furthermore, due to their fundamental differences, concerns have been raised relating to the comparability of utilities derived from CSPBMs and GPBMs (Versteegh et al., 2012). This in turn has contributed to the wider argument of whether all QALYs can be considered as equal (Cookson et al., 2009).

Whilst there is no empirical guidance on when and how to apply CSPBMs, on account of the shortcomings described above, it has been suggested that their use should be limited to interventions where the use of a GPBM would be considered inappropriate (Versteegh et al., 2012, NICE, 2013, Rowen et al., 2017).

CSPBMs are not available for all healthcare conditions. Of relevance to this body of work, there are no reported CSPBMs in the field of paediatric oral health.

Developing condition-specific PBMs

When psychometric evidence indicates that a GPBM is inappropriate, a CSPBM may be used. There are two primary methods that can be used to develop a CSPBM.

Where no existing measure of HRQoL exists, a PBM can be developed ‘de novo’. The classification for a new PBM can be informed by the literature or derived from qualitative interviews, with application of statistical techniques where necessary (Goodwin and Green, 2016). However, where a suitable measure of HRQoL does exist, it may be possible to convert it into a PBM. A six stage process for this was proposed by Brazier and colleagues (2012b) and is shown in **Error! Reference source not found.**. This is the most common method for development of CSPBMs, and hence the possibility of adaptation of the CARIES-QC to a preference-based measure was considered during its development (Gilchrist, 2015, Goodwin and Green, 2016). One advantage of using this approach is that the psychometric properties of the existing measure of HRQoL are likely to have already been assessed, and there is likely to be evidence of previous validation (Rowen et al., 2017). Furthermore, the existing measure of HRQoL may have already been used in a number of studies, and hence utility values can be generated from existing datasets (Rowen et al., 2017).

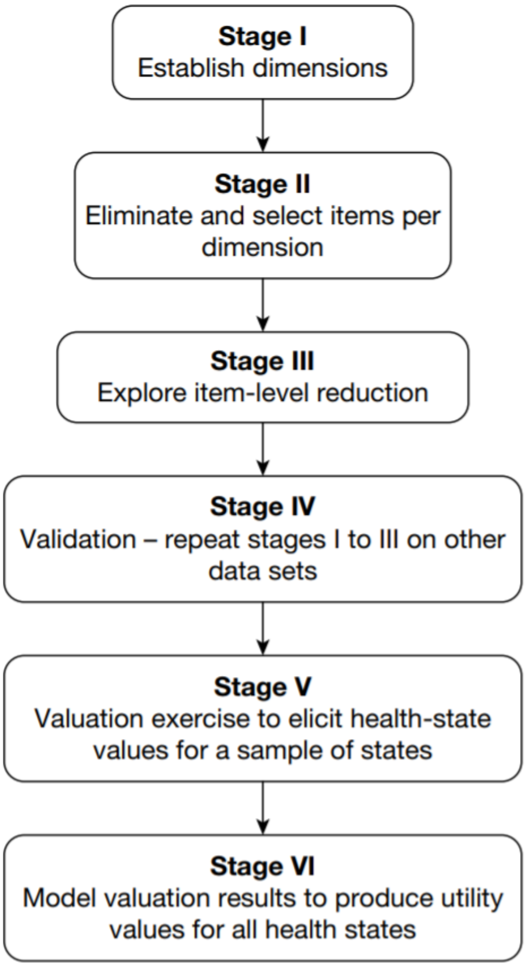


Figure 2-2: The six stages that can be used to derive a CSPBM from an existing non-preference-based measure of HRQoL, from Brazier et al. (2012b)

2.6 Considerations for paediatric preference-based measures

The significance of dental caries as a public health problem has been highlighted, as has the need for economic evaluations in this field. Furthermore, the lack of a preference-based measure suitable for use in children with caries has been identified. When considering the development of such a measure to meet this requirement, there are a number of factors to bear in mind. A recent review of paediatric and adolescent GPBMs highlighted gaps in the evidence regarding whose preferences should be used and the type of preference-elicitation techniques to be used (Rowen et al., 2020). These aspects are explored in greater detail below.

2.6.1 Whose preferences should be used?

The perspective from which preferences are provided is an important factor to consider when developing a measure. Current UK guidance advocates a societal approach, whereby preferences are obtained from a representative sample of the public (NICE, 2012). However, whilst there is little justification for estimating preferences from the literature or clinicians, there may be an argument for the use of patient preferences.

Primarily, as an individual who has actual experience of a health state, a patient can provide a more accurate assessment, as compared to an individual who is having to imagine what it would be like (Shaw, 2011). Nonetheless, some patients with chronic conditions, such as back pain, may become accustomed to the discomfort over time, which may alter their perception of what health is. Whilst some studies have reported no significant differences between the preferences of patients and the populations (Dolders et al., 2006), in general patients' preferences relating to their own condition are higher than those of the general population, though this effect does appear to vary across different health conditions (Brazier et al., 2005). It is acknowledged that further research is required in the application of patient preferences, yet currently the use of population values is widely accepted as best practice (Brazier et al., 2009).

The values of the adult population are typically used in health technology assessment. This approach is advised because adults ultimately fund the healthcare system as taxpayers, and hence their preferences should count (Rowen et al., 2018b). Nonetheless, some practical issues have been reported in relation to obtaining values from adults for paediatric PBMs.

One main issue relates to the viewpoint adopted by the adult when providing preference weight. During valuation of the HUI2, adult respondents were asked to imagine that they were a 10-year-old child when completing the preference elicitation tasks (McCabe et al., 2005). Some concerns were raised following use of this approach, as it became clear that adults were interpreting the instructions differently. Some respondents tried to remember what it was like when they were 10-years-old, some felt they were answering on behalf of their own child, whilst others thought of an imaginary 10-year-old child. This area was explored further by Kind and coworkers (2015), who identified that adult respondents provided lower health state values when asked to consider the states applying to a 10-year-old child, than when ascribing values to adults. The authors concluded that applying adult values to health states for a paediatric PBM could ultimately lead to misrepresentation of health utilities for child health states (Kind et al., 2015).

In order to simplify matters, the CHU9D adopted a different approach during valuation, and asked adult respondents just to imagine themselves in the health state (Stevens, 2012). However, one could question the validity of the preference weights for the CHU9D as the values were not provided with children in mind. This highlights an issue which is difficult to overcome when obtaining adult values for paediatric PBMs. Further to this, it is widely accepted that adult preferences do not reflect the preferences of children and young people (Rowen et al., 2018b).

Attempts to overcome the issues in obtaining values from adults for paediatric measures have been made, through the use of parent/carer proxy values. There is a wealth of evidence to demonstrate that there is substantial variation between child and parent-reported HRQoL through use of PROMs and PBMs, with parents both overestimating and underestimating their child's HRQoL (Cremeens et al., 2006, Vetter et al., 2012, Bray et al., 2017, Galloway and Newman, 2017, Germain et al., 2019). As such, preference elicitation with children themselves may be more appropriate and informative (Wolstenholme et al., 2018).As the health states are experienced by children, one could argue that it is their own preferences that are important. Some aspects of HRQoL impact differently for children, parents/carers and adults, and adults may be unable to account for this. Furthermore, some parts of the classification system, such as the item within CHU9D relating to schoolwork, do not easily translate to adults, and can cause issues during valuation (Stevens, 2012).

In view of these difficulties, and the increased emphasis on involving children in both research and healthcare decisions, the possibility of eliciting preferences from children is being explored.

2.6.2 Which preference elicitation technique should be used with children?

Preference-weights can be elicited in a variety of ways, yet cardinal approaches have featured heavily within the literature. Nonetheless, these may not be the most suitable methods for use with children.

Cardinal approaches

Cardinal methods such as time trade-off (TTO) and standard gamble (SG) are the most commonly used techniques to assign preference weighting and require participants to consider trading a year of their life and the risk of death respectively.

The use of these tasks with children has been criticised, particularly as the tasks may present a range of symptoms that the child may have never experienced, and likewise the concepts of time and death may be difficult for the child to be understand (Petrou, 2003). As such, the validity of the preferences elicited could be questioned (Crump et al., 2017). Moreover, discussing death with children is likely to raise ethical concerns.

Furthermore, whilst one measure, the AQoL6D, has gained preference weights using the TTO method with adolescents aged 16-17 years, in general, these methods are considered to be too cognitively demanding for younger children (Moodie et al., 2010, Ungar, 2011). Ordinal approaches may thus provide a more suitable alternative.

Ordinal approaches

The use of ordinal techniques, such as best-worst scaling (BWS) and discrete choice experiments (DCE), have shown promise as methods to access children and young people's preferences (Ratcliffe et al., 2011, Sung and Regier, 2013, Stevens, 2015).

Introduced to the field of health economics in the 1990's, a DCE task presents the respondent with two or more health profiles which vary with respect to attribute levels (Krucien et al., 2017). Based in random utility theory, the DCE method assumes that respondents choose the health profile with the combination of attribute levels that give him/her the highest utility (Sung and Regier, 2013). DCE approaches are being increasingly used to estimate health state utilities for established adult instruments, such as the EQ-5D, with favourable results in comparison to cardinal methods (Brazier et al., 2012a).

There is a well-developed evidence-base which confirms that DCEs have a sound theoretical basis, and addresses any limitations of this technique (Krucien et al., 2017). One such limitation is that the conventional DCE cannot be anchored on the 1-0 scale of full health to death. To address this, an alternative form of the DCE has been suggested, known as the DCE_{TTO} (Bansback et al., 2012). This form of DCE draws from the aforementioned TTO method through the inclusion of a duration attribute alongside the HRQoL attributes, which enables it to be anchored onto the 1-0 scale. Furthermore, whilst an interviewer is usually required to elicit preferences using the TTO, the DCE_{TTO} could be administered in a less resource-intensive way, as respondents can complete the tasks online (Bansback et al., 2012). Nonetheless, in view of its similarities with the TTO it is likely that the DCE_{TTO} would also be considered to be too cognitively demanding for children and young people.

In contrast, the most commonly employed variant of the BWS method, known as BWS Case 2 task, presents the respondent with one health state profile and ask them to choose the best feature and the worst feature. Through repeating this process numerous times with varying attribute level combinations, preference weights can be estimated. BWS has been used effectively to gain preference weights from 14-17 year olds for the CHU9D (Ratcliffe et al., 2012).

The BWS task is similar to the conventional DCE in that it does not contain any means of anchoring values onto the 1-0 full health to dead scale. However, methods to overcome this have been described in the literature. One such way, is to re-scale preferences obtained using BWS based upon preferences obtained using cardinal approach, such as the TTO (Ratcliffe et al., 2015). However, as the latter methods may be unsuitable for children and young people for the reasons described earlier, the re-scaling values may need to be obtained from adults. A similar approach to re-scaling can also be used to anchor conventional DCEs to the 1-0 scale. It should be noted that a range of approaches exist to facilitate anchoring preferences onto the 1-0 full health to dead scale, and the approach chosen may impact upon the utilities generated (Rowen et al., 2009, Rowen et al., 2015). These approaches are discussed in further detail in section 7.1.1.

Nonetheless, Krucien and coworkers (2017) advise researchers to use the BWS method with caution. They highlight both technical and theoretical limitations to this technique in an adult population when compared to the DCE approach, and hence advocate the use of the DCE in preference (Krucien et al., 2017).

Whilst some of the concerns surrounding BWS may well be upheld if used within a paediatric population, it is possible that further limitations of the DCE could arise when used with children, and as such, no single approach should be advised for children at this stage.

The age at which a child is able to undertake a preference elicitation task is dependent on many factors. The ability to read the text contained within the tasks is of course important, but in order to process a hypothetical scenario they must be able to undertake abstract reasoning (Steinberg, 2005). Furthermore, they require the skills to weigh-up and debate alternative options, and to reach a decision (Yurgelun-Todd, 2007). Whilst there is variation amongst individuals, these abilities are key milestones of cognitive development, which are typically reached during early adolescence (Steinberg, 2005). The child must also be able to communicate their decision, whether verbally, digitally or otherwise, though it is important to acknowledge that there may be other factors not related to age that may impact upon these abilities.

Only one pilot study has qualitatively investigated the two ordinal methods to investigate the feasibility and reliability of each approach with children of different age groups (Stevens, 2015). The author found that adolescents aged 14-years and above were able to manage DCEs and children as young as 10 years could manage BWS tasks, suggesting that BWS tasks are not as cognitively-demanding as DCEs. However, a comparison of the two methods with adults revealed both a preference for DCEs, and greater consistency with underlying theory than BWS (Whitty et al., 2014). As such, further qualitative work is required to determine which type of ordinal task is most appropriate for children and young people.

In the specific context of CSPBMs, ordinal techniques have not yet been used to estimate preference-weights from children and young people, yet they have been used successfully in this way with adults (Netten et al., 2012, Brazier et al., 2012a). One potential issue surrounding eliciting preferences from children for CSPBM, is that some prior knowledge or even experience of the condition itself may be required in order for the children to be able to understand the information within the tasks. One way to overcome this issue, whilst still using a general population-based approach to valuation, is to provide some information about the condition for children to read before they complete the tasks (NICE, 2018).

As mentioned previously, there is potential for the use of ordinal tasks to be self-completed, without the need for an interviewer. Preliminary studies have shown potential for this approach using BWS surveys with adolescents, though there are few studies that have conducted interviews with young participants to determine the reliability of self-completion methods in this age group (Stevens, 2015, Ratcliffe et al., 2016a). With regard to younger children, face-to-face interviewing techniques have been used to elicit values when using ordinal methods (Stevens, 2015). Previous research with adults comparing online and computer-assisted personal interviewing techniques found no effect of the mode of administration on responses to valuation questions (Mulhern et al., 2013a). If shown to be feasible with children and young people, online self-completed formats would be a more time- and cost-efficient valuation process than interviewing, particularly when considering the large sample sizes required. The mode of administration of preference-elicitation tasks is discussed further in section 6.2.3.

2.7 Conclusion

This chapter has outlined the evidence base surrounding dental caries and its impacts on children's OHRQoL. Furthermore, it has explored the ways in which OHRQoL can be measured and has considered the rationale for generating QALYs in oral health research. Lastly, it has reviewed the approaches to involve children in research and has examined the literature regarding preference elicitation with this population. The next chapter will provide justification for undertaking this study and outline the aims and objectives.

3 Aims and Objectives

It is clear that dental caries has substantial impacts on the lives of children and young people, as well as their families and society as a whole. Patient and population-based strategies to prevent caries, as well as interventions to manage the condition, have been shown to improve clinical outcomes. However, as outlined in the previous chapter, the impacts of dental caries extend beyond the mouth, hence measuring outcomes in terms of OHRQoL provides a more holistic approach. The CARIES-QC instrument has inherent advantages over other measures of OHRQoL in that it has been designed specifically to determine the impacts and experiences of children with caries, and has involved children at every stage of development. Despite the increasing body of evidence to show the effectiveness of our interventions in improving OHRQoL, the relative cost-effectiveness of the differing approaches remains largely unknown. Furthermore, measures of OHRQoL such as CARIES-QC are of limited use in economic evaluations as they are not preference-based. Instead, they have a simple summative scoring algorithm that assumes that equal intervals between response levels and attributes are of equal importance (Ratcliffe et al., 2011).

Economic evaluations provide decision-makers with important information that can inform resource allocation, yet there is an acknowledged paucity of evaluations of interventions to improve children's oral health (Lord et al., 2015a, NICE, 2015). As a result, there is a poor evidence-base to inform the current allocation of resources to improve children's oral health (Public Health England, 2014a). Cost-utility analyses can offer some advantages over cost-effectiveness analyses and as a result tend to be preferred by decision-makers. Importantly, determining the effectiveness of interventions through using a common unit of measure, such as QALYs gained, allows comparison of interventions both within and between different oral health conditions.

Preference-based measures (PBMs) are the predominant means of generating the QALY, with the quality adjustment component being derived from HRQoL. There are a number of generic PBMs (GPBMs) available which have been designed for use in paediatric populations, yet only the CHU9D has involved children in the development of the classification system. Nonetheless, in a clinical study of children with dental caries, the

CHU9D was found to lack the degree of sensitivity required to detect changes in oral health status (Foster Page et al., 2015).

Despite there being some debate over the use of condition-specific PBMs (CSPBMs) in healthcare, current UK guidance for health technology assessment in adults states that their use is acceptable where no suitable GPBM exists (NICE, 2018). Whilst there is currently no existing paediatric PBM specific to dental caries, the development of such a measure appears to be both justifiable and feasible. The conversion of CARIES-QC into a caries-specific paediatric PBM was considered during its development, yet would require considerable refinement as well as the application of preference weights.

The importance of involving children in healthcare research has been emphasised over recent years (Marshman et al., 2015). As a result, there is an increasing body of evidence surrounding the ability of children of different ages to elicit preferences using different techniques (Ratcliffe et al., 2011, Stevens, 2015, Ratcliffe et al., 2016a, Chen et al., 2019). Whilst this appears to be both feasible and reliable using ordinal tasks such as DCE and BWS, at present, there is currently insufficient methodological evidence on this topic to determine which method to use (Stevens, 2015). Once the most suitable method has been determined, the adoption of a paediatric population-based approach to preference elicitation would be considered appropriate.

In summary, there is an indication for the development of a child-centred, PBM specific to dental caries. The conversion of CARIES-QC into a CSPBM would appear to be both practical and feasible, though further methodological work is necessary to determine which type of preference-elicitation task is most appropriate for use with children.

3.1.1 Aim

As informed by the literature described above, this study aimed to develop a PBM of dental caries based on the preferences of children and young people.

3.1.2 Objectives

The following specific objectives were identified in order to meet this aim:

6. To identify the quality and scope of published economic evaluations in child oral health research and the measures of benefit currently used

7. To develop and validate a classification system for child dental caries, based upon the CARIES-QC caries-specific measure of OHRQoL, that is amenable to health state valuation

8. To determine what age range of children and young people can complete ordinal health state valuation tasks, whether they prefer DCE or BWS and at what age they can use a computerised format independently

9. To undertake a population-based valuation survey with children and young people to determine their preferences

10. To model their preferences to produce a valuation algorithm that provides preference weights for each health state defined by the classification system

The following chapter aims to address the first of these objectives, through presenting a systematic review of the quality and scope of existing economic evaluations within child oral health research. This area has not previously been explored within the literature, and hence the use of CUA within this field remains unclear, as is the extent of involvement of children in these studies. The findings from this systematic review are intended to inform the wider study.

4 Systematic review of the quality and scope of economic evaluations in child oral health research

4.1 Background

The previous chapters have established the importance of preference-based measures (PBMs) and Quality Adjusted Life Years (QALYs). Furthermore, the need to involve children and young people in research and healthcare decisions has been highlighted. This chapter will present a systematic review of economic evaluations in child oral health research focusing on a quality appraisal of this literature against recognised criteria. This review intends to explore the use of utility measures in this field, and consider barriers and facilitators to their use. The involvement of children in the identified studies will also be examined, with due consideration given to how this could be improved in future economic evaluations. Furthermore, a greater understanding of the strengths and limitations of existing economic evaluations in child oral health research will enable recommendations to be targeted appropriately. It is anticipated that the findings from this systematic review will help to inform the development of this project accordingly.

Dental caries is a largely preventable disease, and there are a range of interventions available to reduce the incidence of caries in children, such as fluoride varnish application and supervised toothbrushing programmes in childhood settings. However, a recent analysis of oral health promotion approaches, commissioned by the National Institute of Health and Care Excellence (NICE), confirms that there is a paucity of economic evaluations with sound methodology (Lord et al., 2015b). Commissioners of dental services and oral health programmes have acknowledged this deficiency and hence local authorities have been recommended to consider economic evaluation of caries prevention programmes (Public Health England, 2014a).

Whilst dental caries is the most prevalent dental problem to affect children, there are various other common oral conditions, each with considerable financial implications. For example, one third of all preschool children have suffered a traumatic dental injury involving the primary dentition, whilst one quarter of all school children have suffered an injury affecting the permanent dentition (Glendor, 2008). Additionally, the prevalence of molar-incisor hypomineralisation (a qualitative enamel defect of systemic origin) ranges in the literature from 3.6 to 25% (Weerheijm, 2003). Both of these conditions frequently require multiple dental visits over a prolonged period of time with multi-disciplinary input. Furthermore, the Child Dental Health Survey in England, Wales and Northern Ireland 2013 found that 9% of 12-year olds and 18% of 15-year olds were undergoing orthodontic treatment, utilising a considerable proportion of the NHS dental budget (Rolland et al., 2016). It is important to determine whether the scope of existing economic evaluations in child oral health research encompasses these conditions also.

As discussed in the previous chapter, there is now persuasive evidence that children and young people are able to report on their own health, and the importance of involving them in healthcare decisions is increasingly acknowledged (United Nations, 2009, Stevens, 2010). A landmark systematic review of paediatric dental literature, published from 2000 to 2005, was undertaken to assess the extent of children's involvement in oral health-related research (Marshman et al., 2007). Disappointingly, only 7.3% of included studies were found to have involved children as active participants. Following dissemination of these findings, recommendations were made to increase involvement of children in oral health-related research. A subsequent review found involvement of children in studies published between 2006 and 2014 had increased to 17.4%; an encouraging finding (Marshman et al., 2015). Whilst this is becoming an increasingly frequent finding in some areas of child oral health research, the involvement of children in economic evaluations of child oral health research has not yet been investigated.

As data from economic evaluations are often used to aid decision-makers, it is essential that these studies are of sufficient scientific quality. A recent systematic review of economic evaluations in all fields of dentistry found many included studies to be flawed, with confusion regarding terminology, inconsistencies and a lack of sound research methodology (Tonmukayakul et al., 2015). Another recent systematic review found a need for improvement in the reporting of economic evaluations of oral health interventions (Hettiarachchi et al., 2017). Nonetheless, neither of these systematic reviews explored both the methodological and reporting quality of the economic evaluations, and neither focused specifically on child oral health research.

The reviewing team for this study was comprised of Helen Rogers (HJR), Professor Helen Rodd (HDR), Professor Zoe Marshman (ZM) and Dr Erik Vermaire (EV). Data extraction for the study was carried out by Rebecca Knapp (BK) and Sarab El Yousfi (SE). Dr Katherine Stevens (KS) provided overall guidance for this review, acting as a third reviewer when required to reach a consensus.

4.2 Aim

This systematic review aimed to examine the quality and scope of economic evaluations in the field of child oral health research

4.2.1 Objectives

The following objectives were set to fulfill this aim:

1. To describe the frequency and trends in the publication of economic evaluations in child oral health research
2. To explore the extent to which children have been involved in economic evaluations of child oral health
3. To examine the quality of published economic evaluations in child oral health research using two quality assessment tools specifically developed for appraisal of economic evaluations
4. To identify which instruments have been used to measure the benefits of child oral health interventions in existing economic evaluations in this field

4.3 Method

4.3.1 Search strategy

A comprehensive search strategy was developed iteratively, combining search terms relating to the key concepts with adaptations of the validated CRD economic evaluation search filter for the databases MEDLINE, EMBASE and CINAHL. The search filters were modified to reflect the inclusion and exclusion criteria, and altered for use in the remaining databases accordingly.

One researcher (HJR) searched the following databases on 17th January 2017: NHS Economic Evaluation Database (CRD York), MEDLINE, EMBASE, CINAHL, Web of Science, Scopus, the Cochrane Library and Econlit. Each search covered the period from commencement of each database system until the initiation of the systematic review. Database searches were re-run on 5th June 2017 prior to final analysis to identify any further studies suitable for inclusion. The search strategy used can be found in Appendix C.

Bibliographic information from identified studies was examined for further applicable titles. Efforts were made to identify relevant unpublished 'grey' literature, theses and conference proceedings through appropriate websites and the databases OpenGrey and EThOS.

4.3.2 Selection criteria

Search results were de-duplicated and organised using EndNote™ X8.2 (Clarivate™ Analytics, Philadelphia, USA). Potentially relevant titles and abstracts were screened against the inclusion and exclusion criteria below by HJR.

Inclusion criteria

- Studies involving children aged 18 years old and under
- Studies including a full economic evaluation in the field of child oral health
- Studies published after 1997

Exclusion criteria

- Studies including participants over 18 years of age
- Decision models extending past 18 years of age
- Studies not in the field of oral health
- Studies published in and prior to 1997

After discussion between all reviewers, it was agreed that studies published in and prior to 1997 should be excluded from this review. This is due to the limited guidance available to researchers before the publication and wider dissemination of the Drummond checklist (Drummond et al., 1997). Furthermore, the reviewing team had some concerns that cost-minimisation analyses (CMA) may not be universally considered as full economic evaluations. Nonetheless, it was agreed that they should be included in this review for completeness. The reviewing team agreed that the exclusion of decision models extending past 18 years of age would be appropriate for this review. Whilst the reviewers agreed that this could be considered as a limitation of this study, for the purposes of this study in informing a larger project, it was justified as being more important to focus on the benefits gained during the childhood and adolescence period, rather than those accrued throughout life.

Full texts were retrieved for all titles appearing to meet the aforementioned criteria (HJR). Two reviewers (HJR and EV) then reviewed the full texts against the inclusion and exclusion criteria independently, with any disagreement resolved by consensus. Input from a third reviewer (KS) was sought where required. No language restrictions were imposed in this review, and translation was requested where necessary to establish whether studies published in other languages met these criteria. A record was maintained of all studies excluded during this stage, with justification for their exclusion.

4.3.3 Data extraction

Relevant data were transferred from each included study into a comprehensive Microsoft® Excel® 2016 (Microsoft Corporation, Redmond, USA) spreadsheet by two reviewers (BK and SE), following provision of a data extraction training exercise by HJR. Data were collected in categorical format, where possible, to facilitate analysis. The data characteristics collected are listed in Table 4-1.

This task was also used to determine the level of involvement of children in the included studies. When extracting data regarding the reporting of outcomes, the reviewers were required to select one of the options stated in Notes: *Only applicable to studies reporting primary data

Table 4-2. These options were adapted from the aforementioned systematic reviews of oral health-related literature to establish involvement of children (Marshman et al., 2007, Marshman et al., 2015). Where a combination of outcome reporting was identified, the data extraction team (BK and SE) documented which types were involved.

For included studies published in languages other than English, data extraction and determination of the level of involvement of children was undertaken by one reviewer (HJR) with an appropriate translator.

Table 4-1: Description of data extracted from each study

Data extracted	Description
Publication characteristics	Publication year Journal title Country in which study was conducted
Economic evaluation characteristics	Type of full economic evaluation Data source Study duration/analytic horizon Cost of intervention Effects Cost-effectiveness results
Aim of study	
Study characteristics*	Setting Oral health condition studied Nature of intervention Outcome reporting and measure used
Participant characteristics*	Age range of participants Number of participants Is there a focus on deprivation? Sample size

Notes: *Only applicable to studies reporting primary data

Table 4-2: Data extraction options regarding outcome reporting

1.	**Clinician-Reported (e.g. DMFT)**
2.	Parent-Reported (e.g. P-CPQ*)
3.	Child-reported (e.g. CARIES-QC, CHU9D)
4.	Combination
5.	Not applicable (e.g. for studies using data from multiple studies/model-based studies)

Notes: Parental-Caregivers Perceptions Questionnaire

4.3.4 Assessment of methodological quality

Since the 1990's, numerous guidelines have been introduced to support researchers and economists in producing high quality economic evaluations, yet the most widely used is the Drummond 10-item checklist (Drummond et al., 1997, Walker, 2012). A simplified version of the detailed 35-item Drummond checklist, it provides comprehensive guidance on the methodological conduct of an economic evaluation. The checklist (Appendix D) is not only used as a guideline, but also as a quality appraisal tool for economic evaluations across different areas of healthcare. Furthermore, this checklist is recommended in the Cochrane Handbook for Systematic Reviews of Interventions (Shemilt et al., 2008).

This 10-item, 13-criteria checklist provides a description of which methodological characteristics should be included in that section of an economic evaluation. When used as a quality appraisal tool, a mark is allocated by the reviewer for each subcomponent in which the study adheres to the recommendations, as described in section 4.3.6.

A calibration exercise led by HJR was conducted with reviewers HDR and ZM prior to commencement of quality appraisal to enable familiarisation with the Drummond checklist, and to gain consistency in scoring. The two reviewers (HDR and ZM) independently marked each included study against the criteria detailed in this checklist. The independent scores assigned against each item of the checklist were assessed for inter-rater reliability through calculation of the percentage agreement. Overall agreement was also determined, alongside Cohen’s kappa (κ) statistics to account for the possibility of a chance agreement (McHugh, 2012). The latter were interpreted against the classification of strength of agreement proposed by Landis and Koch (1977).

The reviewers (HDR and ZM) assessed 10% of the included studies a second time to determine intra-rater reliability. These studies were selected randomly using an online random number generator. The percentage agreement and kappa statistics were calculated and interpreted in the same way as the inter-rater reliability.

Resolution of disagreement in the marks assigned by the independent reviewers was achieved through discussion to reach a consensus decision.

4.3.5 Assessment of reporting quality

When preparing the protocol for this systematic review, personal contact was made with Professor Michael Drummond (University of York, UK) and Don Husereau (University of Ottawa, Canada) regarding the quality appraisal component of the proposed systematic review. These prominent researchers in health economics advised use of both the aforementioned Drummond checklist, to assess methodological quality, alongside the novel Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist (Drummond et al., 1997, Husereau et al., 2013).

The CHEERS checklist was developed by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force, in response to a need for consolidated, updated, user-friendly reporting guidelines (Husereau et al., 2013). Published in 2013, these standards provide a 24-item user-friendly checklist with accompanying recommendations and examples, with the overall aim to encourage more consistent and transparent reporting in this field.

The CHEERS checklist has been used in a number of systematic reviews of economic evaluations of healthcare interventions to date, including one in the field of oral health interventions which was published during the course of this systematic review (Hettiarachchi et al., 2017).

When used as a quality appraisal tool, a mark is allocated by the reviewer for each relevant criterion in which the study adheres to the recommendations, as described in section 4.3.6. A calibration exercise was conducted between two reviewers (HJR and EV) to enable familiarisation with the CHEERS checklist and to gain consistency in scoring. Of note, not all of the 24 criteria were relevant for each type of study. For example, some criteria related specifically to modelling studies, which were not applicable for studies using primary data, and vice versa. This was discussed by the reviewers during calibration to ensure consistency. The two reviewers (HJR and EV) independently marked each included study against the criteria detailed in this checklist. The independent scores assigned against each item of the checklist were assessed for inter-rater reliability through calculation of the percentage agreement. Overall agreement was also determined, alongside Cohen's kappa (κ) statistics to account for the possibility of a chance agreement (McHugh, 2012). The latter were interpreted against the classification of strength of agreement proposed by Landis and Koch (1977).

The reviewers (HJR and EV) assessed 10% of the included studies a second time to determine intra-rater reliability. These studies were selected randomly using an online random number generator. The percentage agreement and kappa statistics were calculated and interpreted in the same way as the inter-rater reliability.

Resolution of disagreement in the marks assigned by the independent reviewers was achieved through discussion to reach a consensus decision.

For studies published in languages other than English, both methodological and reporting quality appraisal were undertaken by one reviewer (HJR) with an appropriate translator.

4.3.6 Strategy for data synthesis

Extracted data were grouped according to the oral health condition studied and simple descriptive comparisons undertaken using SPSS® Statistics software (IBM Corporation, United States, V23).

Synthesis of data from the methodological and reporting quality appraisal followed the precedent set by the aforementioned study by Tonmukayakal and coworkers (2015). The authors applied the Drummond checklist to produce a numerical score for each paper, based upon the total number of criteria that were met. The authors then used this to score to assign each study to an arbitrarily assigned category ('underperformed' and 'well-performed') to facilitate a form of quantitative analysis. The present study employed this same approach to synthesise the results from the methodological quality appraisal to enable comparison with the earlier study. Furthermore, the same approach was applied to the reporting quality appraisal using the CHEERS checklist also, to allow comparison between the two types of quality assessment.

As such, a score of 0, 1 or 2 was allocated by the reviewers for each criterion as follows:

Score 0: Criterion not met

Score 1: Criterion met

Score 2: Criterion not applicable

Simple descriptive statistics were undertaken on the quality appraisal scores using SPSS® Statistics software (IBM Corporation, United States, V23). Percentages of applicable Drummond and CHEERS criteria met by each paper were used as the basis for identifying the median, range and tertiles, to ensure that studies with a larger number of ‘not applicable’ criteria would not be unfairly disadvantaged. The categories of high, moderate and low reporting quality were determined based upon the tertiles. It should be acknowledged that this type of analysis has not been formally acknowledged by the authors of the Drummond or CHEERS checklists, nor have standardised ‘cut-off’ scores been proposed in the literature. As such, this approach was utilised in conjunction with a narrative synthesis.

4.4 Results

The database search was conducted on 17th January 2017, and yielded 2757 studies, with an additional 4 records identified from other sources. Following removal of duplicates, 2315 records remained. Following screening of titles, 2042 records were excluded, leaving a total of 273 records. The abstracts for these records were assessed for eligibility, and 76 appeared to meet the inclusion criteria. Following careful review of full texts, 46 were included in the final analysis. A summary of these studies can be found in Appendix F. The search strategy was repeated on 5th June 2017, to identify further, more recent publications that could meet the search criteria. This search identified four studies, yet three failed to meet the inclusion criteria, and one had already been identified in the initial search.

Figure 4-2 reveals the general trend for an increase in publications in this field, with an apparent peak in 2016 (n=6). It should be noted, however, that publications from 2017 have been excluded from this figure as the review did not cover the full year.

Of the 46 studies included in the final analysis, three were written in languages other than English, namely Portuguese (n=2) and Mandarin (n=1). Most studies (n=16) were undertaken in the USA, where healthcare provision is insurance-based. The rest of the studies were conducted in a range of countries, yet no more than five studies were undertaken in any single country other than the USA (Figure 4-3).

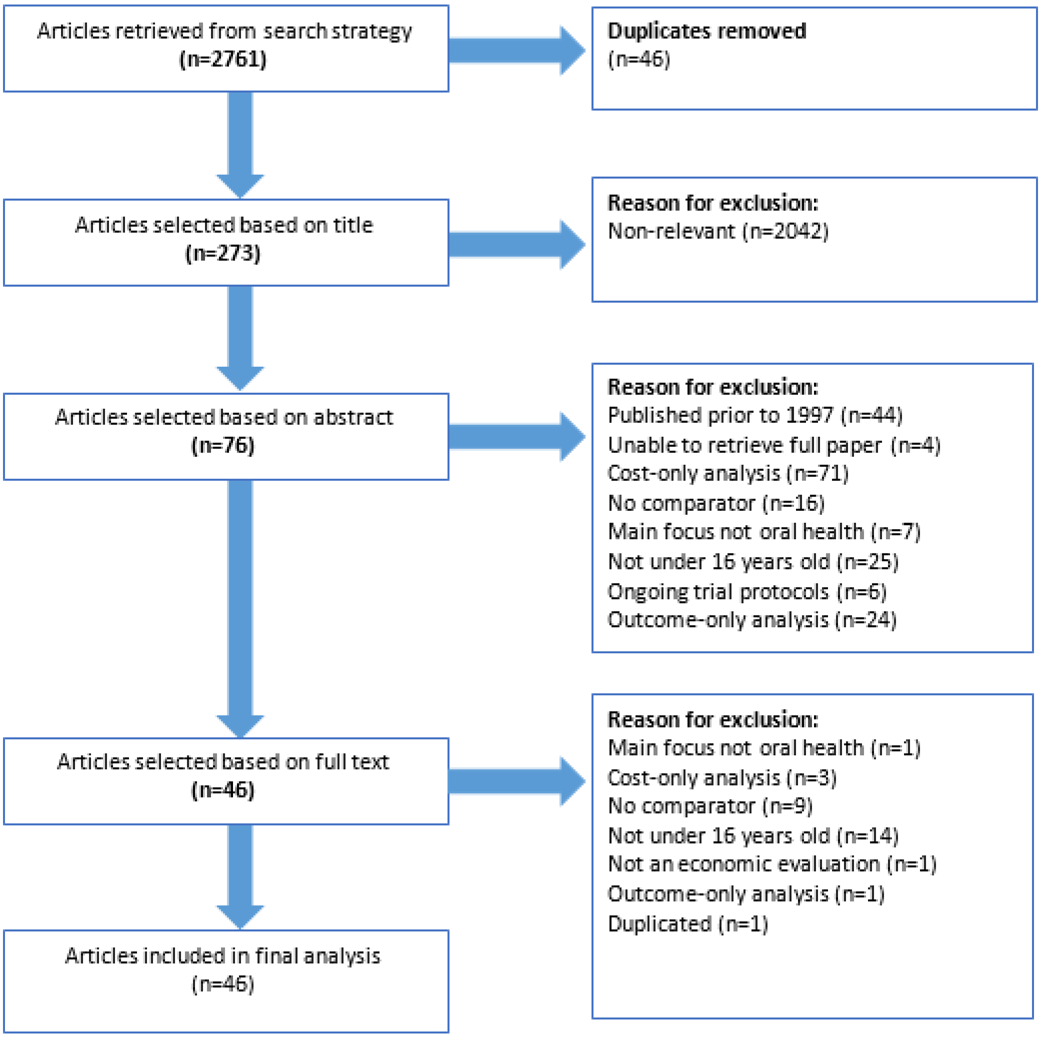


Figure 4-1: Flowchart displaying search results

The overwhelming majority of studies undertook cost-effectiveness analyses (n=38; 82%), whereby the health outcomes were reported in terms of natural units. Health benefits were measured in monetary terms in one study, which carried out cost-benefit analysis alone (n=1; 2%). One study reported a cost-utility analysis alone (n=1; 2%), reporting health outcomes in terms of utilities. Two studies carried out two different types of analyses; both undertook a cost-effectiveness analysis, complemented by a cost-benefit analysis or a cost-utility analysis. Four studies (9%) reported the findings of cost-minimisation analyses. Of the included studies, 61% (n=28) used primary data, with the remainder (n=18) being modelling studies, using data from various sources. The economic outcomes of these studies can be seen in Table 10.2 in Appendix F.

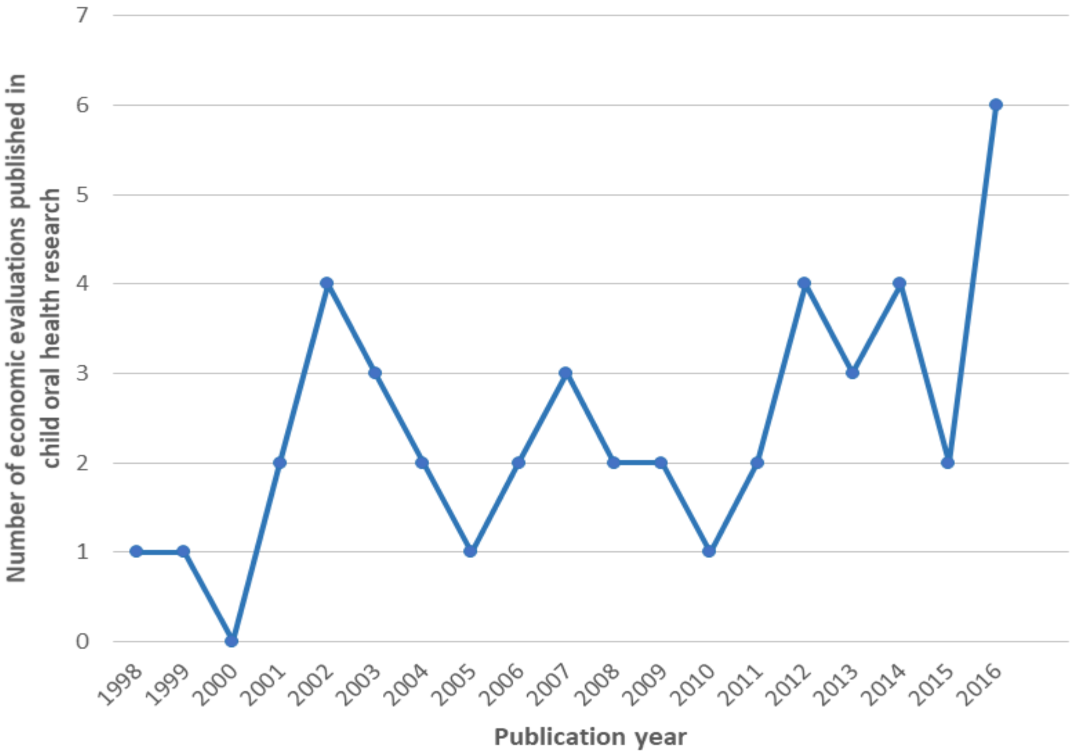


Figure 4-2: Graph displaying trends in the publication of economic evaluations in the field of child oral health research

Whilst the age range of children varied between studies, the majority focused on children in the primary (n=17; 37%) or mixed (n=17; 37%) dentitions. Two studies involved children spanning the mixed and permanent dentitions (4%). Four studies involved children solely in the permanent dentition (9%), and six studies did not provide sufficient information to determine the age range of participants (13%).

As displayed in Figure 4-4, most studies focused on the prevention or management of dental caries (n=42; 91%), with only three studies (7%) relating to malocclusion and one on dental anxiety (2%). The range of interventions provided in the studies varied significantly. Many studies compared the cost-effectiveness of a standard preventive programme with a more comprehensive or targeted preventive programme (n=13). Similarly, a number of studies assessed the cost-effectiveness of an intervention centred on fluoride delivery (n=13).

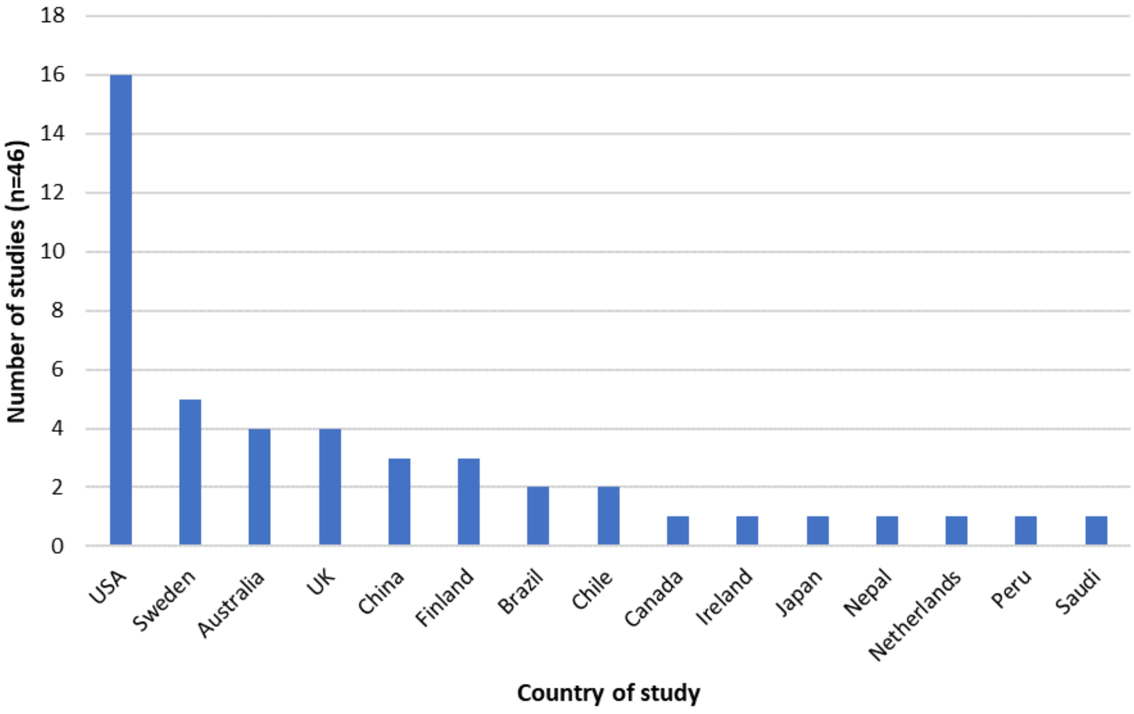


Figure 4-3: Graph displaying the countries where the included economic evaluations were undertaken

The setting for the studies was more varied, with 48% (n=22) relating to interventions in a non-dental community setting, such as schools or homes, and 17% (n=8) based in general dental practice. Fewer studies assessed interventions provided in a clinical community setting (n=3), a specialist dental clinic (n=1) and a hospital setting (n=1). A combination of settings were used in 20% (n=9) of the studies, whilst two studies failed to provide sufficient information to enable the setting to be determined.

Outcomes were reported by clinicians in the majority of studies (n=40; 87%). One study gained parent-reported outcomes, which were used in combination with clinician-reported outcomes. Only one study used a child-reported outcome measure yet this was not a validated tool but rather a satisfaction survey created for the purposes of the study. Furthermore, the findings did not contribute to the cost-effectiveness analysis. A total of four studies (9%) failed to provide sufficient detail regarding who was reporting the outcomes.

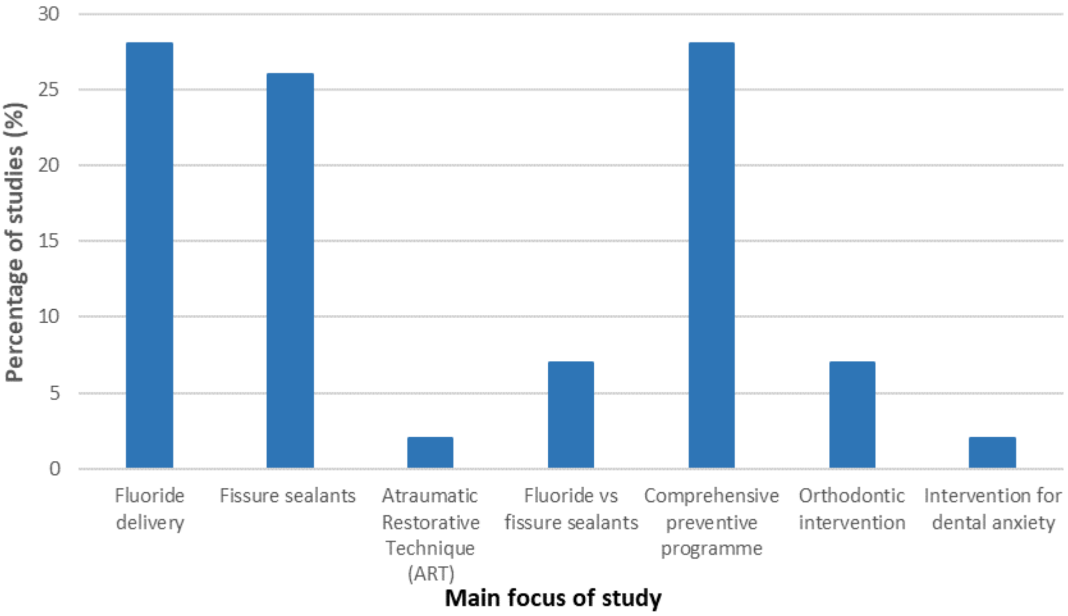


Figure 4-4: Graph displaying the range of interventions provided in included studies

A range of outcome measures were used in the studies. Validated measures were used in 26 studies (56%), including dmfs/DMFS (n=9), DFT (n=3) and dmft/DMFT (n=7). Non-validated outcome measures, such as 'time spent on treatment', and 'presence of caries' were used in 20 studies (44%). Only two studies quantified health outcomes in terms of utilities. One reported outcomes in QALYs, for which data were collected using the aforementioned paediatric PBM known as the CHU9D (Child Health Utility 9 Dimensions). The other study reported outcomes in Quality Adjusted Tooth Years (QATYs), a dental variation of the QALY, discussed previously in section 2.3.1.

Despite the well-reported association between deprivation and caries experience, only 14 studies (30%) acknowledged deprivation as a factor in their evaluations (Pitts et al., 2015).

4.4.1 Methodological quality appraisal results

The overall mean percentage of applicable Drummond checklist criteria met by the studies in this review was found to be 48%, with a range of 0 to 100%. The median score was calculated at 50%. Two studies met all the applicable criteria, scoring 100%, whilst two studies failed to meet any of the applicable criteria. The performance of each study in relation to the Drummond checklist criteria can be seen in

Table 4-3.

Table 4-3: The percentage of applicable Drummond and CHEERS criteria met by each paper, with categorisation to indicate overall quality

Author (year)	% applicable Drummond criteria met	Overall methodological quality	% applicable CHEERS criteria met	Overall reporting quality
Alkhadra (2004)	38	Moderate	65	Moderate
Atkins et al. (2016)	46	Moderate	96	High
Bergstrom et al. (2016)	23	Low	70	Moderate
Bertrand et al. (2011)	54	High	91	High
Bhuridej et al. (2007)	92	High	86	High
Chi et al. (2014)	69	High	96	High
Davies et al. (2003)	54	High	90	High
Frazao (2012)	54	High	48	Low
Goldman et al. (2014)	85	High	90	High
Goldman et al. (2016)	100	High	95	High
Griffin et al. (2002)	77	High	87	High
Hichens et al. (2007)	36	Moderate	84	High
Hietasalo et al. (2009)	69	High	50	Low
Hirsch et al. (2012)	0	Low	39	Low
Holland et al. (2001)	62	High	35	Low
Jokela and Pienihakkinen (2003)	8	Low	60	Low
Kaakko et al. (2002)	15	Low	52	Low
Koh et al. (2015)	77	High	96	High
Kowash et al. (2006)	15	Low	85	High
Leskinen et al. (2008)	8	Low	76	Moderate
Marino et al. (2011)	77	High	87	High
Marino et al. (2012)	54	High	67	Moderate
Marino et al. (2007)	54	High	76	Moderate
Morgan et al. (1998)	38	Moderate	81	Moderate
Neidell et al. (2016)	38	Moderate	77	Moderate
Ney et al. (2014)	38	Moderate	87	High
Oscarson et al. (2003)	85	High	90	High
Ouyang (2009)	69	High	87	High
Petrén et al. (2013)	38	Moderate	62	Low
Pukallus et al. (2013)	54	High	100	High
Quiñonez et al. (2005)	54	High	82	Moderate
Quinonez et al. (2006)	62	High	95	High
Ramos-Gomez and Shepard (1999)	8	Low	55	Low
Sakuma et al. (2010)	62	High	62	Low
Samnaliev et al. (2015)	46	Moderate	100	High
Skold et al. (2008)	62	High	91	High
Stearns et al. (2012)	46	Moderate	100	High

Author (year) continued	% applicable Drummond criteria met	Overall methodological quality	% applicable CHEERS criteria met	Overall reporting quality
Tagliaferro et al. (2013)	77	High	43	Low
Tickle M (2016)	38	Moderate	95	High
Tonmukayakul and Arrow (2017)	46	Moderate	95	High
Vermaire et al. (2014)	100	High	90	High
Weintraub et al. (2001)	0	Low	78	Moderate
Wiedel et al. (2016)	23	Low	71	Moderate
Wu et al. (2002)	8	Low	33	Low
Yee et al. (2004)	23	Low	71	Moderate
Zabos et al. (2002)	31	Low	57	Low

Notes: High methodological quality >50% Drummond criteria met, moderate methodological quality 32-50% Drummond criteria met, low methodological quality <32% Drummond criteria met; high reporting quality > 83% CHEERS criteria met, moderate reporting quality 63-83% of CHEERS criteria met, low reporting quality <63% CHEERS criteria met.

A previous systematic review of dental economic evaluations by Tonmukayakal and coworkers (2015) assigned the reviewed studies into two categories, ‘underperformed’ and ‘well-performed’. Studies were assigned according to whether their overall score was above or below the median score for all included studies, as an arbitrary indicator of the overall quality of each study (Tonmukayakul et al., 2015). In order to facilitate some level of comparison, a similar system was utilised in this review, with the introduction of a third ‘moderate quality’ category.

For this review, the percentage of applicable Drummond criteria met by each paper were used as the basis for identifying the median, range and tertiles, to ensure that studies with a larger number of ‘not applicable’ criteria would not be unfairly disadvantaged. The categories of high, moderate and low reporting quality were determined based upon the tertiles (high methodological quality >50% Drummond criteria met; moderate methodological quality 32-50% Drummond criteria met; low methodological quality <32% Drummond criteria met). These are shown in

Table 4-3. It can be seen that 23 studies were judged to be of high methodological quality, 11 of moderate quality and 12 of low quality.

The following section provides a narrative synthesis of the findings of this review in relation to each of the criteria within the Drummond checklist. The number of studies meeting each of these criteria can be seen in

Table 4-4.

1. Was a well-defined question posed in answerable form?

Of the 46 included studies, 26 were found to examine the costs and effects of the alternatives, stating a viewpoint for the analysis. The remaining studies failed to provide sufficient detail to meet this criterion.

2. Was a comprehensive description of the of the competing alternatives given?

Less than half of the included studies (n=21) adequately described all important alternatives. The remainder either failed to consider all alternatives, including a do-nothing alternative if appropriate, or failed to provide enough detail to make it clear to the reviewers who did what, to whom, where, and how often.

3. Was there evidence that the programme(s) effectiveness had been established?

More than half of the studies (n=27) clearly stated how the effectiveness of the intervention had been established, for example, through a randomised controlled trial, using data from observational studies, or via a literature review. These studies also provided sufficient detail relating to any potential bias arising from these sources, and the generalisability of the findings. Failure to sufficiently establish the effectiveness of an intervention can render an economic evaluation defunct.

4. Were all the important and relevant outcomes and costs for each alternative identified?

Only 14 studies identified costs and outcomes to cover all relevant viewpoints, for example, the societal perspective, or that of patients. Failure to do this can result in an underestimation of the costs of an intervention; a significant methodological flaw.

5a. Were outcomes measured accurately in appropriate units prior to evaluation?

Of the 46 studies, 31 were found to have measured all identified outcomes in appropriate units. The remainder either omitted previously identified outcomes from the analysis, or utilised an inappropriate unit of measurement.

5b. Were costs measured accurately in appropriate units prior to evaluation?

Measurement of costs was found to be significantly inferior to the measurement of outcomes in the included studies, with only 12 studies meeting this criterion. Similar to the previous criterion, the remaining 34 studies either omitted costs from the analysis, or failed to measure them in appropriate physical units.

6a. Were the outcomes valued credibly?

Only 12 studies valued the outcomes appropriately in relation to the study question. The type of analysis selected should reflect the study perspective, for example, a cost-utility, cost-benefit, cost-effectiveness or cost-minimisation analysis. Where patient preferences were utilised, the source of the values should be clearly stated for reproducibility purposes. For the remaining 34 studies, a more suitable type of analysis could have been chosen, or insufficient detail regarding the source of the values was provided.

6b. Were the costs valued credibly?

As above, only 12 studies identified the sources of all costs, such as market values, or made every effort to approximate market values if they were absent.

7a. Were outcomes adjusted for different times at which they occurred?

Criteria 7a and 7b relate to the use of discounting; a process whereby costs and outcomes that occur in the future are adjusted to their present values. Discounting is important due to 'time preference'; the desire to enjoy benefits in the present while deferring any negative effects of doing so (Torgerson and Raftery, 1999). This criterion was deemed to have been met by 13 studies, as either discounting was necessary and undertaken, or not undertaken with due justification. A mark was not allocated to the 33 studies that failed to undertake discounting where it was necessary, or those that claimed discounting was not necessary, but the reviewers disagreed. One study was deemed by the reviewers to not require discounting since the duration of the study was too short, though this was not acknowledged in the study itself (Morgan et al., 1998), hence this criterion was deemed to be not applicable.

7b. Were costs adjusted for different times at which they occurred?

Discounting of costs was undertaken significantly more within the included studies than discounting of outcomes. A total of 29 studies undertook necessary discounting, or provided acceptable justification for not undertaking discounting. A mark was not allocated to the 16 studies that either failed to undertake discounting where the reviewers found it to be necessary, or claimed that discounting was not necessary. The aforementioned study that was deemed to not require discounting, though the authors failed to acknowledge this in the study, was exempt from this criterion.

8. Was an incremental analysis of the outcomes and costs of alternative interventions performed?

Of the 46 included studies, 28 undertook an incremental analysis to compare the costs and additional benefits of each alternative, the conclusion of which is typically presented as an Incremental Cost Effectiveness Ratio (ICER).

9. Was allowance made for uncertainty in the estimates of costs and consequences?

The highest scoring criterion; a total of 32 studies were deemed to have conducted appropriate statistical analyses, as well as a sensitivity analysis where appropriate. A sensitivity analysis is important to assess the robustness of the conclusions drawn from an economic evaluation.

10. Did the presentation and discussion of study results include all of the issues that are of concern to users?

The majority of studies (n=30) were deemed to have presented their conclusions appropriately, with the authors relating the findings to other studies in their field, and commenting on the generalisability of their work. Limitations of these studies were also considered, including any difficulties related to the adoption of a new dominant intervention. The remaining 16 studies failed to consider all of these factors in the presentation of the study findings.

Table 4-4: Frequency of studies meeting each Drummond checklist criteria

Criterion number	Criterion description	Studies meeting criterion n=46 (%)	Studies failing to meet criterion n=46 (%)	Studies criterion not applicable n=46 (%)
1	Was a well-defined question posed in answerable form?	26 (57)	20 (43)	0 (0)
2	Was a comprehensive description of the of the competing alternatives given?	21 (46)	25 (54)	0 (0)
3	Was there evidence that the programme's effectiveness had been established?	27 (59)	19 (41)	0 (0)
4	Were all the important and relevant outcomes and costs for each alternative identified?	14 (30)	32 (70)	0 (0)
5a	Were outcomes measured accurately in appropriate units prior to evaluation?	31 (67)	15 (33)	0 (0)
5b	Were costs measured accurately in appropriate units prior to evaluation?	12 (26)	34 (74)	0 (0)
6a	Were the outcomes valued credibly?	12 (26)	34 (74)	0 (0)
6b	Were the costs valued credibly?	12 (26)	34 (74)	0 (0)
7a	Were outcomes adjusted for different times at which they occurred?	13 (28)	33 (72)	1 (2)
7b	Were costs adjusted for different times at which they occurred?	29 (63)	17 (37)	1 (2)
8	Was an incremental analysis of the outcomes and costs of alternatives performed?	28 (61)	18 (39)	0 (0)
9	Was allowance made for uncertainty in the estimates of costs and consequences?	32 (70)	14 (30)	0 (0)
10	Did the presentation and discussion of study results include all of the issues that are of concern to users?	30 (65)	16 (35)	0 (0)

4.4.2 Inter-rater and intra-rater reliability

Two reviewers (HDR and ZM) independently assessed each study against the Drummond checklist. Whilst differences in scoring were agreed by consensus, the inter-rater reliability for each Drummond criterion can be seen in Table 4-5. Little variation in inter-rater agreement can be seen across the criteria, with overall agreement measured at 90%. The lowest level of agreement was noted for Drummond criterion 6b, which relates to the valuing of costs. Nonetheless, at 81% agreement, the inter-rater reliability for this criterion can still be considered relatively strong. Cohen's kappa (κ) was calculated at 0.8 for overall inter-rater agreement. According to the classifications proposed by Landis and Koch (1977), this is a substantial strength of agreement.

The reviewers (HDR and ZM) assessed 10% of the included studies (n=5) a second time to determine intra-rater reliability. For reviewer HDR, intra-rater reliability was weakest for Drummond criteria 4 and 5a (both 60% agreement), whilst for reviewer ZM, agreement was weakest for criterion 10 (60% agreement). Overall intra-rater reliability was slightly higher (κ = 0.87) for ZM than for HDR (κ = 0.64), hence the strength of agreement could be deemed as ‘almost perfect’ for ZM, and ‘substantial’ for HDR using the aforementioned classification.

Table 4-5: Inter- and intra-rater reliability for quality appraisal using Drummond checklist

Drummond criteria	Inter-rater reliability (% agreement)	Intra-rater reliability HDR (% agreement)	Intra-rater reliability ZM (% agreement)
1	93	100	100
2	86	80	80
3	91	100	100
4	91	60	100
5a	95	60	100
5b	88	80	100
6a	81	80	100
6b	93	80	100
7a	91	80	100
7b	86	80	100
8	93	100	80
9	91	80	100
10	88	100	60
Overall agreement	**90**	**83**	**94**

4.4.3 Reporting quality appraisal results

The overall mean percentage of applicable CHEERS criteria met by the studies included in this review was calculated at 77%, with a range of 33-100% and a median of 83%. A total of three studies met all applicable criteria (100%). These results can be seen alongside the Drummond checklist results in Table 4-3.

The percentage of applicable CHEERS criteria met by each paper were used as the basis for identifying the median, range and tertiles, to ensure that studies with a larger number of ‘not applicable’ criteria would be unfairly disadvantaged. The categories of high, moderate and low reporting quality were determined based upon the tertiles (high reporting quality > 83% CHEERS criteria met; moderate reporting quality 63-83% of CHEERS criteria met; low reporting quality <63% CHEERS criteria met). From this categorisation, it can be seen that 23 studies were deemed to be of good reporting quality, 11 of moderate quality and 12 of low quality.

The following section provides a narrative synthesis of the findings of this review in relation to each of the criteria within the CHEERS checklist. The number of studies meeting each of these criteria can be seen in Table 4-6 below.

1. Title

A total of 44 studies correctly identified the study as a form of economic evaluation in the title, and provided sufficient detail regarding the interventions provided. The majority used wording to further clarify the type of economic evaluation undertaken, for example ‘A cost-effectiveness study of…’. Two studies could not be easily identified as an economic evaluation in the title. One contained the words ‘effectiveness’ and ‘cost’ but failed to unite them, and the other had a title based upon the intervention itself (Holland et al., 2001, Kaakko et al., 2002).

2. Abstract

Of the 46 included studies, 45 provided a suitably structured abstract. The study that failed to meet this criterion did provide an abstract, but the reviewers agreed it did not contain sufficient details of the objectives, perspective, setting, methods, results and conclusions (Ramos-Gomez and Shepard, 1999).

3. Background and objectives

The overwhelming majority of studies (n=44) clearly stated the broader context for the study and provided a description of how the study question was relevant for health policy. The two studies that were deemed not to have met this criterion were both published in languages other than English (Wu et al., 2002, Tagliaferro et al., 2013) and hence it may be that the clarity of the background and objectives were somewhat lost in translation.

4. Target population and subgroups

Most studies met this criterion (n=41) by providing detail regarding the base case and subgroup populations, and stating why these groups were chosen. The five studies that failed to meet this criterion stated the populations they used, but provided limited details and lacked justification.

5. Setting and location

Whilst many studies met the requirements for this criterion (n=40), a total of six studies failed to provide sufficient detail to enable the reviewers to identify relevant aspects of the system in which the decision needed to be made.

6. Study perspective

A total of 33 studies described the perspective of the study and related this to the costs. The perspective of the economic evaluation was unclear in 13 studies, which can limit the applicability of the findings.

7. Comparators

The majority of studies (n=41) provided a clear description of the interventions being compared with a justification for their selection. The five studies that failed to meet the requirements for this criterion did provide a description of the interventions but did not supply a reason for their selection.

8. Time Horizon

Most studies (n=40) provided details regarding the time horizon during which the costs and comparisons were being compared, and why this time period was appropriate. Studies that gave the time period (e.g. three years), without stating the year when the intervention was initiated, did not meet the requirements for this criterion. Likewise, studies that did provide details regarding the time horizon, but failed to justify the time period were not allocated a mark for this criterion.

9. Discount rate

A degree of flexibility was afforded to the reviewers for this criterion. During calibration, the reviewers felt that some acknowledgment by authors of the need to discount both costs and outcomes was required, though it became apparent during the review that a number of studies falsely stated that discounting of costs and outcomes were not necessary for their particular study. As the CHEERS checklist is intended to improve reporting quality, rather than methodological quality, studies that reported discounting of both costs and outcomes were deemed to meet this criterion, regardless of whether the discounting was undertaken accurately and appropriately. A total of 29 studies were considered to have reported the discounting of costs and outcomes.

The reviewers agreed that the discounting of costs and outcomes for studies of a duration less than two years was not necessarily required, hence for one study, this criterion was deemed not applicable.

10. Choice of health outcomes

Most studies (n=42) provided sufficient detail regarding the outcomes that were used to measure benefits and their relevance. Of the four studies that failed to meet this criterion, two studies were focused on orthodontic interventions, and two were published in languages other than English.

11. Measurement of effectiveness

A total of 41 studies provided adequate detail regarding the study design, whether this was based upon a single source of data, or a synthesis-based estimate (frequently used in modelling studies). Five studies failed to provide sufficient detail.

12. Measurement and valuation of preference-based outcomes

Only two studies utilised preference-based outcomes, hence for the overwhelming majority of studies (n=44), this criterion was not applicable. The two studies using this approach both provided information regarding the population and methods used to gain preferences for outcomes.

13. Estimating resources and costs

Of the 46 included studies, 38 clearly stated the approaches taken to estimate resource use and the methods used to value these in terms of unit cost. The remaining eight studies failed to provide adequate detail regarding how resource use was determined.

14. Currency, price date, and conversion

A total of 36 studies supplied the dates when resource use estimations and unit costs were identified, and provided sufficient detail regarding conversion of costs into a common currency. This information was not present in the remaining ten studies.

15. Choice of model

Of the 18 studies using a model-based approach, only 13 provided a suitable justification for selection of the model. The remaining 28 studies did not use this approach, and hence this criterion is not applicable.

16. Assumptions

The majority (n=15) of the 18 studies using a decision analytic model provided enough detail regarding the underpinning assumptions. Failure to provide sufficient information regarding the assumptions can cause the study findings to be misinterpreted, as well as difficulties in comparing findings with other studies. This criterion was not applicable for the 28 studies that did not use a model-based approach.

17. Analytic methods

Of the 46 included studies, 33 fully described all analytical methods supporting the evaluation.

18. Study parameters

Just 29 studies reported the values, references and ranges adequately for all parameters.

19. Incremental costs and outcomes

Mean values for the main categories of estimated costs and outcomes were reported in 40 of the included studies, as were the mean differences between the comparator groups. Incremental cost-effectiveness ratios were provided where applicable in these studies.

20. Characterising uncertainty

A full description of the effect of uncertainty was reported in only 26 of the 46 included studies. Of the 20 studies that failed to characterise uncertainty, six were studies using a decision-analytic model approach.

21. Characterising heterogeneity

This criterion was not applicable for 20 studies of the 46 reviewed studies, as their design did not incorporate different subgroups. Of the 26 studies for which this criterion was relevant, only eight reported differences between subgroups of patients relating to heterogeneity. The existence of significant heterogeneity can limit the generalisability of the findings from a study, and hence failure to report this could result in misinterpretation.

22. Study findings, limitations, generalizability, and current knowledge

The reviewers found 36 of the studies met this criterion in full, by summarising the key findings and relating these to the conclusions. These 36 studies also discussed the study limitations to some degree, yet some failed to report all potential limitations. Nonetheless, the reviewers found that ten studies did not mention any study limitations. Failure to highlight study flaws can cause significant misinterpretation of the findings.

23. Source of funding

Of the 46 included studies, the reviewers found 31 studies to meet this criterion, by acknowledging any source of funding, or stating that it was not funded.

24. Conflicts of interest

The poorest performing criterion; just nine studies reported whether a conflict of interest was present or not. The remaining 37 made no comment regarding conflict of interest, failing to acknowledge potential introduction of bias.

Table 4-6: Frequency of studies meeting each Consolidated Health Economic Evaluation Reporting Standards (CHEERS) criteria

Criteria number	CHEERS criteria description	Studies meeting criteria n=46 (%)	Studies failing to meet criteria n=46 (%)	Studies criterion not applicable to n=46 (%)
1	Title	44 (96)	2 (4)	0 (0)
2	Abstract	45 (98)	1 (2)	0 (0)
3	Background and objectives	44 (96)	2 (4)	0 (0)
4	Target population and subgroups	41 (89)	5 (11)	0 (0)
5	Setting and location	40 (87)	6 (13)	0 (0)
6	Study perspective	43 (93)	3 (7)	0 (0)
7	Comparators	41 (89)	5 (11)	0 (0)
8	Time Horizon	40 (87)	6 (13)	0 (0)
9	Discount rate	29 (63)	16 (35)	1 (2)
10	Choice of health outcomes	42 (91)	4 (9)	0 (0)
11	Measurement of effectiveness	41 (89)	5 (11)	0 (0)
12	Measurement and valuation of preference-based outcomes	2 (4)	0 (0)	44 (96)
13	Estimating resources and costs	38 (83)	8 (17)	0 (0)
14	Currency, price date, and conversion	36 (78)	10 (22)	0 (0)
15	Choice of model	16 (35)	2 (4)	28 (61)
16	Assumptions	15 (33)	31 (67)	29 (63)
17	Analytic methods	33 (72)	13 (28)	0 (0)
18	Study parameters	29 (63)	17 (37)	0 (0)
19	Incremental costs and outcomes	40 (87)	6 (13)	0 (0)
20	Characterizing uncertainty	26 (57)	20 (43)	0 (0)
21	Characterizing heterogeneity	8 (17)	18 (39)	20 (43)
22	Study findings, limitations, generalizability, and current knowledge	36 (78)	10 (22)	0 (0)
23	Source of funding	31 (67)	15 (33)	0 (0)
24	Conflicts of interest	9 (20)	37 (80)	0 (0)

4.4.4 Inter-rater and intra-rater reliability

Two reviewers (HR and EV) independently assessed each study against the CHEERS checklist. Whilst differences in scoring were agreed by consensus, the original inter-rater reliability for each CHEERS criterion can be seen in Table 4-7.

Table 4-7: Inter- and intra-rater reliability for quality appraisal using CHEERS checklist

CHEERS criteria	Inter-rater reliability (% agreement)	Reviewer HJR intra-rater reliability (% agreement)	Reviewer EV intra-rater reliability (% agreement)
1	98	100	100
2	93	80	100
3	98	80	100
4	86	100	100
5	86	100	80
6	77	100	100
7	88	100	100
8	84	80	80
9	79	80	80
10	91	80	80
11	93	80	80
12	95	100	80
13	86	100	80
14	79	80	80
15	91	80	60
16	84	100	80
17	72	80	40
18	70	60	80
19	88	80	100
20	86	100	100
21	67	100	80
22	81	100	100
23	95	60	80
24	91	80	80
Overall agreement (%)	**86**	**88**	**85**

Variation in inter-rater agreement can be seen across the criteria. Exceptionally high levels of agreement between reviewers could be seen for criteria 1, 3, and 23. Lower levels of agreement could be seen in criteria 18 (70%) and 21 (67%). Cohen’s kappa (κ) was calculated at 0.7 for overall inter-rater agreement. According to the classifications proposed by Landis and Koch (1977), this is a substantial strength of agreement.

The reviewers (HJR and EV) assessed 10% of the included studies (n=5) a second time to determine intra-rater reliability. These studies were selected randomly using an online random number generator. The studies selected were the same studies also used to determine intra-rater reliability for use of the Drummond checklist. Intra-rater reliability was lowest for reviewer HJR for CHEERS criteria 18 and 23 (both 60% agreement), whilst for reviewer EV, agreement was weakest for criteria 15 and 17 (60% and 40% agreement respectively). Overall intra-rater reliability was slightly higher (κ = 0.72) for HJR than for EV (κ = 0.67), nonetheless the strength of agreement for both reviewers can be deemed as 'substantial' using the aforementioned classification.

4.5 Discussion

This chapter has outlined the findings from a systematic review of the quality and scope of economic evaluations in child oral health research. This discussion will initially reflect upon the key findings arising from this review, before considering the study design and the strengths and limitations. Finally, a series of recommendations are provided, with suggestions for future research efforts.

4.5.1 Reflections on key findings

Scope of economic evaluations

It is clear from this review, as the focus of 91% (n=42) of the included studies, that the prevention and management of dental caries remains at the forefront of paediatric oral health research. This fits the current agenda, as the consequences of a 'western' diet on both oral health and general health have gained further media and political attention in recent years (RCS England, 2015, BSPD, 2015). The need for economic evaluations of interventions for dental caries in children is acknowledged globally, with studies in this review arising from multiple continents. Nonetheless, neither the quantity nor overall quality of economic evaluations in this field is currently sufficient to adequately guide decision-makers.

Measures of benefit

The recent review of economic evaluations in oral health interventions from Hettiarachchi and coworkers (2017) reported an increase in the publication of cost-utility analyses over recent years, yet this was not reflected in the present review. Only two of the 46 included studies in this review used a cost-utility approach, whilst the overwhelming majority of studies conducted a cost-effectiveness analysis (83% n=38).

For the latter, a wide range of natural units were used to report the benefits of interventions. Many were a variation of the DMFT index (Decayed, Missing and Filled Teeth), which has been widely used for over half a century as a means of collecting easily comparable data on caries prevalence and treatment provision from different populations (Klein et al., 1938). Unfortunately, there are so many variations stemming from this index alone, such as the DMFS (Decayed, Missing and Filled Surfaces), DFS (Decayed and Filled Surfaces), DFT (Decayed and Filled Teeth) indices, that comparisons between studies can be incredibly difficult to conduct. With further clinical units used, such as 'number of caries-free teeth', 'number of caries averted' and 'number of caries-free months', drawing comparisons can be an impossible task, preventing data from being maximised through systematic reviews, and ultimately disrupting the dissemination of study findings across the world (Ricketts et al., 2013).

The impact of this variation in reporting clinical outcomes amongst researchers has not gone unnoticed, and has been highlighted by authors of systematic reviews in previous years (Ricketts et al., 2013, Schwendicke et al., 2013). This has led to the initiation of the Outcomes in Trials for Management of Caries Lesions (OuTMaC) study, which aims to develop a core outcome set for trials investigating clinical management of caries lesions in primary or permanent teeth (Schwendicke et al., 2015b). This study is currently in progress, though it intends to use Delphi methods to facilitate panel agreement upon a maximum of seven units in which to measure clinical outcomes for use in this field. It is anticipated that the findings from this study will ultimately improve the measurement of benefits in CEA within child oral health research.

As outlined in the previous chapter, there can be a number of benefits to using CUA over CEA. Of the two studies in this review that did use a cost-utility approach, one generated

QALYs, whilst the other generated the lesser-known Quality Adjusted Tooth Year (QATY). As discussed in section 2.3.1, the QATY was developed as a dental variation of the QALY (Fyffe and Kay, 1992), yet its use within the literature has been minimal due to a number of limitations. Primarily, the QATY cannot be used for all dental interventions, and was not designed for use in the primary dentition. Furthermore, it does not fully acknowledge the strong and important link between oral health and general health. Considering this, along with the benefits of the QALY as detailed in the previous chapter, the QALY remains the primary means of representing strength of preference.

The study using the QALY as an outcome gained utility data through use of the CHU9D (Child Health Utility 9 Dimensions), a generic paediatric multi-attribute instrument, which was developed with involvement of children and young people. The measure is suitable for use with children and young people aged 7 to 17 years, yet the authors of this study altered the questionnaire to enable completion by a parent proxy, so that they could gain data relating to very young children (Koh et al., 2015). Unfortunately, further research has found the CHU9D to be unresponsive to the changing components of dental caries experience, which may limit the applicability of this measure in child oral health research (Foster Page et al., 2015).

Involvement of children and young people

Lack of meaningful involvement of children was a key flaw of the included studies. No single study prioritised involvement of children, and as such, we cannot be confident that the findings from this body of research clearly reflects the issues affecting children’s oral health. As discussed in the previous chapter, the importance of involvement of children in both research and healthcare decisions is increasingly acknowledged, hence there is a need for a substantial improvement within this field. One way to accomplish this would be to gain preferences from children in the development of a utility measure. Whilst this methodology is not yet widely used in healthcare, preliminary research indicates that it is both feasible and reliable (Stevens, 2015).

Methodological quality

The methodological quality of economic evaluations in dentistry as a whole has been reported previously by Tonmukayakal and coworkers (2015). The undertaking of discounting was highlighted as a key weakness of the included studies, with 39% failing to discount outcomes and 21% failing to discount costs. This fits with the present systematic review, with 72% failing to discount outcomes and 35% failing to discount costs. This indicates that, whilst the need to discount costs is acknowledged by most authors, some confusion was evident surrounding the need to discount outcomes, which may be a reflection of the ongoing debate amongst health economists on this topic (Severens and Milne, 2004, NICE, 2012).

The studies in the present review performed better, however, with regard to considering uncertainty, with the majority of included studies undertaking a sensitivity analysis (70%), as opposed to the 47% in the study by Tonmukayakal and colleagues (2015).

Overall, whilst 50% (n=23) of the studies in the present review were judged to be using high methodological quality, it must be noted that the median percentage of applicable criteria met by the included studies was used to determine this category. The median was selected to facilitate comparison with the systematic review by Tonmukayakal and coworkers (2015). Nonetheless, the median score was much lower in this review than in the aforementioned review, hence a study could be deemed to be of high methodological quality by meeting only 50% of the Drummond criteria. Moreover, only two studies met all the relevant criteria, hence the remaining 44 were flawed in some way. The findings from this review have similarities to those of a systematic review of economic evaluations of caries prevention programs by Mariño and coworkers, (2013) who reported inconsistency, confusion in terminology and a lack of sound research methodology. The consistently poor performance of economic evaluation in dentistry with regard to discounting and sensitivity analyses in particular, indicates the need for clear guidance to be targeted at authors publishing in dental journals.

Reporting quality

The overall reporting quality of economic evaluations in the present study was found to be greater than the methodological quality, with a median score of 83%. However, this is less than the median from a recent systematic review of economic evaluations of oral health interventions (92%) by Hettiarachchi and coworkers (2017) which used the same CHEERS checklist. This suggests that reporting of economic evaluations in the narrower scope of child oral health research is of reduced quality than dentistry overall.

The area of least compliance was identified as the reporting of conflicts of interest, with only nine of the 46 studies meeting this criterion (20%). This finding was also reported by the authors of the aforementioned study (Hettiarachchi et al., 2017), indicating that this may be a reflection of dental health economic literature overall. The CHEERS checklist was designed to be used for economic evaluations in the same way that the CONSORT checklist is used for quality appraisal of publications arising from trials (Husereau et al., 2013). Whilst a number of medical journals have openly endorsed the CHEERS checklist, and expect submitting authors to comply with the requirements, this does not appear to be the case for dental journals. Further improvement in reporting quality could be expected following endorsement of a checklist such as this.

4.5.2 Study Design

It should be noted that all of the reviewers were dentally-trained, with an interest in health economics, with additional support provided by health economists where required. The reviewers found the CHEERS checklist to be easier-to-use that the Drummond checklist, which may relate to its similarities with the CONSORT checklist. Nonetheless, the CHEERS checklist contained more criteria that were not applicable to some studies, which can result in difficulties conducting a quantitative analysis. Given the relatively low concordance between the two checklists, there is an indication for both to be used to enable comprehensive appraisal of the methodological and reporting quality of this body of evidence.

4.5.3 Strengths

Novel research

To the researchers’ knowledge, this is the first systematic review to explore both the methodological and reporting quality of economic evaluations in child oral health research. Furthermore, this is the first review to investigate the extent of involvement of children and young people in economic evaluations in this field.

Language inclusivity

This review did not apply any language restrictions, though the studies published in languages other than English were only reviewed by one calibrated reviewer (HJR), working alongside a translator. The translators used were native language speakers, and either dentists or health economists, so the terminology used within the studies was familiar to them.

4.5.4 Limitations

Exclusion criteria

The restrictions on the inclusion of studies published prior to 1997, and modelling studies with a horizon extending into adulthood can be considered a limitation of this study.

It was considered inappropriate to quality appraise studies conducted prior to 1997 against a standard that had not been set at that time. Nonetheless, the inclusion of these studies may have helped to establish the impact of the Drummond criteria on the quality of economic evaluations in this field.

Similarly, this review intended to focus on studies that explored the benefits of interventions gained, and associated costs incurred, solely during childhood. Nonetheless, these modelling studies play an important role in acknowledging that oral health interventions administered during childhood can have benefits (and associated costs) that extend far beyond childhood.

Screening

Title and abstract screening was undertaken by just one researcher (HJR). Screening by a single researcher is not ideal as it may increase the possibility of a potentially relevant study being excluded at an early stage. Nonetheless, the researcher undertaking the screening used an inclusive approach, hence whenever there was any doubt over the relevance of the study being screened, it was retained until the full text screening stage, which may have minimised the potential for error.

Categorisation

The categorisation of studies as being of low, moderate or high methodological or reporting quality could be considered a limitation of this study. This was undertaken in an attempt to support the interpretation of the appraisal scores for each checklist, with the use of tertiles to determine the ‘cut-offs’ for each category, though these could be considered somewhat arbitrary. It is also understood that the application of these same parameters to future updates of this systematic review could be considered contentious as they would be based upon the original dataset. This could restrict the researchers’ ability to compare the categorical results.

Assessment of study design

This systematic review did not assess the methods employed in the underlying study designs of the included studies. There was the potential for the underpinning study methods to have been assessed using well-known tools, such as the CONSORT (Consolidated Standards of Reporting Trials) checklist, which would have added further value to this review, particularly in combination with the quality appraisal of the economic evaluation components.

4.5.5 Recommendations

As anticipated, the majority of economic evaluations in child oral health research are focused on caries prevention and management, confirming that this condition remains at the forefront of the global paediatric oral health agenda. Nonetheless, it was surprising that economic evaluations relating to other common, and potentially burdensome, childhood dental conditions were sparse. It is proposed that molar-incisor hypomineralisation and traumatic dental injuries also present considerable societal and healthcare impacts, and thus should be prioritised for future economic research.

A lack of core outcome measures for studies of caries can result in significant difficulties in the comparison of findings across cost-effectiveness analyses in particular. Unlike other areas of oral health research, there does not appear to be an increase in the publication of cost-utility analyses in child oral health research. Given the unsuitability of the CHU9D for use in this field, this scarcity of cost-utility analyses is likely to be due to the lack of a suitable paediatric preference-based measure. Furthermore, there may be a long-standing confusion amongst dental researchers on whether to generate QATYs or QALYs.

There is substantial room for improvement in both the methodology and reporting of economic evaluations in child oral health research, though there is scope to address this. The endorsement of guidelines such as the Drummond and the CHEERS checklists by dental journals would help to educate authors, and consequently increase the quality of these publications.

An additional concern relates to the lack of involvement of children as active participants in these studies. Whilst significant improvements have been made in other aspects of child oral healthcare, this has not yet translated to economic evaluations in this field. Involvement of children and young people in the development and valuation of a utility measure that could be used in economic evaluations of interventions to improve children's oral health would be an important step forward in addressing this deficit.

This systematic review was conducted to inform the development of this study, hence it was not updated on completion of the wider project. A repeat systematic review in ten years' time would be recommended.

4.6 Conclusion

In conclusion, this systematic review has identified a paucity of high quality economic evaluations in the field of child oral health research. This deficiency could be partly addressed through the endorsement of economic evaluation guidelines by dental journals.

Few cost-utility analyses have been conducted in child oral health research, which may be due to the lack of a suitable instrument with which to generate utilities in this population. This review also highlighted a notable lack of involvement of children and young people as active participants. The development of a utility measure for use in paediatric oral health research may help to address these issues, through facilitating the engagement of children and young people in future economic evaluations of dental interventions, and enabling QALYs to be generated.

The following chapter describes the first stage in the development of this measure, specifically the identification and validation of the classification system.

4.7 Publications arising

4.7.1 Prospective study registration with PROSPERO

Helen Rogers, Zoe Marshman, Helen Rodd, Fiona Gilchrist, J.H (Erik) Vermaire, Katherine Stevens. A systematic review of the quality and scope of economic evaluations in child oral health research. PROSPERO 2017 CRD42017054607 Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017054607.

4.7.2 Published abstract from oral presentation at 2017 International Association of Paediatric Dentistry Congress (Santiago, Chile)

Helen Rogers, Zoe Marshman, Helen Rodd, Fiona Gilchrist, J.H (Erik) Vermaire, Katherine Stevens. A systematic review of the quality of economic evaluations in child oral health research. *International Journal of Paediatric Dentistry* (2017) 27(Suppl. 2): 11–35.

4.7.3 Peer-reviewed journal article

H.J. Rogers, H. D. Rodd, J. H. Vermaire, K. Stevens, R. Knapp, S. El Yousfi, and Z. Marshman. A Systematic Review of the Quality and Scope of Economic Evaluations in Child Oral Health Research. *BMC Oral Health* (2019) 19.1:132. Web.

5 Identification and validation of the classification system

5.1 Background

A preference-based measure (PBM) consists of two key components: a multi-attribute classification system, and an index of health state utility values for all health states described by that classification system (Goodwin et al., 2015a). The classification system for a PBM contains a series of specific items (these are often referred to as dimensions) that are relevant for the population and/or condition in question.

Some preference-based measures are administered directly, such as EQ-5D, where the classification system is completed by participants, and the classification system is identical to the questionnaire. However, for other measures, such as SF-6D, the classification system is derived using responses to a larger non-preference-based measure, the SF-36, that contains a larger set of items than those included in the classification system and that has several different items capturing the same domain. The items included in these measures could be derived through use of so called ‘top down’ approaches, such as reviewing the literature or seeking opinions of expert clinicians in the relevant field (Patrick, 1973, EuroQol, 1990). Alternatively, the views of patients with experience of the relevant condition could be used to identify the items, in what is referred to as a ‘bottom up’ approach (Grewal et al., 2006, Stevens, 2009). For paediatric measures, previous classification systems have been derived from parental opinion, as a proxy for gaining views from children themselves (Stevens et al., 2005). Whilst this may still be considered appropriate in certain circumstances, such as when the population is especially young, or where the condition limits children’s ability to express or communicate their own views, it is now widely acknowledged that the views of children and their parents differ, and that children themselves should be involved (Stevens et al., 2005, Ratcliffe et al., 2012b, Stevens, 2015, Marshman et al., 2015, Ratcliffe et al., 2016b). Some researchers have also used a combination of these contrasting ‘top down’ and ‘bottom up’ approaches (Palfreyman and Brazier, 2012).

The incorporation of patient views in the development of a classification system is considered to offer a number of advantages to a PBM, such as greater face and content validity (Mokkink et al., 2010, Stevens and Palfreyman, 2011). Furthermore, this approach ensures that the language and terminology used is appropriate for the population in question; a feature which is of particular importance when developing a paediatric PBM (Stevens, 2010, Stevens and Palfreyman, 2011). Alternatively, the items can be identified from an existing non-preference-based measure specific to the relevant condition (Brazier et al., 2012b).

The use of an existing non-preference-based measure to generate the classification system for a PBM (typically a measure of HRQoL) offers a number of advantages to other methods of identifying items for inclusion in a classification system. Firstly, the existing measure may have undergone rigorous assessment during its development, and hence its psychometric properties will already have been determined. Secondly, the existing measure may already be widely used by scientific and clinical communities in the respective field (Goodwin and Green, 2016). This means that the new PBM is likely to be well accepted. Furthermore, retrospective economic evaluations can also be conducted by applying the new PBM to HRQoL datasets that have previously been collected using the existing measure (Brazier et al., 2012b, Goodwin and Green, 2016).

The potential benefits afforded by this approach are heavily dependent on the selection of an appropriate measure of HRQoL from which to derive the classification system. The chosen measure should be clearly applicable to both the population and condition in question, and its psychometric properties should be carefully reviewed.

The derivation of a classification system from an existing non-preference-based measure in this way may be considered to be a ‘top down’ approach (Stevens and Palfreyman, 2011). In many cases this would be an accurate description, in that the development of the existing measure may have been informed solely by clinician opinion or the literature. As such, any limitations of the existing measure arising from this, such as poor face or content validity, are in turn inherited by the new PBM. Nonetheless, the use of an existing measure that has in itself been developed using a ‘bottom up’ approach, such as CARIES-QC, could challenge this assumption (Gilchrist et al., 2018).

In order to generate QALYs, a preference-weight or utility value must be assigned to every health state described by the classification system (Goodwin and Green, 2016). However, before a valuation survey can take place to achieve this, the classification system should be validated (Brazier et al., 1999, Brazier, 2017). The use of a qualitative approach for validation of a classification system may be an endorsed approach, but it is not frequently used by researchers in this field (Patrick et al., 2011, Brazier, 2017). This is discussed in further detail in the following section (5.2).

5.2 Background to methodology

5.2.1 CARIES-QC

The decision to select CARIES-QC as the basis from which to derive a classification system for this PBM was taken after a critical review of alternative instruments. As outlined in section 2.2.4, there are a number of measures of OHRQoL designed for paediatric populations that have been used widely by researchers around the world. Nonetheless, these were not developed specifically to capture the impacts of caries and hence may lack the psychometric properties to detect changes in caries status arising from an intervention. Notably, few of these measures have involved children in their development.

As a paediatric measure of OHRQoL specific to caries that was developed using a ‘bottom up’ methodology with involvement of children at every stage, CARIES-QC offered a unique advantage over other potential measures. Furthermore, as it was developed in a UK setting, the features of CARIES-QC are directly relevant to the population in the present study. The psychometric properties of CARIES-QC are also favourable; it has good face, content and construct validity, responsiveness and reliability (Gilchrist et al., 2018).

It is, however, acknowledged that CARIES-QC is a relatively novel measure, and that there are other measures of OHRQoL that are much more established worldwide, with large datasets from varied populations. Nonetheless, global interest in CARIES-QC is certainly growing. At the time when plans for the present study were conceived, CARIES-QC was being used in New Zealand and the Netherlands, with a Dutch translation having been completed (Rogers et al., 2019, Foster Page et al., 2019). Since then, it has been translated into four other languages (Chinese, Arabic, Hindi, Brazilian-Portuguese), validated in a Chinese population and applied in a randomised controlled trial in Australia (He and Wang, 2020, Arrow and Forrest, 2020). Furthermore, its successful use in a number of UK-based studies prompted the use of CARIES-QC as an outcome measure for a large multi-centre trial investigating the use of toothbrushing reminders to improve children's oral health, which was ongoing at the time this thesis was being prepared (AlBader, 2019, Knapp, 2019, Marshman et al., 2019). This growing repertoire of studies and demand for cross-cultural adaptation supports the choice of CARIES-QC as the most appropriate measure on which to base the classification system for the present study. A summary of the advantages and disadvantages of CARIES-QC alongside the alternative measures considered for use in this study is provided in Table 5-1.

Table 5-1: Summary of the advantages and disadvantages of alternative measures considered for use as the basis for a caries-specific preference-based measure, based upon Gilchrist et al, 2016

Measure	Advantages	Disadvantages
CARIES-QC	• Caries-specific • Developed with involvement of children at every stage • Developed in line with COSMIN standards • Relatively short (13 questions)	• Not yet widely used • Requires further testing with larger population to confirm responsiveness
CPQ	• Numerous versions available, including a short form • Good evidence of construct validity • Most widely used of these measures	• Generic measure of OHRQoL • Children not fully involved in item generation, which may affect content validity • Internal consistency unclear, as no assessment using item response theory or factor analysis • Clinical significance of scores unclear • Varied reliability and validity of short form versions
C-OIDP	• Short (only 8 items) so little burden on participants • Useful in epidemiological studies • Children involved in face and content validity testing	• Generic measure of OHRQoL • Adapted from existing Thai version of OIDP
COHIP	• Good evidence of content validity • Involvement of children during item impact stage of development	• Generic measure of OHRQoL • Not widely tested • Children not involved during item

	- Factor analysis used during item reduction	generation - Original version contains 37 questions, which may burden participants

Notes: CARIES-QC: Caries Impacts and Experiences Questionnaire for Children; CPQ: Child Perceptions Questionnaire; C-OIDP: Child Oral Impacts on Daily Performances; OIDP: Oral Impacts on Daily Performances; COHIP: Child Oral Health Impact Profile; OHRQoL: Oral health-related quality of life; COSMIN: Consensus-based Standards for the Selection of Health Measurement Instruments

CARIES-QC is a unidimensional measure containing 13 questions that relate to the items shown in Table 5-2. The final question is a global question to summarise the respondent's impacts and experiences: *'how much of a problem are your teeth for you?'*

CARIES-QC is based on the definition of OHRQoL proposed by Locker and Allen (2007) to encompass "the impact of oral diseases and disorders on aspects of everyday life that a patient or person values, that are of sufficient magnitude, in terms of frequency, severity or duration to affect their experience and perception of their life overall". It was developed using a seven-step methodology proposed by Guyatt and coworkers (1986), involving: item generation; item reduction; questionnaire design; testing of face and content validity, and testing of validity, reliability and responsiveness.

In comparison to other non-preference-based measures of HRQoL, CARIES-QC contains a relatively small number of items (Brazier et al., 2002). There are also three response options (henceforth referred to as levels) for each item in CARIES-QC, that were identified by children during its development (Table 5-2).

To form a PBM it is necessary to gain preference-weights or values for every health state defined by the classification system. Health states are formed through the selection of one level for each item. A greater number of items and levels within a classification system produces more health states that require valuation.

If CARIES-QC was used as a classification system in its original format (excluding the global question, as it is too broad to define a specific dimension), with 12 items and 3 levels, it would contain a total of 531,441 possible health states that would require valuation. This can be calculated using the formula n^r where n=levels and r=repetitions (in this case it would be 3^{12}=531,441). To gain preference weights for such a large number of health states would require an unfeasibly large sample of participants to provide values. Alternatively, each participant would be required to complete a very large number of tasks. Furthermore, previous research suggests that both adults and children are unable to consider more than nine items simultaneously (McCabe et al., 2005, Stevens, 2015). As such, it would be impossible to use CARIES-QC as the classification system in its current form.

Table 5-2: The 13 questions within CARIES-QC, shown here alongside the related items and response levels.

Question from CARIES-QC	Item	Response levels
How much do your teeth hurt you?	Hurt	Not at all A bit A lot
Do your teeth make it hard to eat some foods?	Hard to eat	Not at all A bit A lot
Do you have to eat on one side of your mouth because of your teeth?	One side	Not at all A bit A lot
Do you get food stuck in your teeth?	Food stuck	Not at all A bit A lot
How much do you get kept awake by your teeth?	Kept awake	Not at all A bit A lot
How much do your teeth annoy you?	Annoy	Not at all A bit A lot
How much do your teeth hurt when you brush them?	Brushing	Not at all A bit A lot
Do you have to eat more carefully because of your teeth?	Carefully	Not at all A bit A lot
Do you have to eat more slowly because of your teeth?	Slowly	Not at all A bit A lot
Do you feel cross because of your teeth?	Cross	Not at all A bit A lot
How much have you cried because of your teeth?	Cried	Not at all A bit A lot
Do your teeth make it hard to do your schoolwork?	School	Not at all A bit

		A lot
How much of a problem are your teeth for you?	Global	Not at all A bit A lot

There are three potential ways to reduce the number of health states that could be defined by a classification system (Brazier et al., 2012b). Firstly, the number of items could be reduced. Secondly, the number of levels for each item could be reduced. Lastly, a combination of these approaches could be used.

The current number of levels within CARIES-QC is already considered to be the minimum number of levels for each item. In fact, there is some suggestion in the literature that having as few as three levels within a classification system could limit the sensitivity of an instrument (McDowell, 1996, Harper, 1997, Brazier, 2017). In light of this, it is important that the number of levels is not reduced further, hence the only remaining option to reduce the number of health states defined by the classification system, in the case of CARIES-QC, would be to reduce the number of items.

Numerous approaches to achieve this have been proposed and utilised, though essentially the process involves the identification of one item (or sometimes two items) to represent each dimension represented within the measure (Young et al., 2009, Brazier et al., 2012b). The selection of these items is key to the sensitivity of the PBM, and should aim to minimise the loss of descriptive information.

It has been suggested that the first stage in this process should be the establishment of domains within the existing non-preference-based measure, where these have not been proposed by the original authors (Young et al., 2009). No formal domains were outlined to categorise the items within CARIES-QC during its development so it was necessary to determine these through the use of factor analysis.

The next stage of this proposed process involves the use of Rasch analysis and classical psychometric testing as a means of identifying which items to eliminate and which to retain within the classification system (Young et al., 2009). The Rasch model is a mathematical model first proposed by Georg Rasch (Rasch, 1960). It has been used predominantly in education research, though more recently has been applied to health sciences, particularly in the development of patient-reported outcome measures (PROMs) (Tennant and Conaghan, 2007). Rasch analysis allows the comparison of responses from an outcome scale against the expectations of the model. The models expectations are based upon a probabilistic form of Guttman scaling which, when applied to health outcomes, is a logistic function of the difference between a person's severity level and the severity level of the item (Twiss et al., 2016). The items can be considered to fit the model if the observed data does not deviate significantly from the expected responses.

The use of Rasch analysis in PROM development allows the new measure to be designed to fit the models expectations from the outset. This approach was used in the development of CARIES-QC (Gilchrist et al., 2018), allowing confirmation that all items within the measure have an acceptable fit to the model.

Further to its application in the development of a measure, it is also a useful tool for evaluating the psychometric properties of an existing ordinal scale or PROM (Tennant and Conaghan, 2007). For the purposes of this study, the application of the Rasch model was not to determine whether the items in CARIES-QC had an acceptable fit to the model (as this had already been confirmed during its development), but to use a more discriminatory approach, such as using the 5% significance level rather than a Bonferroni adjustment, to determine which items fit the model best. These items would be then considered as stronger candidates for inclusion within the classification system.

Similarly, classical psychometric testing was also conducted to evaluate the overall functioning of the newly-developed CARIES-QC, though it remains necessary to review this again on an individual item level, with particular regard to the feasibility, internal consistency, response distribution and responsiveness of each item. As with the Rasch analysis, those items with more beneficial properties would be more appropriate to include within the classification system.

A number of condition-specific preference-based measures (CSPBMs) have been developed using a similar approach to select the classification system, incorporating Rasch analysis and classical psychometric analyses, but with the addition of input from the developers of the original measure of HRQoL (Young et al., 2011, Mulhern et al., 2013b, King et al., 2016, Mulhern et al., 2017). The inclusion of the developer's opinion can provide valuable insight into the functioning of each item, particularly where a qualitative approach was used in the identification of items for the original measure, as with CARIES-QC. However, a more novel addition to this multi-staged approach is the inclusion of the opinions of patients and the public. This was considered to be particularly important for the present study, given that the measure to be created is intended for a specific population and condition. More detail of the involvement of patients and the public for other aspects of this stage is provided in section 5.5.7.

5.2.2 Validation approaches

Validation of the classification system is recommended prior to proceeding with a valuation survey (Brazier et al., 2012b). Existing guidance on the validation of a classification system for a CSPBM advises repeating the same analyses that were conducted in identifying the classification system once again, using either alternative data from the same dataset, an alternative timepoint from the same dataset, or an alternative dataset altogether (Brazier et al., 2012b). This guidance contains no mention of using qualitative techniques to validate a classification system. In such a rapidly advancing field with growing emphasis on involvement of patients and the public in research, these recommendations could be considered somewhat outdated. In contrast, the recommendations from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) on the development of patient-reported outcome instruments strongly endorse the use of qualitative approaches such as Cognitive Interviewing, as being essential to the evaluation of content validity (Patrick et al., 2011).

In recent years, there appears to have been an increase in the use of qualitative techniques in the development of PBMs. Nonetheless, primarily these techniques have been used as a means of identifying a pool of items from which the classification system could be determined (for a ‘de novo’ PBM), rather than as a way of validating the choice of items for a classification system from an existing measure (Conway et al., 2011, Sutton and Coast, 2014, Powell et al., 2019).

The use of Cognitive Interviewing (CI) techniques can provide a unique insight into the functioning of a questionnaire, by allowing the researcher access to the respondent’s thoughts during completion (Patrick et al., 2011). The benefits of using it to validate a classification system are two fold; firstly, to determine the participant's comprehension and interpretation of the items in relation to their intended meaning, and secondly to identify whether any important items have been left out (Patrick et al., 2011). This is the primary justification for the validation of a classification system with participants from the relevant population group, with experience of the condition in question. For the present study, it was essential to involve children and young people who have experienced dental caries in the validation process.

The term CI encompasses a range of different techniques, including concurrent probing (questions asked during each item response), retrospective probing (questions asked after all item responses) and concurrent verbalisation ('thinking aloud' during each item response) (Willis, 2005, Pepper et al., 2018). The latter approach was primarily employed by the present study, given that it is less disruptive to task performance and less reliant on participant recall than the former techniques (Willis, 2005). Furthermore, it has been used successfully with children and young people in previous research in education and healthcare respectively (Vandevelde et al., 2015, Joffer et al., 2016).

Whilst cognitive approaches enable the interviewer to access the thoughts of the participant whilst they complete the questionnaire, for the present study it was necessary to complement this with semi-structured interviews also. This combination of techniques allowed greater exploration of children and young people's views surrounding the items.

Qualitative research typically adopts one of a broad range of methodologies that are linked to differing beliefs, theories and contexts. Each approach has implications for the inductive and deductive processes used, as well as the way in which the data are analysed and interpreted (Ritchie et al., 2014). As such, the methodology chosen should be suited to answering the research question, avoiding confinement to specific traditions aligned with a particular epistemological stance (Ritchie et al., 2014). In turn, through using the approaches described above, a pragmatic stance was considered to be most appropriate for the present study.

5.3 Aim

This part of the study aimed to identify and validate a classification system based upon CARIES-QC that would be suitable for use in a valuation survey. The following specific objectives were set:

1. To identify a preliminary classification system from CARIES-QC suitable for use in a valuation survey, using Rasch analysis, classical psychometric tests, involvement of children and parents, as well as the team who developed CARIES-QC
2. To use a qualitative approach to validate the preliminary classification system with children and young people who had a diagnosis of dental caries

5.4 Method: identification of the classification system

The classification system was derived from CARIES-QC using a combination of Rasch Analysis, classical psychometric testing, Patient and Public Involvement (PPI), and developer opinion.

The first two of these approaches utilised a dataset that was collected from the initial validation study of CARIES-QC (Gilchrist et al., 2018), which is described in more detail below. The subsequent approaches did not require the use of a dataset. Findings from each approach were then discussed by the study team, particularly where stakeholder views were found to conflict with the results of statistical analyses. Where this occurred, agreement was sought by consensus on which items should be selected for inclusion in the preliminary classification system.

5.4.1 Dataset

The dataset used in the Rasch analysis and psychometric tests in the present study were originally collected for the validation of the CARIES-QC measure in 2014 and 2015 (Gilchrist et al., 2018). Data were obtained from a sample of 200 children recruited from new patient clinics at a dental hospital (Charles Clifford Dental Hospital) and community dental site (Firth Park Salaried Dental Service) in Sheffield, UK. Children were aged 5 to 16 years and had a diagnosis of active dental caries. Children were asked to complete the CARIES-QC measure at three different timepoints: baseline (T0), prior to the start of treatment (T1) and following a course of dental treatment to manage the caries (T2). Sociodemographic data were collected, including age, gender, ethnicity and postcode. Postcode data were used to determine the participant's level of deprivation, using the Index of Multiple Deprivation tool (Index of Multiple Deprivation, 2015). A range of clinical data were also collected to establish the number of decayed, missing and filled teeth in the primary or permanent dentition (dmft/DMFT), the presence of caries in the anterior teeth, whether the child experienced pain from their teeth, and whether the caries had extended to involve the pulp. (Gilchrist et al., 2018).

All analyses conducted on this dataset used the T0 baseline timepoint, as this had the highest number of observations, with the exception of responsiveness testing which utilised data from the T0 and T2 timepoints. was sufficient. A formal sample size calculation was not undertaken, though the literature suggests that Rasch analysis is known to be sensitive to larger sample sizes, with very large samples causing an increase in the frequency of statistical significant findings, causing difficulties in item reduction (Tesio, 2003, McTaggart-Cowan et al., 2010). Whilst a sample of 200 is relatively small compared to those that have been used in the development of other HRQoL instruments and PBMs, which have seen samples with around 400 to 700 participants being used successfully, it is still considered to be sufficient (Young et al., 2009, McTaggart-Cowan et al., 2010). Sample adequacy for factor analysis was confirmed prior to conducting the psychometric tests, through use of the Kaiser-Meyer-Olkin test.

Error! Reference source not found.

5.4.2 Rasch analysis

Prior to conducting the analyses, a likelihood-ratio test was undertaken to determine whether the rating scale or partial credit model would be the most appropriate version of the polytomous Rasch model to use. Whilst these are similar, there are slight differences in that the rating scale model expects the distance between thresholds to be the same for all items, whilst the partial credit model does not (Tennant and Conaghan, 2007).

For the purposes of this study, the Rasch analysis focussed on the spread of items across the three levels (response categories) at logit 0 (as this is where the Rasch model is centred), using threshold probabilities, whereby a greater spread indicated the respondent was able to distinguish between the item levels (Young et al., 2011). Item (χ^2) goodness-of-fit statistics were also conducted, with the items having the best fit to the underlying model (and hence having a high χ^2 value and a low non-significant P-value) being the best candidates for inclusion in the classification system (Young et al., 2011). Fit residuals were identified, to provide a summation of individual person and item deviations from the model (Pallant and Tennant, 2007). Items with residuals closer to 0 indicated a better fit to the model and hence possible candidates for inclusion in a classification system. Conversely, items with values above +2.5 were considered to misfit the model and items with values below -2.5 were considered to be redundant (Pallant and Tennant, 2007). Differential Item Functioning (DIF) was also assessed to determine whether each item was working the same across respondents of different age groups (5-7 years, 8-12 years, 12-16 years), genders, ethnicities (white British and non-white British) and levels of deprivation according to Index of Multiple Deprivation (IMD) quintiles (Young et al., 2011, Index of Multiple Deprivation, 2015). Analysis of variance (ANOVA) statistics were used to compare scores across each of these factors and across each class interval. Items with DIF, either uniform (constant across attribute levels) or non-uniform (varying across attribute levels) were considered for exclusion from the classification system (Young et al., 2011).

Disordered thresholds can arise when respondents are inconsistent in their selection of response levels, which can occur as a result of having too many response levels, or unclear labelling for the response levels (Ramp et al., 2009). Disordered thresholds were not anticipated given that the original CARIES-QC measure was developed with Rasch analysis, though it was necessary to confirm this prior to item selection. Further to this, an assessment of local dependencies between items was conducted through use of the residual correlation matrix, with positive correlations of 0.2 or more being indicative of response dependency (Christensen et al., 2016).

Rasch analysis was conducted using the Rasch Unidimensional Measurement Model RUMM2030™ software (©Rumm Laboratory Pty Ltd, Version 5.3). The researcher who undertook this analysis (HJR) attended a two-and-a-half day Introduction to Rasch Analysis workshop led by the Psylab Group at Leeds University to gain training on how to use this software.

5.4.3 Classical psychometric testing

The global question from CARIES-QC was excluded from all psychometric analyses as the study team agreed from the outset it was not appropriate to include within the classification system.

In line with recommendations in the literature, initial factor analysis was conducted to establish the dimensional structure of CARIES-QC. This approach looks for patterns of correlations within a set of observations to identify structurally independent dimensions (Chatfield and Collins, 1981, Young et al., 2011). Principle component analysis was undertaken, using the varimax rotated component matrix to identify domains. The suitability of the sample for use of factor analysis was confirmed using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity to establish the presence of correlations between the items. Factor loadings of more than 0.3 were considered to be contributing significantly to the factor, as recommended in the literature (Marchioni et al., 2005).

This was followed by four classical analyses in line with other studies of this type (Young et al., 2009, Young et al., 2011). These analyses were conducted using baseline data (T0 timepoint) with the exception of the responsiveness test, as described below.

Firstly, analyses to determine the rate of missing data were undertaken, to evaluate item feasibility. Items with more than 5% missing data were considered to be poor candidates for inclusion within the classification system (Schafer, 1999).

Internal consistency would usually be determined by comparison of the item with its respective domain score, though in the absence of established domains, correlations between each item with the global question and total score were determined using Spearman's correlation coefficient. Furthermore, correlations between items were assessed to identify items that were capturing the same aspect of quality of life, where one of the items may be selected in the classification system to reflect the wider set of items.

The distribution of responses was also analysed. Floor and ceiling effects were deemed to be present if more than 15% of participants chose the best ('not at all') or worst ('a lot') responses (Terwee et al., 2007). It was acknowledged, however, that in a measure with only three response options, most of the items would have some degree of a floor or ceiling effect, or both. Items with strong floor effects were considered to be poor candidates for the classification system given that they would not be able to capture a deterioration in health. Conversely, items with strong ceiling effects were considered for selection as this suggested an ability to capture the impacts of higher disease severity.

The responsiveness of each item, defined as its ability to detect a change when it has occurred, was estimated using the Standardised Response Mean (SRM) in line with similar studies (Angst et al., 2008, Young et al., 2011). This was determined to be the most appropriate indicator of effect size given the presence of a correlation greater than 0.5 (Pearson correlation coefficient = 0.529) between baseline (T0) and follow-up (T2) scores (Norman, 2014). The SRM (also known as Cohen's *d*) was calculated by dividing the mean score change (follow-up score (T2) minus the baseline score (T0)) by the standard deviation of the change (Norman, 2014). The SRM were interpreted using Cohen's criteria, whereby an SRM of <0.2 is deemed inconsequential, 0.2-0.5 is considered small, 0.5-0.8 is considered moderate and above 0.8 is considered large (Cohen, 1992, Durlak, 2009). A higher SRM indicated greater sensitivity to change.

Classical psychometric analyses were carried out using SPSS® software (IBM Corporation, United States, Version 25).

5.4.4 Views of patient and parent representatives

A panel of children and young people comprising personal contacts, local schoolchildren and patients from a paediatric dental clinic were invited to give their views at one of two informal meetings, held on 10th May 2017 and 9th June 2017, to determine their views on the items within CARIES-QC. The panel was comprised of children from a range of ages, genders and ethnicities, with differing experiences of dental caries. The meetings were moderated by HJR, whom, as a paediatric dentist, had experience in communicating with children. These discussions were focused on how important each item was felt to be, whether any items were considered to overlap, and whether any items were felt to be too similar. Two parent representatives were also involved in these discussions, to provide their thoughts on the items within CARIES-QC from their perspective. Written notes were taken by HJR throughout these meetings. Further details of the panel and the parent representatives are provided in section 5.5.7.

5.4.5 Developer opinion

The final stage of this process involved an informal meeting held on 9th May 2018 with researchers who had led the development of the original CARIES-QC measure. It was considered important to acknowledge any issues or concerns identified by the research team during the development of this instrument, particularly since children were involved throughout this process. Furthermore, it was essential that any issues surrounding the use of the instrument in different settings and languages were taken into account. Written notes were taken by HJR throughout this meeting.

5.4.6 Discussion to determine the preliminary classification system

The findings from the aforementioned four steps were discussed at a meeting which involved clinicians (specialists and consultants in paediatric dentistry), a senior health economist and researchers who led the original development of CARIES-QC. The results from each of the four approaches were deliberated for each item, weighing up the advantages and disadvantages of each. Where the statistical approaches (Rasch analysis and classical psychometric testing) generated results that conflicted with the stakeholder views (PPI and developer opinion), a consensus was sought from the members of the study team. Each approach was weighted equally (i.e. no single approach provided results that were valued more highly than another). The outcome from this meeting was an agreed preliminary classification system.

5.5 Method: validation of the classification system

Before the preliminary classification system could be deemed suitable for use in a valuation survey, it was necessary to validate it with a sample of children who had experience of the condition in question to ensure it remained relevant. As such, the classification system was shown to children and young people who had a diagnosis of dental caries. A qualitative approach was used to determine whether the most important items had been included and to ensure that the refinement process had not affected the face or content validity of the instrument. The reporting of this qualitative stage of the study was undertaken in accordance with the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist (Tong et al., 2007).

5.5.1 Setting

Potential participants were identified via referral letters to the Paediatric Dental Department at the Charles Clifford Dental Hospital, Sheffield. The vast majority of patients seen in this department have been referred by their general dental practitioner for the management of dental caries. Referred patients are given a 'new patient' appointment, where they undergo a dental assessment, radiographs are taken if necessary, and a treatment plan is devised. Patients attend these appointments from a wide geographical area, including South Yorkshire, Lincolnshire, Nottinghamshire and Derbyshire.

5.5.2 Participants

The inclusion and exclusion criteria detailed below were applied when considering potential participants for this stage.

Inclusion criteria

- Children aged 5 to 16 years;
- Children with active dental caries, including if in conjunction with Molar Incisor Hypomineralisation (MIH);
- Children who were able to understand spoken English, i.e. able to understand and undertake the research with support;
- Children with parents/ caregivers who were able to understand spoken English, i.e. able to understand the research with support.

Exclusion criteria

- Children with caries in conjunction with dental anomalies such as amelogenesis imperfecta or dentinogenesis imperfecta;
- Children who were unable to understand and undertake the research even with support;
- Children with parents/caregivers who were unable to understand the research even with support.

5.5.3 Sample size and sampling methods

A maximum variation purposive sampling approach was used, guided by the sampling framework outlined in Table 5-3, to facilitate a range of ages, ethnicities and gender (Sandelowski, 1995). This was important to ensure a broad range of views were obtained from participants from different backgrounds and with varied experiences. A similar sampling approach was used in both formulating the classification system and testing the content validity of CARIES-QC (Gilchrist, 2015). In line with this previous research, it was expected that approximately 20 interviews would be required to reach data saturation. Data saturation was considered to have been met when nothing new emerged from the interviews; a point of so called ‘informational redundancy’ (Sandelowski and Given, 2008).

Table 5-3: Purposive sampling framework for validation of the classification system

Ethnicity	**White British**		**Non-White British**		
Age (years)	**Male**	**Female**	**Male**	**Female**	**Total**
5-10	2	2	2	2	8
11-16	3	3	3	3	12
Total	5	5	5	5	20

5.5.4 Participant recruitment

Potential participants, who met the aforementioned eligibility criteria and purposive sampling requirements (Table 5-3) were approached on ‘new patient’ clinics at the Charles Clifford Dental Hospital (CCDH) paediatric department between 1st October and 30th November 2018. A poster explaining the study in simple terms was placed in the waiting area to inform potential participants in the first instance, and advising them that they may be approached during their visit. Following their initial examination with the dentist, HJR outlined the purpose of the study in simple terms to potential participants and their parents. Children and their parents were made aware that they would receive a shopping gift card to the value of £5 on completion of the interview in acknowledgement of their time and commitment, in line with INVOLVE recommendations (INVOLVE, 2020).

Information sheets containing further detail about the study were given to children and their parents to read whilst they were waiting to have their radiographs taken. On their return from radiography, children and their parents were asked if they wanted to take part, or whether they needed more time to consider their participation. In the department, children with caries are typically offered a prevention appointment, where oral hygiene and dietary advice are provided rather than any invasive dental treatment. This prevention appointment presented a second opportunity to approach potential participants who required further time to consider whether they wanted to take part.

Written consent was sought from the parents of participating children, whilst children and young people completed an age-appropriate assent form. All participant-facing documents (seen in appendices G, G and I) were developed with involvement of the Young People Panel and parent representatives for the wider study, to ensure the language used was both age-appropriate and unambiguous. The recruitment process for this stage of the study can be seen in Figure 5-1.

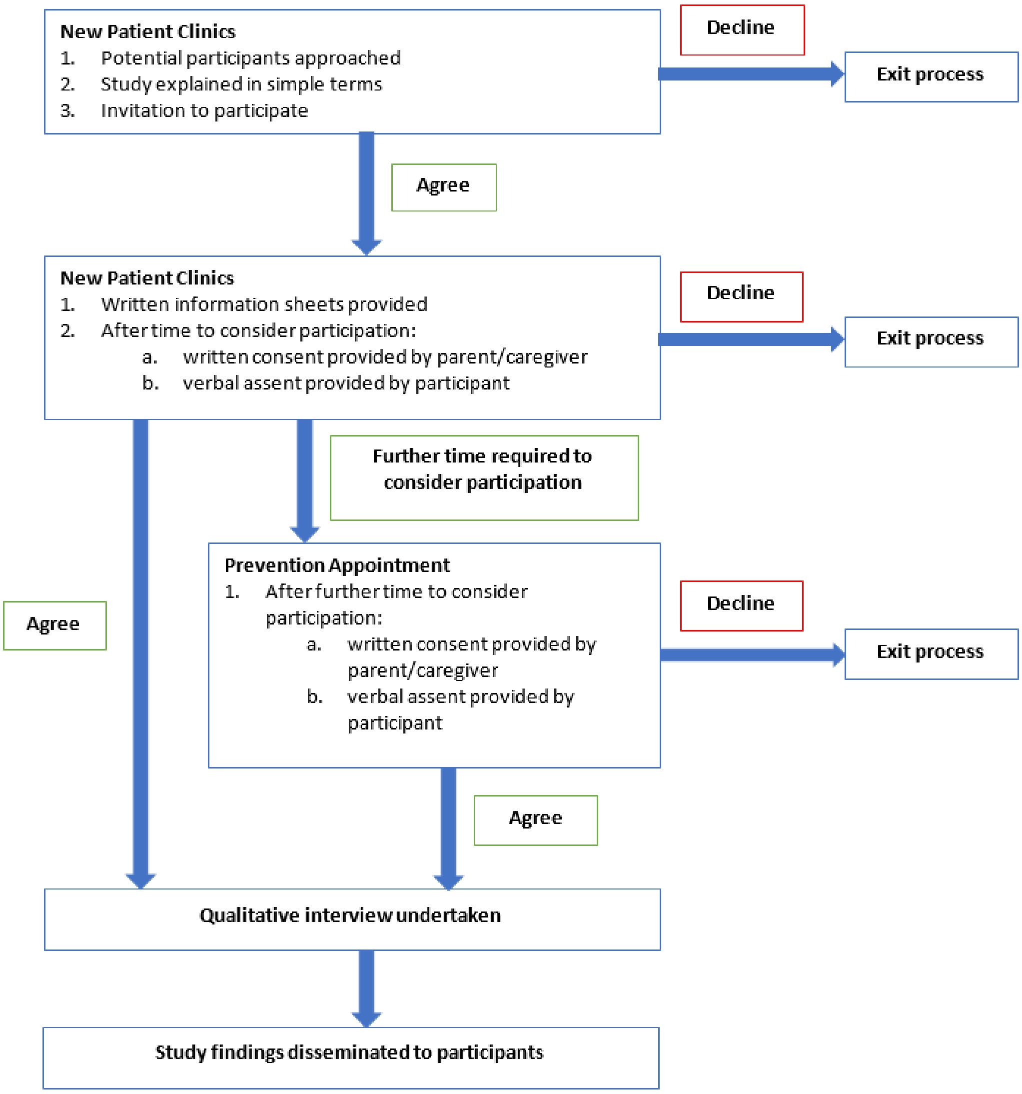


Figure 5-1: Flowchart displaying participant recruitment process for validation of the classification system

5.5.5 Data collection

Qualitative, semi-structured interviews were conducted with participants in a quiet room, separate from the main paediatric dental clinic. Interviews were conducted by HJR, a female paediatric dentist with some previous experience and formal training in qualitative techniques, provided by the Social Research Association, London, UK. HJR introduced herself to participants as a ‘student’ or ‘researcher’ doing a project about teeth, and wore casual attire to minimise the possibility of her status as a clinician from influencing children’s responses (Harden et al., 2000). A 'think aloud' approach was used initially (as discussed in section 5.2), whereby participants were encouraged to talk through their thought process whilst they completed a reduced version of the CARIES-QC questionnaire (Willis, 2005). This was comprised of the questions from CARIES-QC relating to the five items in the preliminary classification system. Participants were then given the seven questions from CARIES-QC relating to the items that had not been included in the preliminary classification system. Similar items were placed next to each other for the purposes of the validation study to avoid participants needing to switch back and forth between different concepts (Patrick et al., 2011).

Children were prompted to discuss their thoughts whilst completing the items, and give their opinion on the items themselves. Following this, the interviewer asked the participant a series of questions guided by a topic guide (appendix J) which was developed iteratively. The questions were designed to explore whether the five items selected for the classification system included the most important questions, or whether any that had been removed should be reinserted. Further questioning investigated children's preferences regarding the wording of similar items, such as those regarding the impacts of eating. Participants were also asked whether they thought any further questions could be removed, whilst ensuring the questionnaire still made sense.

Children and young people were advised that they could ask for support with reading or filling in the questionnaire at any time, and could stop the interview at any point without any repercussions, in accordance with the ethical considerations discussed in section 5.5.8. All interviews were recorded using an Olympus WS-853 digital voice recorder and field notes were taken where required. The duration of each interview was also documented.

On completion of the interview, the clinician who assessed the participant during their appointment was asked to complete a short data collection form. This documented the participants' age, gender, and ethnicity; details of the latter were based upon information provided by the parent upon registration at the hospital using the ethnic category codes used within the NHS (NHS Digital, 2020). Participants' full postcodes were also recorded to enable a calculation of Index of Multiple Deprivation (IMD) scores to determine the national deprivation quintile of the area in which they reside (Index of Multiple Deprivation, 2015).

Clinical data were also collected in the form of the dmft index, which combines the number of decayed (d), missing (m) and filled (f) teeth (Klein et al., 1938). This was assessed using a combination of findings from the clinical examination as well as any appropriate radiographs that were available. Caries was deemed to be present in any tooth with caries extending into dentine either clinically, radiographically or both. The dmft score (and corresponding DMFT score for the permanent teeth) was determined by both the assessing clinician and HJR, with any disagreement resolved by discussion.

5.5.6 Analysis and interpretation

Descriptive statistics were undertaken for socio-demographic data using SPSS® software (IBM Corporation, United States). Further descriptive statistics were undertaken to determine deprivation levels within the sample, according to IMD scores derived from participant postcodes.

All interview recordings were transcribed verbatim. Transcription of the first three interviews was undertaken by HJR, to gain an understanding of the processes involved. The remaining transcriptions were completed by an external company (©Dictate2us Ltd., Manchester, UK). All transcripts produced externally were checked by HJR alongside the original digital voice recordings for accuracy. Data from each participant were analysed contemporaneously. Qualitative data were organised using a Framework approach in NVivo 12 software (©QSR International Pty Ltd., Chadstone, Australia) and analysed thematically by two researchers (HJR and ZM) independently, followed by a discussion to agree upon the key themes (Ritchie et al., 2014). HJR had limited previous experience in qualitative analysis, though had received formal training on qualitative analysis from the Social Research Association, London, UK. Researcher ZM had extensive experience and expertise in this field. The analysis focused on identifying children's level of understanding for each item, the amount of importance participants placed upon each item and whether they considered any as redundant or overlapping. The topic guide was modified as required to guide the interviewer in eliciting further information surrounding emerging themes.

The data synthesised from this stage were discussed by all members of the study team at a meeting held on 5th December 2018. The discussions aimed to determine whether the preliminary classification system derived through the processes described above, had retained the face and content validity of the original measure, or whether further refinements were necessary. Further discussions were held with young patient and public involvement (PPI) representatives (see section 5.5.7 for more detail) to ensure the data had been interpreted appropriately and whether any suggested modifications were suitable. Any modifications to the items within the preliminary classification system deemed appropriate by both the study team and the PPI representatives were then undertaken.

The wording of the health state descriptors pertaining to these items was then determined with the involvement of PPI representatives to ensure they used the same tense, and were easy to read. The final classification system with health state descriptors could then be confirmed.

5.5.7 Patient and public involvement

Patient and public involvement (PPI) was a priority throughout the research described in this chapter, as well as the wider study. Funding to support PPI within this study was provided through a grant from the Research Design Service Yorkshire and Humber (October 2015). This funding was used to form the Young People Panel; a group of children and young people of different ages, genders and ethnicities, as well as differing experiences of caries, whom agreed to provide their views on various aspects of the study as required. Some members of the panel were personal contacts of the researchers involved with the study, whilst others were identified from a clinical setting. All members of the panel had parental approval to assist with the study.

Communication with panel members took place either face-to-face, or via their parents email address, which was provided at the initial meeting. Face-to-face meetings were arranged to minimise the burden on panel members e.g. for panel members attending the dental clinic for an examination or treatment, a meeting was held directly before or afterwards as appropriate. The Young People Panel initially comprised of ten children, though a number of members left the panel over the course of the study, for reasons such as upcoming exams at school, whilst others could no longer be contacted by email. The Young People Panel was supplemented by other involvement of other children and young people throughout the study as required.

Two parent representatives (one mother, one father) also agreed to be involved in the study. The parents were both personal contacts of members of the research team, though had no clinical background. Each parent had two children, one of whom had experience of dental caries. The parent representatives continued to be involved throughout the duration of the study.

Panel members and parent representatives were provided with £5 shopping gift vouchers to thank them for their time after contributing to each part of the study, in accordance with INVOLVE guidelines (INVOLVE, 2020).

5.5.8 Ethical Approval

Ethical approval for this part of the study was obtained from Yorkshire and the Humber Research Ethics Committee (Reference: 18/YH/0148). The details of this are provided in appendix K. This stage was not considered to raise any significant ethical concerns other than those inherent to conducting health research with children and young people. Nonetheless, the following specific ethical issues were given due consideration as detailed below.

Informed consent

Parental information sheets and age-appropriate child and young person information sheets were developed with involvement of PPI representatives. Further information about PPI can be found in section 5.5.7

Information sheets were given to potential participants and their parents at their new patient appointment for them to read whilst waiting to have radiographs taken, hence not unnecessarily prolonging their visit. On returning from radiography, children and parents were asked whether or not they would like to participate, or whether they needed more time to consider. Previous experience in this setting suggested that most parents and participants were happy to provide consent at this appointment, without requiring additional time.

Due to the age of participants involved in this part of the study, informed consent was requested from participant’s parents or carers. Nonetheless children and young people were also asked to assent to take part. Consent or assent was required from both parties for the child to be able to participate.

Potential participants were advised that the decision to participate was entirely theirs. It was also made clear they may withdraw from the study at any point, how to do this, and that this would not make the researcher, or their dentist, upset with them (Hurley and Underwood, 2002). Furthermore, they were advised that there was no requirement for them to justify this decision and there would be no consequences.

Participant distress

As a paediatric dentist, the researcher (HJR) conducting the interviews had extensive experience in interacting and communicating with children and young people. Whilst the topics discussed during the interviews were not expected to cause distress, it was important that the researcher was able to determine whether the participant's body language indicated that they were unhappy or uncomfortable with any element of the research, and communicate with them accordingly (Kirk, 2007).

Confidentiality

To ensure participant confidentiality, all interviews were conducted in a quiet room away from the main dental clinic to ensure that they were not overheard by others.

Participants were anonymised through the assignment of a participant identifier. No participant identifiable information was recorded during the interviews. Data from the digital voice recorder were held on a secure network accessible only by the direct research team, via a password-protected computer in a secure room in the dental school. All data in paper format, including field notes, completed questionnaires and socio-demographic data collection forms were stored in a locked drawer, in a locked room within the dental school, accessible only by the direct research team. It was agreed that all research data generated by this stage of the study would be kept for five years after all publications arising from this work had been accepted, and then destroyed. Pseudonyms or participant numbers were assigned whenever direct quotes from participants were used during dissemination of the study findings.

Parents were not given access to the responses that their child gave during the qualitative interviews. Children could be harmed or embarrassed if their confidentiality was breached in this way. Furthermore, this could negatively affect their ability to trust researchers and participate in future studies.

Safeguarding children and young people

Children's safety was a priority, superseding the need to maintain confidentiality in this study. Participants and their parents/carers were advised that the researcher would act in the best interest of the child should any safeguarding concerns arise during this part of the study. The researcher who conducted the interviews with children and young people (HJR) had completed Level 3 training in Safeguarding Children and Young People and maintained the relevant skills and knowledge by undertaking annual training updates (RCPCH, 2019). There were already robust processes in place in the paediatric dentistry clinic and the wider Sheffield Teaching Hospitals NHS Foundation Trust to ensure that any concerns about a child were appropriately managed. The lead researcher (HJR) was already familiar with department and trust policies relating to safeguarding children and young people who may be subject to potential neglect or abuse. Further to these, a project-specific protocol was developed (appendix L) for this part of the study, to guide the researcher should a participant have disclosed anything that raised concern. This pathway also provided clarification on how to manage a non-safeguarding issue, such as if a child participant had concerns about their treatment, or was being bullied because of their dental appearance. A copy of this pathway was maintained in the study site file.

Children were provided with contact details for the research team in case they had any concerns or issues to discuss between the hours of 9am and 5pm. This ensured that the research team would not be required to manage an out-of-hours safeguarding issue without having the support/advice from the Safeguarding teams that is available during normal working hours.

Protection from harm

The researcher that had direct contact with participants (HJR) held the appropriate level of disclosure from the Criminal Records Bureau, UK.

5.6 Results: Identification of the classification system

The findings from the four approaches to identify which items to select for the classification system are provided below, followed by a summary of the discussions held by the research team to agree upon the preliminary classification system. The results from the qualitative study to validate this classification system are detailed subsequently.

The first two of these approaches involved analyses based upon existing data derived from the original CARIES-QC validation study, which was described earlier in section 5.4.1. The sociodemographic characteristics and caries experience of the participants in this dataset are provided in Table 5-4.

Table 5-4: The sociodemographic and clinical characteristics of participants in the dataset on which Rasch analysis and classical psychometric testing was conducted, adapted from Gilchrist et al. (2018)

Participant characteristics	**Original CARIES-QC validation study T0 baseline *n*=200 (%)**
Age (years)	mean: 8.1; range: 5-16
Gender	
Male	95 (47.5)
Female	105 (52.5)
Ethnicity	
Asian background	31 (15.5)
Black background	5 (2.5)
Mixed background	9 (4.5)
White British background	130 (65.0)
Other background	9 (4.5)
Unknown background	16 (8.0)
Deprivation quintiles	
Most deprived	119 (59.5)
More deprived	37 (18.5)
Average	20 (10.0)
Less deprived	13 (6.5)
Least deprived	11 (5.5)
Total dmft*	mean: 6.24 (SD: 3.45); range: 0-16
Total DMFT*	mean: 1.57 (SD: 2.18); range: 0-13

Notes: *dmft indicates the number of decayed, missing and filled teeth in the primary dentition, whilst DMFT indicates the number of decayed, missing and filled teeth in the permanent dentition

5.6.1 Rasch Analysis

Rasch analysis was conducted using the partial credit model. Overall, the CARIES-QC data were found to have a good item (mean 0.385 ± 0.902) and person fit (mean 0.254 ± 0.999) to the Rasch model, as the means were close to 0 and the standard deviations close to 1.

A summary of the key results from the Rasch analysis for each item is provided in Table 5-5. The items with the highest spread across the three levels at logit 0 were *'food stuck'* (1.632), *'hurt'* (1.605), *'hard to eat'* (1.585) and *'cried'* (1.466) respectively. Those with the lowest item spread, and hence candidates for exclusion from the classification system, were *'cross'* (0.705), *'one side'* (0.858), *'school'* (0.894) and *'brushing'* (0.913) respectively.

Regarding goodness-of-fit, the items '*food stuck*' and *'annoy'* did not fit the Rasch model at the 5% significance level (p=0.036 and p= 0.013 respectively). Conversely, the best-fitting items were *'hurt'* (χ^2=5.142), *'carefully'* (χ^2=4.367) and *'cried'* (χ^2=4.237).

'Hard to eat' (0.048), *'cross'* (0.130) and *'brushing'* (0.379) had fit residuals closest to 0 respectively, indicating a good fit to the model. The items *'annoy'* and *'carefully'* were found to have high negative item fit residuals (-1.802 and -1.801 respectively) and the item '*cried*' was found to have a high positive fit residual (1.112). Whilst these are notable, and could potentially indicate item redundancy (associated with Item-Total Correlation), a level of +/-2.5 should normally be reached for this to cause concern (Pallant and Tennant, 2007).

The items *'hard to eat'* (F=4.78; p=0.03) and *'cross'* (F=5.53; p=0.02) were found to have uniform DIF with regard to age at the 5% significance level. *'Hard to eat'* also showed DIF at this level (p=0.014), as did *'eat on one side'* (p=0.049). *'Food stuck'* appeared to be working differently for variations in age groups (F=-0.293) and genders (F=-0.126).

Regarding the individual items, none were subjected to local dependency with all values being less than 0.2 above the average correlation (Christensen et al., 2016). Furthermore, no items had disordered thresholds, although the thresholds did appear to be less even for some items, particularly *'food stuck'* and *'school'* as seen in Figure 5-2.

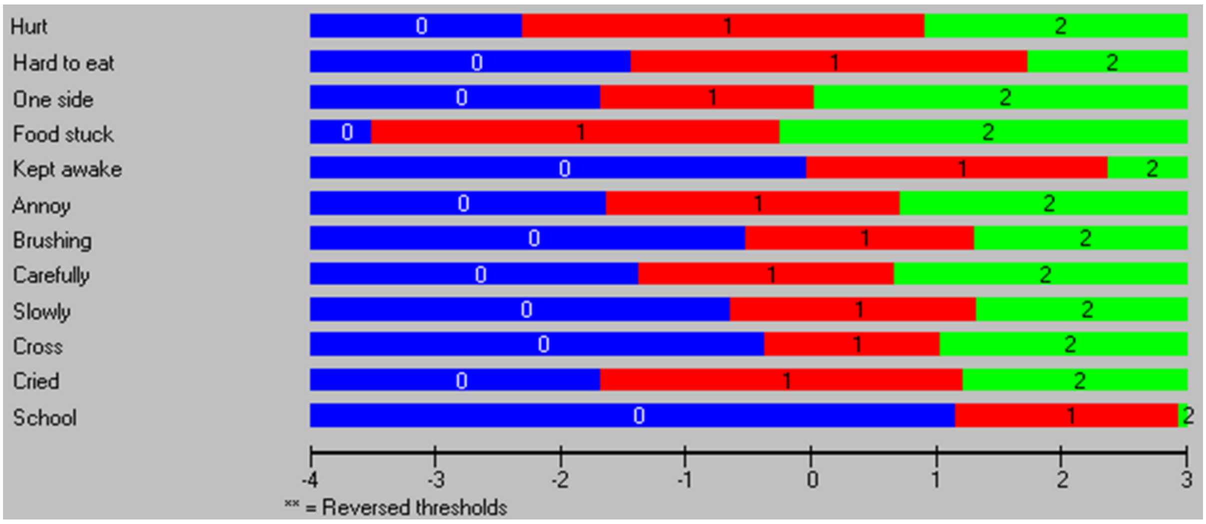


Figure 5-2: Threshold map from the items within CARIES-QC, showing that no thresholds are disordered

5.6.2 Classical Psychometric Testing

Principal component factor analysis identified only one factor to be present. This factor accounted for 45.54% of the total variance. The high Kaiser-Meyer-Olkin measure of sampling adequacy result of 0.914 determined that the sample was suitable for factor analysis. The statistically significant Bartlett's Test of Sphericity provided confirmation that the variables were correlated; a degree of correlation is necessary for factor analysis. The Scree plot is shown in Figure 5-3, whilst further findings from the factor analysis can be found in appendix M.

No items were found to have missing values greater than 5% suggesting there were no issues surrounding feasibility (Schafer, 1999).

There were moderate levels of correlation (between 0.3 and 0.5) between most items within CARIES-QC. Strong correlations (between 0.5 and 0.9) were found between the item '*annoy*' and five other items, namely '*hurt'* (*r*=0.59)*, 'one side'* (*r*=0.58)*, 'kept awake'* (*r*=0.52)*, 'carefully'* (*r*=0.55)*,* and *'cross'* (*r*=0.51). Similarly the item *'carefully'* had strong correlations with four other items, namely '*hard to eat* (*r*=0.51)*', 'one side'* (*r*=0.63)*, 'annoy'* (*r*=0.55)*,* and *'slowly'* (*r*=0.60). This suggests that a smaller number of items within the classification system could reflect what is captured by the wider measure.

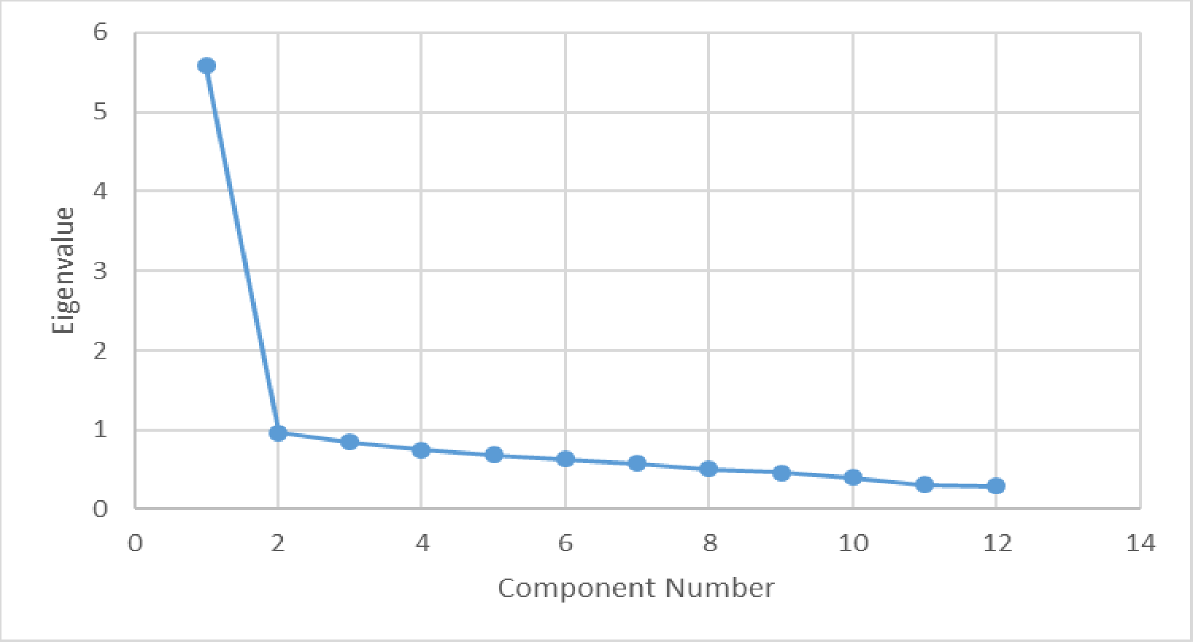


Figure 5-3: Scree plot from principal component factor analysis, demonstrating the presence of one factor

As the factor analysis did not identify multiple domains within CARIES-QC, correlations were undertaken between each item and the global question and total score at baseline (T0). All items had positive correlations with both the global question and the total score. *'Hurt'* had the strongest correlation (*r*=0.64) with the global question, whilst *'school'* (*r*=0.38) had the lowest. *'Annoy'* (*r*=0.79) and *'carefully'* (*r*=0.77) had the strongest correlations with the total score, whilst *'school'* again had the lowest (r=0.47). All correlations can be seen in appendix N.

Regarding the distribution of responses, *'food stuck*' was the only item to have a floor effect (32% responded 'a lot') without also having a ceiling effect. High ceiling effects were noted for '*kept awake*' and '*cross'*, with 67% and 59% of respondents reporting no experience of these impacts. A particularly high ceiling effect (82%) was observed in the item *'school'*, suggesting it was possibly misinterpreted by participants.

Data were available for 38 participants at follow-up (timepoint T2) after receipt of treatment. These data were used to calculate the SRM. The SRM for each item can be seen in Table 5-5. A strong SRM (>0.8) was found for *'annoy'* (0.93), followed by moderate effect sizes for *'food stuck'* (0.68) and *'hurt'* (0.61). Trivial effect sizes were observed for *'school'* (0.09) and *'slowly'* (0.16).

5.6.3 Patient and Public Involvement

A total of 17 children of differing ages and genders were involved in the PPI meetings, including members of the Young People Panel (outlined in section 5.5.7). Through the discussions it became apparent that some children and young people had experienced caries and treatment for it, whilst others had not. Two parent representatives (details provided in section 5.5.7) also had a discussion with the researcher (HJR), which was held separately from the children, to avoid any potential influences over the children.

Children and young people noted that there were multiple items within CARIES-QC relating to eating, and many suggested that one item alone could encompass the others on this topic. Children thought the items *'carefully'* and *'hard to eat'* had the broadest remit, and that one of these could be considered in place of the rest.

Children expressed some uncertainty about whether the item *'food stuck'* related to getting food stuck in their teeth in general, or getting food stuck in the holes in their teeth.

Children felt the term *'annoy'* was too similar to *'cross'*. Older children in particular thought they would be less likely to use the word *'cross'*, and hence would prefer the item *'annoy'*.

Older children thought that their peers would not be likely to admit to crying about their teeth.

Child and parent representatives expressed some confusion about how schoolwork could be affected by teeth. They reasoned that if dental pain was causing the impacts on schoolwork, this may be captured elsewhere under the category of *'hurt'*.

Parent representatives thought that pain related to toothbrushing, could also come under the umbrella term *'hurt'*. They also considered whether *'hurt'* and *'annoy'* might mean the same thing, though children and young people disagreed.

5.6.4 Developer opinion

A discussion with members of the research team who were involved in the development of the original CARIES-QC measure revealed a number of important points.

An issue was highlighted regarding the item *'food stuck'* during its translation into Dutch for the development of the CARIES-QC-NL. There was a need to clarify that the question related to getting food stuck in the cavities, or holes, rather than food getting stuck between the teeth in general, and this caused the translation to be somewhat lengthy and complex. Furthermore, the developers of CARIES-QC noted that children had a varied understanding of the item *'school'*, with some expressing confusion as to why their teeth would affect their schoolwork. The developers advised that these two items should be considered for exclusion from the classification on this basis.

Children and young people of different ages viewed the concepts of *'hurt'* and *'annoy'* to be different during development of CARIES-QC, although both terms were used to describe the physical sensations that they felt. This suggests it may be important to retain both of these items within the proposed classification system.

After a discussion regarding the issues raised by the PPI representatives, the developers of CARIES-QC highlighted that during their qualitative research, older children had indeed admitted to crying about their teeth, in contrast to the suggestions from child PPI representatives.

5.6.5 Discussion to determine the preliminary classification system

The findings from all four steps outlined above were discussed on 23rd May 2018 between all members of the study team, and the preliminary classification system (Table 5-6) was agreed by consensus. A summary of the key discussion points surrounding each item is provided below, based upon the results seen in Table 5-5 and appendices M and N.

Hurt

This was considered by both the developers and PPI representatives to be a key item and essential to retain within the preliminary classification system. It displayed the second highest item spread at logit 0, the best fit to the Rasch model and there was no DIF present. A moderate SRM also suggested it was sensitive to change, and a strong correlation with the total CARIES-QC score at baseline implied good internal consistency. *‘Hurt’* did have a strong correlation with the item *‘annoy’*, which is explored further below, though this alone would not be sufficient to justify the exclusion of an important item.

Hard to eat

PPI representatives felt that it was only necessary to retain one item surrounding eating, and felt that this item, or *‘carefully’* would be the best candidates. Whilst *‘hard to eat’* had a better item spread and residual than *‘carefully’*, the item fit was worse and DIF was present. Similarly, *‘hard to eat’* had a relatively low correlation with the global question score. Taking this into consideration, alongside the psychometric advantages of the alternative item, *‘hard to eat’* was excluded from the preliminary classification system.

One side

This was not considered by the PPI representatives to be the best item surrounding eating. This item also had poor item spread and a weaker fit to the model when compared to the other items covering this topic. It also had more missing data (2.5%) than any other item, and whilst this alone did not reach a level to cause feasibility concerns, when combined with the aforementioned issues it was considered appropriate to exclude this item from the preliminary classification system.

Food stuck

The item *‘food stuck’* had the highest item spread (1.632) in the Rasch analysis, suggesting it could be a good item to retain within the classification system. Similarly, *‘food stuck’* had the second highest SRM, indicating good responsiveness. Nonetheless, this item also had the highest floor effect (32%), and although this did not meet the threshold of 50% (set within the present study as a level of note), it was acknowledged that this would not allow a declining OHRQoL to be captured. Furthermore, this suggests the item may have been misinterpreted by respondents, or that a high proportion of children experienced this impact regardless of whether they had caries or not. When considered alongside the concerns raised by the developer and PPI representatives regarding the translation and interpretation of this item respectively, the panel agreed this item should be excluded from the preliminary classification system.

Kept awake

This item did not have a particularly strong item spread or fit to the Rasch model, though it had a good fit residual. It was apparent from the large ceiling effect that few participants experienced the highest severity level for this item. The panel considered that this was likely to be one of the less common impacts of caries, but one that could significantly affect the OHRQoL of those with more advanced disease. It was acknowledged that the preliminary classification system would require an item such as this in order to represent one of the worst health states, and hence *'kept awake'* was included.

Annoy

Whilst it was noted that the item *'annoy'* was not found to fit the Rasch model at the 5% level (p=0.013), it was considered important to retain, given its strong correlations with other items. It was anticipated that *'annoy'* could potentially capture some of the impacts of the items that it strongly correlated with. Furthermore, *'annoy'* had the highest SRM, indicating it was the most sensitive item to change, and had the strongest correlation with the total CARIES-QC score at baseline. Although parents expressed concerns that *'annoy'* could be too similar to *'hurt'*, during the development of CARIES-QC it was noted that children considered them to be differing concepts (Gilchrist et al., 2015). Whilst *'annoy'* did have strong correlations with *'hurt'*, the data suggested these items were independent of each other.

Brushing

The item *'brushing'* had a poorer spread at logit 0 than most other items within CARIES-QC and displayed a comparatively poor fit with the Rasch model also. It had a good residual, but performed differently for variations in ethnicity. The high ceiling effect suggested that this impact may not be commonly experienced by children. Whilst the developers considered *'brushing'* to be quite distinct from the other items, parent representatives felt that the discomfort experienced when toothbrushing could be encompassed within the broader item *'hurt'*. With these points in mind, the panel decided to exclude *'brushing'* from the preliminary classification system.

Carefully

As the PPI representatives expressed a need for only one item relating to eating within the classification system, the panel felt that *'carefully'* would encompass this best. This was in part due to its strong correlations with other items surrounding the impacts and experiences from caries upon eating. Furthermore, it had a good fit with the Rasch model and was not affected by DIF. As such, *'carefully'* was included in the preliminary classification system.

Slowly

Item spread and item-level fit were both low for *'slowly'* in comparison to other items in CARIES-QC, and this item experienced uniform DIF in relation to variations in deprivation levels. Furthermore, *'slowly'* had an extremely low SRM, indicative of poor responsiveness. As it was strongly correlated with another item related to eating that had better properties (*'carefully'*), the panel agreed to exclude *'slowly'* from the preliminary classification system.

Cross

The item *'cross'* had the second-worst item spread out of all of the items in CARIES-QC, and had DIF relating to age and ethnicity. Child PPI representatives expressed some uncertainty regarding their use of the word itself, and some felt that it was too similar to the term *'annoy'*. Given the strong correlations between these two items, the panel considered it to be advantageous to have *'annoy'* within the preliminary classification system instead of *'cross'*, given that it had better properties. As such, *'cross'* was excluded from the preliminary classification system.

Cried

Whilst adolescent PPI representatives thought their peers would be reluctant to admit to crying about their teeth, the developers of CARIES-QC found the opposite to be true during their study. The clinicians on the panel felt that *'cried'* could be a valuable item to have within a classification system, to represent one of the worst health states for caries. As there were no significant psychometric issues affecting this item, the panel agreed to include it within the preliminary classification system.

School

The item *'school'* had a relatively low item spread and was affected by uniform DIF relating to ethnicity. This item had a very high ceiling effect, and was subject to misinterpretation amongst the child PPI representatives. Importantly, *'school'* had almost no effect size, indicating poor responsiveness and the lowest correlations with the global question score and the total CARIES-QC score at baseline. The panel agreed that this item should not be included within the preliminary classification system.

Table 5-6 shows the five items that were selected to form the preliminary classification system. These broadly represent the domains shown in Table 5-7.

Table 5-5: Summary of results from the four approaches to identify the preliminary classification system from CARIES-QC

Item	Item Spread at Logit 0	Item level fit χ^2 (P-value)	DIF	Residual	SRM	Missing Data (%)	Floor effects (%)	Ceiling effects (%)	Strong correlations with other items	PPI views	Developer views
Hurt	1.605	5.142 (0.076)	✗	-0.757	0.61	✓ (1.5)	✓ (17% 'a lot')	✓ (31% 'not at all')	✓ (annoy)	✗	✗
Hard to eat	1.585	1.288 (0.525)	✓ (age* and gender**)	0.048	0.30	✓ (2)	✗	✓ (43% 'not at all')	✓ (carefully)	✗	✗
One side	0.858	0.868 (0.648)	✓ (age**)	-0.793	0.33	✓ (2.5)	✓ (25% 'a lot')	✓ (37% 'not at all')	✓ (annoy; carefully)	✓	✗
Food stuck	1.632	6.646 (0.036)	✓ (ethnicity* **)	0.661	0.68	✓ (2)	✓ (32% 'a lot')	✗	✗	✓	✓
Kept awake	1.202	0.612 (0.736)	✗	-0.393	0.39	✓ (1.5)	✗	✓ (67% 'not at all')	✓ (annoy)	✗	✗
Annoy	1.174	8.699 (0.013)	✗	-1.802	0.93	✓ (2)	✓ (18% 'a lot')	✓ (40% 'not at all')	✓ (hurt; one side; kept awake; carefully; cross)	✗	✗
Brushing	0.913	1.362 (0.506)	✓ (ethnicity**)	0.379	0.49	✓ (0.5)	✗	✓ (57% 'not at all')	✗	✓	✗
Carefully	1.019	4.367 (0.113)	✗	-1.801	0.38	✗	✓ (18% 'a lot')	✓ (43% 'not at all')	✓ (hard to eat; one side; annoy; slowly)	✗	✗

Item	Item Spread at Logit 0	Item level fit χ^2 (P-value)	DIF	Residual	SRM	Missing Data (%)	Floor effects (%)	Ceiling effects (%)	Strong correlations with other items	PPI views	Developer views
Slowly	0.988	1.775 (0.412)	✓(deprivation*)	-0.874	0.16	✓ 0.5	×	✓ (55% 'not at all')	✓ (carefully)	✓	×
Cross	0.705	2.368 (0.306)	✓ (age* and ethnicity**)	0.130	0.51	×	×	✓ (59% 'not at all')	✓ (annoy)	✓	×
Cried	1.466	4.237 (0.120)	×	1.112	0.44	×	×	✓ (38% 'not at all')	×	✓	×
School	0.894	1.339 (0.512)	✓ (ethnicity*)	-0.536	0.09	✓ (0.5)	×	✓ (82% 'not at all')	×	✓	✓

Notes: DIF: Differential Item Functioning SRM: Standardised Response Mean *Uniform DIF ** Non-uniform DIF

✓ indicates presence of potentially problematic feature × indicates absence of potentially problematic feature

Table 5-6: The preliminary classification system for CARIES-QC-U

Items	Response levels
Hurt	Not at all A bit A lot
Kept awake	Not at all A bit A lot
Annoy	Not at all A bit A lot
Carefully	Not at all A bit A lot
Cried	Not at all A bit A lot

Table 5-7: The domains represented by the items selected to form the preliminary classification system for CARIES-QC-U

Domain	Item(s) from CARIES-QC
Physical impacts	Hurt Annoy
Impacts on daily activities	Carefully
Impacts on sleep	Kept awake
Emotional impacts	Cried

5.7 Results: Validation of the classification system

5.7.1 Participants

Interviews were conducted with 20 participants, of which 6 were male, and 14 female, at which point data saturation was considered to have been met. Two potential participants declined to take part; one parent felt their child was too shy to take part, whilst the other reported a lack of time.

Further details of the sociodemographic and clinical characteristics of these participants are provided in Table 5-8. The majority of participants (n=14) were White British, whilst the rest (n=6) identified with a variety of different ethnicities. The age of participants ranged from 6 to 15 years with a mean of 10 years (median 10 years). Using home postcodes to gain Index of Multiple Deprivation scores, half of the participants (n=10) were found to reside in the 'most deprived' areas of England. Typically the mean dmft was 2.85 (median 2.5; range 0-12) and DMFT was 1.7 (median 0; range 0-11). The mean length of interview was 8 minutes, though this ranged from below 5 minutes to upwards of 16 minutes.

Table 5-8: The sociodemographic and clinical characteristics of participants in the qualitative validation of the preliminary classification system for CARIES-QC-U

Participant characteristics	***n*=20 (%)**
Age (years)	mean: 10.1; range: 6-15
Gender	
Male	6 (30.0)
Female	14 (70.0)
Ethnicity	
Asian background	2 (10.0)
Black background	1 (5.0)
Mixed background	2 (10.0)
White British background	14 (70.0)
Other background	1 (5.0)
Unknown background	0 (0.0)
Deprivation	
Most deprived	10 (50.0)
More deprived	0 (0.0)
Average	3 (15.0)
Less deprived	3 (15.0)
Least deprived	4 (20.0)
Total dmft	mean: 2.85 (SD: 3.05); range: 0-12
Total DMFT	mean: 1.7 (SD: 2.88); range: 0-11

Notes: *dmft indicates the number of decayed, missing and filled teeth in the primary dentition, whilst DMFT indicates the number of decayed, missing and filled teeth in the permanent dentition

5.7.2 Qualitative findings

The qualitative findings from this validation study are described below, with quotes to illustrate each aspect. Key themes were identified surrounding the complexity of the reduced questionnaire, the interpretation, importance and appropriateness of items within it, as well as those that had been removed.

Complexity

Children found the questions relating to the proposed classification system straightforward to complete and did not appear to experience much difficulty in choosing an answer for each question. Furthermore, they believed the questions covered a range of impacts:

"They're kind of easy…but they mean a lot" (Participant 8, 11 years old)

Children thought their schoolfriends would be able to answer the questions easily too:

"They'd probably….they'd probably find it easy actually" (Participant 9, 11 years old)

Children found the questions that had been removed from the classification system to be more difficult:

"This [side] was definitely harder" (Participant 9, 11 years old)

On questioning, younger children struggled to make decisions between items and found it difficult to communicate a clear preference:

"*Both…I like them both*" (Participant 5, 6 years old)

Overlapping items

During the development of the proposed classification system, parent representatives for the study had raised some concern that the items *'hurt'* and *'annoy'* were too similar and potentially overlapping. Nonetheless, these interviews suggest the contrary, as children felt *'hurt'* and *'annoy'* described different things, and considered them both to have value:

"I think they're very different because annoying and hurt are two different meanings" (Participant 12, 13 years old)

Interviewer: "Some people think this question [points to *'hurt'*] and this question [points to *'annoy'*] are almost the same, some people think they are different. What do you think?"

Participant: "I think they're different" (Participant 5, 7 years old)

Importance of items

Children had conflicting views on the item *'cried'* relating to the question 'have you ever cried because of your teeth'. Those who had experienced this impact placed greater importance on this item:

"'Cause sometimes if they really hurt, I do cry…..I actually think that is important" (Participant 2, 6 years old)

However, those who had never experienced this impact expressed confusion:

"I don't really know why people would cry about their teeth" (Participant 3, 14 years old)

Appropriateness of items

Children thought the question 'do you have to eat more carefully because of your teeth?' did not adequately describe the dietary restrictions resulting from caries. They displayed a clear preference for one of the questions that had been removed from the classification system, which asked whether their teeth made it *'hard to eat'* some foods:

"*If you eat more carefully you can still eat but if you find it hard to eat you can't really eat much*" (Participant 6, 9 years old)

"Because if you have to eat more carefully it's like how you eat whereas "Does your teeth make it hard to eat some foods?" would like eliminate foods out" (Participant 3, 14 years old)

Items that were excluded from the classification system

Children and young people raised concerns with a number of items that had been excluded from the classification system.

They identified ambiguity regarding the excluded item *'food stuck'* and thought that it was unclear that it related to caries:

"I'd also say try to rephrase them instead of just putting teeth, like put like, 'do you get food stuck in your cavity or the hole in your tooth'…like be a bit more specific" (Participant 3, 14 years old)

Older children also felt that they would not use the word *'cross'*:

"It's quite a young word. You'd say like you're angry, you wouldn't really go like, 'Oh, I'm really cross with someone'" (Participant 3, 14 years old)

Children expressed issues with other excluded items:

"I don't think that's very important to be honest" [talking about the item '*slowly*'] (Participant 9, 11 years old)

Interviewer: *"So what do you think about that question?"* [referring to the item *'school'*]

Participant: *"I think it's not really…well, I think it is important because it's important that you do your school work, but I don't think your teeth has anything to do with it"* (Participant 8, 11 years old)

5.7.3 Interpretation of qualitative findings

The qualitative findings described above suggested that the preliminary classification system was not valid in its present state. Children highlighted an issue with the item regarding eating more *'carefully'*, and felt that it failed to encompass the extent to which their diet could be affected by caries. They expressed a clear preference for the item *'hard to eat'*, and thought this item should be reinserted in the place of the problematic item.

The rest of the items within the preliminary classification system were easily understood and considered to be both important and appropriate. Furthermore, children believed these items to be independent of each other and not overlapping. Similarly, children identified issues with items that had been excluded from the classification system, confirming that this decision was appropriate.

The results from the qualitative validation study were discussed at a meeting held on 5th December 2018 attended by all members of the study team, who agreed that the item *'carefully'* should be replaced with *'hard to eat'* within the classification system. The findings were discussed with members of the Young People Panel who agreed that this proposed modification to the preliminary classification system was in line with the findings from the qualitative interviews. The items validated for inclusion in the classification system are shown in Table 5-9.

5.7.4 Confirmation of health state descriptors

As shown earlier in Table 5-2, CARIES-QC contains both the items and three response levels; ‘not at all’, ‘a bit’ and ‘a lot’. The wording for these response levels was determined using a child-centred approach during the development of CARIES-QC (Gilchrist et al., 2018). It was important to ensure that this wording was retained as much as possible within the health state descriptors for the classification system. Nonetheless, the priority was to ensure that the health state descriptors were easily comprehendible by children and young people. Furthermore, it was important to ensure that these were in the same tense. Incorporating descriptors with differing tenses could cause confusion amongst participants completing tasks comprised of these health state descriptors; it could suggest that the impacts were experienced over different, unstated time periods, or that some impacts were historical and no longer occurring. For this reason it was considered most appropriate to bring all descriptors into the present tense.

Potential health state descriptors were discussed with members of the Young People Panel. They felt it would be easier for their peers to imagine a descriptor from a personal perspective, in that the term ‘my teeth’ should be used, rather than ‘your teeth’. The panel members thought it would be most appropriate to use this term at the start of each descriptor, for consistency.

Children and young people found the items *‘hurt’* and *‘annoy’* straightforward to convert into health state descriptors for the three levels. They found it reasonable to separate ‘not at all’ around the verb. Similarly, the panel agreed upon present tense descriptors for *‘kept awake’* and *‘cried’*. The most challenging item to convert into a descriptor was *‘hard to eat’*, as the panel found it could not easily be combined with the level ‘a lot’. They felt that ‘really’ was the most appropriate replacement term to use.

Children and young people confirmed that these descriptors were easy to read and understand.

The final validated classification system is shown in Table 5-9 and the suggested domains these items could be categorised into are shown in Table 5-10.

Table 5-9: The final five-item classification system for CARIES-QC-U, validated by children and young people

Items	Health state descriptors
Hurt	My teeth do not hurt me at all My teeth hurt me a bit My teeth hurt me a lot
Annoy	My teeth do not annoy me at all My teeth annoy me a bit My teeth annoy me a lot
Kept awake	My teeth do not keep me awake at all My teeth keep me awake a bit My teeth keep me awake a lot
Hard to eat	My teeth do not make it hard to eat some foods My teeth make it a bit hard to eat some foods My teeth make it really hard to eat some foods
Cried	My teeth do not make me cry at all My teeth make me cry a bit My teeth make me cry a lot

Table 5-10: The domains represented by the items within the validated classification system for CARIES-QC-U

Domain	Item(s) from CARIES-QC
Physical impacts	Hurt Annoy
Impacts on sleep	Kept awake
Impacts on daily activities	Hard to eat
Emotional impacts	Cried

5.8 Discussion

This chapter outlines a novel methodology to identify a classification system for a paediatric condition-specific preference-based measure from a condition-specific patient-reported outcome measure. This discussion will initially reflect upon the key findings arising from this body of work, before considering matters surrounding the participants in this study and the ethical and governance issues involved. Finally, the strengths and limitations will be summarised, incorporating an appraisal of the study design used, before highlighting areas for future research efforts.

5.8.1 Reflections on key findings

Rasch analysis

The Rasch analysis played a key role in identifying the preliminary classification system. Item spread is considered to be the most important test for this purpose and hence it is not surprising that the four items with the lowest item spread (*'one side'*, *'brushing'*, *'cross'* and *'school'*) were all excluded from the classification system (Young et al., 2011). Further to this, the three items with the best fit to the Rasch model (*'hurt'*, *'carefully'* and *'cried'*) were all retained within the preliminary classification system.

Nonetheless, there were some outcomes of the Rasch analysis that were considered, but not prioritised in the decisions surrounding the preliminary classification system. One example of this relates to the item *'annoy'*, which was found to not fit the model at the 5% significance level but was still included in the classification system. For the purposes of a classification system, whereby the items should represent the domains measured by the wider measure, this is not a reason for exclusion from consideration since it is acceptable that items can represent different concepts. Conversely, the item *'hard to eat'* was initially excluded from the preliminary classification system for similar reasons, despite performing well in the Rasch analysis, though later was reintroduced as a result of the qualitative validation study. This latter example highlights a limitation of this study, namely the subjectivity surrounding the inclusion of items. This is discussed further in section 5.8.5.

It is important to note that the 5% significance level was employed during the present study, rather than a Bonferroni adjustment, to allow a more critical review of each item. For example, the DIF that was identified by the present study was not identified during the original development of CARIES-QC, where a Bonferroni adjustment was applied (Gilchrist et al., 2018). Similarly, the item *'annoy'* was found to not fit the model at the 5% level, but did fit the model when using a Bonferroni adjustment. This is important to clarify, since although the Rasch analysis in the present study identified a number of issues, under normal standards these items would be considered to function acceptably.

Further to this, an analysis to determine the level of bias exhibited by the items with DIF that were identified could have been undertaken, as this may have influenced whether these items were selected for inclusion in the classification system or not (Tennant and Pallant, 2007). However, given that the DIF would not have been identified had Bonferroni adjustments been applied, it is likely that the impact of this DIF would be minimal.

As anticipated, the thresholds for all items were ordered; a testament to the use of Rasch analysis in the development of CARIES-QC. Despite this, whilst the thresholds were ordered, the thresholds were seen to be less 'even' for the items *'food stuck'* and *'school'*; this relates to the strong ceiling effect and floor effect that these items had respectively. This suggests that the full range of levels were not being utilised for these items, providing further justification for their exclusion from the classification system.

Classical psychometric testing

An important finding from the psychometric testing came from the factor analysis. With the presence of just one factor or domain within CARIES-QC, it was not possible to determine correlations between the items and domain scores, as would typically be conducted in this process (Young et al., 2011). Instead, correlations were compared between items, and between each item and the global question from CARIES-QC and the total CARIES-QC score.

The presence of moderate correlations between all items in CARIES-QC was unsurprising, given the unidimensionality of the measure and the obvious relationship between the impacts and experiences it covers. The strong correlations that the items *'annoy'* and *'carefully'* had with numerous other items, suggested that the retention of these two items within the preliminary classification system could reflect a wider set of items within CARIES-QC.

Similarly, the presence of floor and ceiling effects was expected, given the three response levels in CARIES-QC. Concerns were raised by the panel regarding the item *'food stuck'*, the only item to have a substantial floor effect and no concurrent ceiling effect, as it would be difficult for this item to detect deterioration. Whilst this finding influenced the decision to exclude *'food stuck'* from the preliminary classification system, it is interesting to note that an item with a very high ceiling effect, *'kept awake'*, was retained within the classification system. The literature suggests that items with high ceiling effects could fail to detect improvement in a patient's condition (Young et al., 2011). However, an item that is less frequently experienced by patients may be able to represent the more severe impacts of a condition and hence could be considered a valuable addition to a classification system.

The feasibility of each item was determined through analysis of the missing data. Whilst *'one side'* had the most missing data, this was still well below the 5% threshold beyond which concerns would arise (Schafer, 1999). This is important, since a high level of missing data can reduce the usefulness of an item (Young et al., 2011).

The SRM was the most appropriate indicator of each item's responsiveness, as determined by the strength of the correlation between the baseline and follow-up scores. Had this correlation been less strong, it would have been more appropriate to use Glass's Δ method, whereby the difference in means would be divided by the standard deviation of the baseline scores only (Norman, 2014).

Interestingly, the item *'food stuck'* had the second highest SRM, indicating good sensitivity to change, despite functioning poorly in other areas of the psychometric testing and Rasch analysis. The qualitative findings discussed in section 0, and the issues highlighted by the developers relating to the Dutch translation of this item, suggested perhaps children did not interpret this question as intended, by failing to relate the food getting stuck to the holes in their teeth. Nonetheless, this high SRM suggests this may not be the case. Regardless, there were too many other issues affecting this item for it to be included in the classification system.

Stakeholder views

The views of PPI representatives and the developers of CARIES-QC played an important role in the identification of the preliminary classification system in this study.

A classification system derived from an existing measure of HRQoL would typically contain only one item to represent each domain (Goodwin and Green, 2016). Nonetheless, the classification system identified by the present study contains two items within the suggested domain of 'physical impacts'; *'hurt'* and *'annoy'*. This relates to the relative importance that children placed on the item *'annoy'*, and how it appears to capture a different impact from that of *'hurt'*. This could explain why feeling *'annoyed'* and experiencing *'pain'* are also both seen within the child-centred classification system developed for the CHU9D (Stevens, 2010). Whilst parent representatives for this study thought these two terms were very similar, it seems that children interpret them very differently. A similar finding was noted during the development of CARIES-QC, where it was also acknowledged that the word *'annoy'* is considered by children to be a niggling sensation, and more of a physical impact, rather than something psychological such as an irritation, as adults may interpret the term (Gilchrist et al., 2015). It is also important to note that the four domains suggested within this study were provided to give an idea of suggested categories or themes that the items may be considered within. These cannot be considered as formal domains given that the factor analysis revealed the presence of only one statistical domain, and that children were not involved in the identification of these categories.

A conflict in stakeholder views arose regarding the item *'cried'*, whereby older children were unsure whether their peers would admit to crying about their teeth. In contrast, the developers of CARIES-QC were able to demonstrate that this did not occur in their evaluation study, and older children did report impacts regarding crying about their teeth. It is possible that the PPI representatives had not understood that the questionnaire was intended to be filled in confidentially, and perhaps thought that it would be possible for other people outside of the research team to view their answers. As such, the item *'cried'* was retained within the preliminary classification system and subsequently validated.

Validation of the classification system

The attempt to validate the preliminary classification system highlighted a key issue with one item, namely *'carefully'*. Children and young people eloquently described how this word did not address all the impacts from eating as well as the psychometric tests had suggested. Despite the strong correlations between *'carefully'* and other items relating to eating, participants felt it failed to acknowledge how there are some foods that simply cannot be consumed when an individual has caries, and implied that anything could be eaten, providing it was done with care. This is an important distinction to make, and one that the panel involved in determining the preliminary classification system had not considered. This provides support for the use of qualitative methodologies to validate a classification system, as a typical quantitative validation process would not have been able to identify this important issue. This is discussed further in section 5.8.3.

Children and young people felt that *‘hard to eat’* was preferable to *‘carefully’* as it covered this wider impact on eating. The decision to replace *‘hard to eat’* for ‘carefully’ within the final classification system was well justified, given that the former had actually outperformed *‘carefully’* in a number of tests conducted in the Rasch analysis. Whilst it lacked the strong correlations with so many other items, its relevance and importance to children and young people was the priority.

Interestingly, throughout the qualitative approach to validation it became apparent that children who had not experienced dental pain severe enough to cause them to cry were unable to understand relevance of this impact. The range of responses surrounding this item from a sample who all have diagnosed dental caries confirms previous research highlighting the variation in impacts that children can experience, and how many suffer no symptoms at all (Tickle et al., 2002). Furthermore, the association between the number of carious teeth and the impacts experienced is often not as linear as one might expect (Tickle et al., 2002). Nonetheless, it is important for a utility measure to contain an item that represents an impact that is only experienced by those with the greatest severity of the condition, as this plays an important role in the formation of what is known as the ‘PITS’ – the worst health state (Ratcliffe et al., 2011, Ratcliffe et al., 2015).

5.8.2 Participants

Sample size

Rasch analysis is known to be sensitive to larger sample sizes, with very large samples causing an increase in the frequency of statistical significant findings, causing difficulties in item reduction (Tesio, 2003, McTaggart-Cowan et al., 2010). The present study used a dataset comprising a sample size of 200 participants on which to conduct the Rasch analysis. This is a relatively small sample compared to those that have been used in the development of other HRQoL instruments and PBMs, which have seen samples with around 400 to 700 participants being used successfully, though is still considered to be sufficient (Young et al., 2009, McTaggart-Cowan et al., 2010). Similarly, the sample size was adequate for the conduct of the psychometric tests, primarily the principal component factor analysis, for which the KMO test revealed sample adequacy.

Regarding the qualitative validation of the classification system, data saturation was considered to have been met after interviews had been held with 20 participants. The point at which this is reached is heavily dependent on how the term data saturation is defined (Saunders et al., 2018). In the present study, data saturation was considered to have been met when consecutive participants began to raise the same comments about the items and nothing new became apparent, in line with the concept of 'informational redundancy' described by Sandelowski (2008).

Response rates and diversity

The parents of only two potential participants declined their child's participation in the qualitative validation study. A lack of time was cited by the parent of one potential participant as a reason to not participate, whilst the other felt their child was too shy to take part in an interview. As the vast majority of potential participants agreed to take part, this suggests that the study was considered to pose little burden. The promise of a voucher to thank participants for their time may have acted as an incentive for children to take part, though it was not intended to be considered as such.

A potential limitation of the validation study is that it sampled disproportionately more female participants than males. This does not reflect the wider population, whereby there is a trend for boys to have a slightly higher prevalence of caries than girls (Pitts et al., 2015). Whilst a purposive sampling strategy was used to guide recruitment of participants, an element of convenience sampling had to be applied for pragmatic reasons. Whilst the researcher (HJR) had earmarked referral letters relating to patients that appeared to fit the sampling strategy, a large number of these potential participants were not brought to their appointments. Unfortunately this is a common scenario in this field, as approximately 10% of new referrals to this paediatric department fail to attend. For practical and ethical reasons, it was important for recruitment to continue, to ensure that study timescales were adhered to, hence more female participants were recruited than originally intended.

Postcodes are frequently collected from participants in caries research, as it provides an indication of deprivation. Many participants within the qualitative validation of this study were found to reside in deprived parts of the country, which corresponds with existing knowledge on the association between caries prevalence and deprivation (Schwendicke et al., 2015a, Slade and Sanders, 2017).

Caries experience

The severity of caries experience of participants in this study was higher than the national average identified by the most recent Child Dental Health survey, undertaken in 2013, that reported an average DMFT of 0.8 for 12-year-olds in England, Wales and Northern Ireland (Pitts et al., 2015). Similarly, a more recent oral health survey of five-year-olds in England reported a dmft of 0.8 (Public Health England, 2020). The caries experience in five-year-old children in Yorkshire and the Humber is known to be greater than the national average (28.7% compared to 23.4%), though the higher caries experience in the present sample is more likely to be explained by the recruitment of participants from a tertiary referral centre (Public Health England, 2020). These participants are likely to have been referred to the dental hospital due to the extent of their disease, and resulting symptoms. Whilst this could be considered a limitation of the study due to the lack of representativeness of the sample, it could be argued that those experiencing the impacts described in CARIES-QC would be the most appropriate sample to validate the classification system. Furthermore this approach ensured that those experiencing the most severe, and perhaps less frequently encountered impacts (e.g. crying) were involved.

5.8.3 Ethical and governance considerations

For the qualitative validation of the classification system, consent was provided by parents or carers, whilst children were asked to assent. Whilst assent is generally viewed as a positive aspect of research with children, by encouraging their engagement with the study and promoting their participation in the decision-making process, there is some debate over the concept of assent (Wilkinson, 2012). Firstly, there are numerous ways in which assent is defined, ranging from ‘acquiescence’; a mere lack of dissent, to affirmative agreement; the expression of positive assent, resulting in a lack of universal clarity (Baines, 2011). Further to this, there is a wide variation in the proposed limits of assent, with the upper limit being the ability to consent, and the lower limit (lacking any formal definition) ranging from as low as

7 years to as high as 14 years (Baines, 2011). Moreover, varying opinions exist as to how a conflict between parental consent and child assent should be dealt with.

For the younger age groups in this study, it was necessary for consent to be sought from parents as they would not be considered able to make the more significant decisions in their lives (Baines, 2011). Nonetheless, many older children in this study would have been capable of providing consent themselves. Given that there were a range of ages of children involved, it was considered acceptable to ask parents of all children to provide consent, for practical reasons. On reflection, whilst the researchers in the present study would have preferred to allow older children to consent for themselves, considering the difficulties encountered when attempting to use this approach in the following stage of this study (described in section 6.6.2), this would not have been likely to receive ethical approval.

5.8.4 Strengths

Study design

The systematic and varied approaches used to identify the classification system can be considered one of the strengths of this part of the study.

The method used to select items for the classification system was based upon those that have been used in the development of a large number of condition-specific preference-based measures including measures for dementia, diabetes and cancer (Rowen et al., 2011, Brazier et al., 2012b, Mulhern et al., 2012, Mulhern et al., 2013b, King et al., 2016, Mulhern et al., 2017). This study further built upon this process by also incorporating child and parent views.

Child-centred approach

The involvement of children and young people was a priority throughout this study, from the identification of the preliminary classification system, through to the interpretation of the qualitative validation interviews. This level of involvement is rarely employed in the development of classification systems for paediatric preference-based measures, such as the generic EQ-5D-Y and HUI2, or condition-specific measures such as those for atopic dermatitis and asthma (Torrance et al., 1996, Stevens et al., 2005, Chiou et al., 2005, Wille et al., 2010). The incorporation of the views of children, young people and parents addresses ongoing calls to prioritise patient and public involvement in research, and ensures that the face validity of the classification system is maximised, and the health state descriptors are appropriate.

Qualitative validation

Whilst qualitative approaches have been used in the identification of items to form classification systems preference-based measures, particularly for older and younger populations, they have not been used in the validation of classification systems (Stevens, 2009, Sutton and Coast, 2014, Canaway et al., 2017). As demonstrated earlier, this novel approach offers many benefits over a quantitative approach, through ensuring that the items within the classification system are considered important to the relevant population. Whilst this study could have employed a combined quantitative/qualitative approach to validation, incorporating analyses on post-treatment (T2) data, it would have been difficult to address any conflict between findings from the two approaches, particularly since there was no hierarchy to the approaches used to identify the preliminary classification system. This is discussed further in section 5.8.5.

5.8.5 Limitations

Applicability of classification system

The applicability of this classification system is likely to be limited to populations within the UK only. The dataset used was derived from a UK sample and hence the findings from the psychometric tests and Rasch Analysis can only be applied to this population. Nonetheless, as CARIES-QC is now being used in other countries around the world, more datasets are likely to become available over the coming years. These in turn could be used as the basis on which to conduct statistical tests to validate the classification system in other populations. Furthermore if the current classification system is not appropriate for other countries, an alternative classification system can be identified based upon these data. Similarly, the PPI representatives from the UK may have differing views to those from other countries, and hence involvement of children, young people and parents from local populations would also be necessary.

Reproducibility

The discussions conducted when interpreting and deciding upon which items to include in the classification system involved clinicians, a senior health economist, and the researchers who led the development of CARIES-QC. This could be considered a strength of this part of the study, as a range of viewpoints from an interdisciplinary panel were incorporated into the identification of the classification system. The discussion raised a number of important points, and a consensus was achieved. Nonetheless, the reproducibility of this approach is clearly limited, and a different group of researchers may well have selected different items for inclusion in the classification system.

This partly relates to the lack of hierarchy in which each approach and test was considered. Whilst a similar study by Young and coworkers employed a more ordered approach to appraising the statistical findings, prioritising analysis of item spread, this study did not incorporate the views of PPI representatives (Young et al., 2011). Incorporation of the views provided by these representatives is critical; it would not be ethically justifiable to overlook their opinion simply if it contrasted with the results from a statistical test. Conversely, it would be considered precarious to include an item that was favoured by PPI representatives but performed poorly in all statistical analyses. The aforementioned discussion and consensus agreement was conducted in order to try to balance the findings from each of the approaches on an item-by-item basis, rather than applying a blanket hierarchical system to all items that could potentially result in oversight. Nonetheless, the impacts of this method on the reproducibility of this study are acknowledged.

Exclusion criteria

It is important to note that children and young people with dental trauma were not specifically mentioned within the exclusion criteria for the qualitative validation interviews. This meant that children with dental trauma could have participated in the interviews, and may have discussed factors that related to symptoms from teeth that had been damaged due to trauma, rather than caries. Whilst it is possible that this occurred, it is unlikely.

Children and young people were identified as potential participants by the referral letters from their dentist, which stated caries as the reason that they required assessment at the hospital. Should a child have attended their assessment appointment with evidence of dental trauma, this would have likely occurred since the time of the original referral. Alternatively, it is possible that the referral letter omitted the details of a concurrent dental injury. Nonetheless, should a participant have been recruited with both caries and dental trauma, the interviewer ensured each child was aware that the focus was about what it was like to have holes in their teeth. Those who had been referred for trauma, but on examination, also happened to have caries were not approached. On reflection, whilst it is unlikely to have had any significant impact on the present study, it may have been beneficial to have excluded participants with concomitant traumatic dental injuries from this part of the study.

Qualitative analysis

Thematic analysis using a Framework approach was chosen for use in this study as it allows the data to be interpreted easily alongside quantitative data in mixed-method studies. Furthermore, it is considered to be a pragmatic technique when the qualitative research is being conducted to meet a relatively narrow study aim (Ritchie et al., 2014).

Whilst the thematic analysis of qualitative data from the validation interviews was conducted by two researchers (HJR and ZM) independently, with discussion and agreement on the themes identified, the interpretation of the qualitative findings was not confirmed with the original participants, as is considered best practice, for pragmatic reasons. Nonetheless, the researcher who conducted the interviews (HJR) made every effort to clarify anything that was unclear with the participant at the time of the interview, and took field notes to document any other observations that would support the interpretation of the qualitative data. The involvement of representatives from the Young People Panel to review the interpretation of the qualitative data acknowledged the need to gain input from the relevant population prior to the adaptation of the classification system. Nonetheless, the involvement of the original participants for this purpose would have been a stronger approach.

5.8.6 Further research priorities

Short-term priorities

It is important to address the aforementioned limitation regarding the applicability of the classification system in other populations. At the time of writing, data were available from an interventional study in New Zealand and epidemiological studies in the Netherlands and China (Schuller et al., 2018, Foster Page et al., 2019, He and Wang, 2020). Analysis of these data, ideally complemented by the views of PPI representatives from the local populations, would indicate whether the UK classification system for CARIES-QC described here would be applicable in other cultures, languages and populations, or whether adaptations would need to be made. This process should continue as CARIES-QC is translated into more languages and validated for specific populations.

Longer-term goals

The involvement of children and young people should become standard practice in the development of classification systems for this population, whether they are condition-specific or generic. This approach not only increases the validity of the classification system, it also addresses the need to involve children and young people in research and healthcare decisions. Furthermore, researchers should be encouraged to consider the use of qualitative validation approaches, given the advantages described above, rather than relying solely on the use of quantitative methods.

5.9 Conclusion

In conclusion, following child-centred modification as detailed above, the preliminary classification system can now be considered valid, since it has been derived using a combination of Rasch analyses, classical psychometric testing, PPI and developer views, clinical input as well as involvement of children with dental caries. The five-item classification system can now be considered suitable for use in a valuation survey with children and young people. Furthermore, the novel methodology used to develop and validate this classification system can be used in the development of other preference-based measures.

There are a number of different tasks that can be used in valuation surveys to elicit preferences. The next chapter will explore which type of task is most suitable to elicit preferences from children and young people, and will address further gaps in the evidence-base surrounding the administration of these tasks to this specific population.

5.10 Publications arising

5.10.1 Published abstract from poster presentation at 2018 International Association of Dental Research General Session (London, UK)

A classification system for a paediatric caries-specific preference-based measure. Rogers HJ, Rodd H, Marshman Z, Gilchrist F, Foster Page L, Rowen D. *Journal of Dental Research* (2018) 97(Spec Iss B):1869 (www.iadr.org).

5.10.2 Published abstract from oral presentation at 2019 International Association of Paediatric Dentistry Congress (Cancun, Mexico)

Child-centred validation of a classification system for a caries-specific utility measure. Rogers HJ, Marshman Z, Gilchrist F, Rodd H, Rowen D. *International Journal of Paediatric Dentistry* (2019) 29(Suppl. 1):p139.

5.10.3 Peer-reviewed journal article

Rogers HJ, Gilchrist F, Marshman Z; Rodd H, Rowen D. Selection and validation of a classification system for a child-centred caries-specific preference-based measure. *Journal of Patient-Reported Outcomes* 2020:4(105) DOI: 10.1186/s41687-020-00268-9

6 Child-centred evaluation of preference-elicitation tasks

6.1 Background

A validated classification system is just one component of a preference-based measure (PBM); it is also necessary to obtain preference weights for every health state described by the classification system. Before it is possible to generate this latter component, it is important to clarify who will be providing these preference weights.

Paediatric PBMs can be valued by either children and young people themselves, or by adults. As detailed in section 2.6.1, there are arguments to be made for and against each approach. Those that support the use of adult values are centred around their position as taxpayers. Through this contribution to society, it is considered that this population should have a role to the decision-making process by providing these values. Nonetheless, concerns have been raised regarding the viewpoint that adults should be asked to take when providing preference weights for paediatric measures, as discussed previously in section 2.6.1 (McCabe et al., 2005, Kind et al., 2015, Rowen et al., 2018b).

Conversely, arguments that support the use of children and young people's values tend to surround their ability to report on their own health, particularly since it is they who would experience the health states described by the classification system (Rowen et al., 2020). Furthermore, the ability of children and young people to complete health state valuation tasks has been demonstrated previously (Ratcliffe et al., 2011, Stevens, 2015, Ratcliffe et al., 2016a, Rowen et al., 2020).

For the purposes of the present study, and in acknowledgement of the need to involve children and young people in research, it was agreed that children and young people would be the most appropriate population to value the classification system for this caries-specific PBM. Moreover, it would be difficult to claim that the PBM created by this study was truly child-centred if it were to be based solely upon the preferences of adults. However, it remains unclear as to how well this population actually understand these tasks.

Stevens undertook an initial investigation into the reliability of valuation with children and young people by using cognitive debriefing with New Zealand schoolchildren aged 7-17 years as they completed tasks based upon the classification system from CHU9D (Stevens, 2015). The results suggested that children's ability to understand the tasks was dependent on both their age and the type of task being undertaken, with children under the age of 10 struggling to understand both DCE and BWS tasks. Stevens also noted that even amongst children of the same age, there would be some who were able to complete the tasks with ease, and others who struggled to understand them. Stevens acknowledged the need for further research in this area, using different classification systems. In view of this, it was important to clarify which age range of this population would be most suitable to provide the preference weights for the present study, using the classification system identified in the previous stage.

This in turn, is heavily dependent on the type of task to be used to gain the preference weights. As outlined in section 2.6.2, cardinal tasks, such as the TTO and SG, are not considered to be ethically appropriate for use with children and young people as they require participants to consider trading a year of their life and the risk of death respectively (Rowen et al., 2020). A key purpose of the ethical approval process is to prevent harm to research participants. Harm can take a variety of forms, though it is important to note that the risk of an individual experiencing harm, and the subsequent impact of that harm will vary across the population (Canadian Paediatric Society, 2008). Children are considered to be a unique group that require additional protection from harm due to the inherent vulnerabilities posed by their youth (Canadian Paediatric Society, 2008). Whilst there is no known evidence suggesting harm has been caused by undertaking cardinal tasks with adults, the potential exists for harm to be caused through the consideration of death at such a young age.

Further to this, the concepts of time and death are complex, hence these tasks require substantial cognitive processing that may exceed the abilities of younger populations (Petrou, 2003). Whilst one measure, the Assessment of Quality of Life instrument (AQoL6D), has gained preference weights using the TTO method with adolescents aged 16-17 years, the validity of these values could be questioned, in light of the concerns outlined above (Moodie et al., 2010, Ungar, 2011, Crump et al., 2017).

More recently, the use of ordinal techniques that have no mention of death, such as BWS and DCE, have shown promise as more appropriate methods to access children and young people’s preferences (Ratcliffe et al., 2011, Stevens, 2015, Ratcliffe et al., 2016a). The benefits of these tasks in terms of facilitating access to children and young people's preferences were initially outlined in section 2.6.2 and are discussed further below in relation to the present study.

6.2 Background to methodology

6.2.1 Type of choice task

For the purposes of preference elicitation with children and young people, the ideal task would not require an excessive degree of cognitive processing and would exclude topics deemed unsuitable for this population, such as death. Furthermore, it would have a strong consistency with underlying theory and would allow preferences to be anchored onto the QALY scale. It is important to consider how DCE and BWS tasks might compare to this ideal.

As introduced in section 2.6.2, pairwise DCE tasks require the respondent to state their preference between two hypothetical health states, each with described characteristics (Krucien et al., 2017). DCE tasks have been increasingly used in health state valuation with both adults and adolescents, and have been reported to have good consistency with underlying Random Utility Theory (Whitty et al., 2014, Krucien et al., 2017, Whitty and Oliveira Gonçalves, 2018). Nonetheless, the aforementioned pilot study found that adolescents under the age of 14 years struggled to understand the pairwise DCE, suggesting that there is still a substantial degree of cognitive processing required to complete them (Stevens, 2015).

It has been postulated that the cardinal method, TTO, could be considered as having a similar format to a DCE, given that respondents are also asked to make a series of pairwise choices between two discrete alternatives. A variant of the DCE, known as the $DCE_{TTO,}$ which includes a duration attribute for each profile in the pairwise DCE, would enable values to be anchored onto the 1-0 full health to death scale required to determine QALYs (Bansback et al., 2012). Whilst this would be of great benefit to a valuation survey with children and young people, unfortunately the DCE_{TTO} remains affected by the same issues as the standard TTO, with regard to greater requirements for cognitive processing as well as ethical concerns due to the mention of dead or death in the exercise (after the specified duration of the profile). As such, it is difficult to anchor the values obtained using DCE with this population onto the 1-0 scale, and hence the views of adults may need to be incorporated for this purpose (Brazier et al., 2012a).

In contrast to the DCE, the profile case BWS task provides the respondent with one health state profile from which they are requested to select the best attribute and the worst attribute, as outlined in section 2.6.2. The profile case BWS tasks have not been used as frequently as DCE in the literature, though Ratcliffe and coworkers (2016a) used this technique to elicit preference weights from 14-17 year olds for the CHU9D, as did Chen and colleagues (2019) for the estimation of an adolescent scoring system for the CHU9D in China. There was little qualitative evidence to inform these studies as to whether children and young people could understand the tasks and find them appropriate. Nonetheless, Ratcliffe and coworkers (2011) did undertake a quantitative feasibility study to determine how profile case BWS tasks derived from the CHU9D classification system compared to modified versions of SG and TTO in terms of ease of adolescents' understanding. The results suggested that BWS were more easily understood than the cardinal approaches, supporting the findings of the aforementioned pilot study, which found that children as young as 10 years could manage BWS tasks based upon the CHU9D (Stevens, 2015). Despite the indications from these studies, it would be necessary to determine whether these findings would arise with different populations and classification systems, with incorporation of a qualitative approach to further explore the understanding and views of participants.

Unfortunately, despite their alleged simplicity, there is no existing 'hybrid' combination of the BWS task with a cardinal task that could allow the values to be anchored onto the 1-0 full health to death scale. As such, the values of an adult population would be required in order to facilitate the anchoring of BWS preference weights from children and young people. Furthermore, a number of researchers have raised concerns about the use of BWS tasks in valuation exercises to determine utilities, citing both theoretical and technical reasons, as well as a lack of research surrounding their limitations (Krucien et al., 2017, Krucien et al., 2019). Nonetheless, given the unsuitability of cardinal tasks for eliciting preference weights from children and young people, a compromise may be required, and should take the form of the task best understood by this population.

As the above summary demonstrates, neither task meets the ideal criteria for a preference elicitation task for use with children and young people. Whilst acknowledging this, it is important for the tasks to be compared in terms of their suitability for this population, rather than just their characteristics alone. Whilst both DCE and BWS have been used in large-scale valuation surveys with adolescents with apparent success, it should be noted that there was little evidence available to inform the design of these surveys; for example, which age range of adolescents are able to understand these tasks, and how many tasks they can complete (Ratcliffe et al., 2011, Ratcliffe et al., 2012b, Ratcliffe et al., 2016a, Chen et al., 2019, Mott et al., 2019). Moreover, whilst the pilot study by Stevens confirmed that younger children were not able to understand ordinal tasks, it remains unclear as to which type of ordinal task is most appropriate for adolescents to complete (Stevens, 2015).

6.2.2 Number of choice tasks

There have been a number of concerns raised in the literature surrounding the burden of choice tasks on participants, and the threat of survey satisficing (Krosnick, 1991). When the survey respondent employs effort-saving heuristics, such as rushing through the tasks or choosing a response simply due to its position on the page, the quality of responses is likely to deteriorate (Berinsky et al., 2014, Bansak et al., 2018). Whilst one study, notably not in the field of healthcare, reported there to be no significant deterioration in response quality as the number of conjoint tasks increased, it is widely accepted that the number of tasks presented to participants should be as low as necessary to prevent undue burden and response fatigue (Krosnick, 1999, Johnson et al., 2013). This is even more important for an adolescent sample, who may already find the tasks to be challenging.

It is important to note that the burden of completing tasks will also be heavily dependent on the number of attributes and levels within the underlying classification system. A complex task with many attributes and levels will clearly require greater effort from a participant to read, process and make a decision. For the purposes of the present study it was considered important to determine how many tasks based upon the classification system for CARIES-QC-U a sample of adolescents felt they could manage.

6.2.3 Mode of task administration

As discussed in section 2.6.2, the literature has demonstrated it is possible for ordinal tasks to be self-completed by adults online, without the need for an interviewer (Mulhern et al., 2013a, Rowen et al., 2016). The use of online surveys to gain preference-weights has a number of advantages, as evidenced by their widespread use (Bansback et al., 2012, Singh et al., 2012, Ratcliffe et al., 2016b, Flood et al., 2018). Firstly, they do not require the presence of an interviewer, which makes them substantially cheaper to conduct. Secondly, multiple surveys can be completed in a short amount of time, with instantaneous return of data, as opposed to slower postal surveys. Furthermore, large sample sizes can be accommodated with ease.

The derivation of preferences from child populations raises a key conceptual issue relating to parental influence. This has been acknowledged by researchers previously when investigating the relationship between parent/carer and child health state valuation and measurement of HRQoL (Rosenbaum, 1996, Finkelstein, 1998, Ronen et al., 2003, Ungar et al., 2012). They identified that parents may become involved in the child's decision-making process, and in some cases the children relied upon this involvement. Whilst the extent of this influence is clearly difficult to quantify, researchers have made attempts to minimise this influence through directing questions to the child specifically, or asking parents to allow the child to respond (Ronen et al., 2003). Similarly, researchers have considered encouraging joint responses from children and parents in the form of a dyad, though this risks allowing the parent to take the lead and could potentially minimise the child's own views (Ungar et al., 2012). Nonetheless, it should be acknowledged that the majority of these studies have involved younger children, and the extent of parental influence over adolescents remains unclear. Similarly, these studies have assessed parental influence using an interviewer-led approach, rather than online self-completion surveys.

Whilst online self-completion BWS surveys have been conducted with adolescents, there is little available evidence to demonstrate whether this mode of administration is feasible and reliable in this age group (Stevens, 2015, Ratcliffe et al., 2016a). Before proceeding to an online valuation survey, it was considered important to determine whether adolescents were able to complete the survey independently, through observing the ease with which they navigated the different pages, and their ability to problem-solve when they encountered difficulties.

To explore these aspects, the present study employed qualitative methods, combining both cognitive 'think aloud' interviewing whilst participants completed the online survey, and subsequent semi-structured interviews. It was anticipated that this combination of techniques, alongside researcher observations, would help to illustrate how the survey results apply, whilst exploring participants' perspectives and beliefs regarding its content. The qualitative approach was considered as the priority in this stage, whilst the survey provided the basis for the cognitive interviewing, talking points for the semi-structured interviews, and a means of collecting sociodemographic and summary data. The use of qualitative methods to exemplify the survey findings in this way is known as elaboration (Morgan, 1998, Brannen, 2005).

As outlined in section 5.2.2 for the previous stage, there is a need to acknowledge the methodological stance that this study adopted, due to the implications it has on the qualitative research process (Tong et al., 2007, Ritchie et al., 2014). Whilst there are a number of different 'schools' of qualitative research, the researcher must decide which approach is most relevant to their own study. Once again, a pragmatic methodology was considered to be the most appropriate stance for this study, as it enabled the research question to be addressed without requiring conformity to specific traditions aligned with other epistemological viewpoints (Ritchie et al., 2014).

6.3 Aim

This part of the study aimed to identify whether DCE or BWS preference elicitation tasks are more suitable for use in an online valuation survey for adolescents. The specific objectives required to meet this aim are as follows:

1. To determine adolescents' level of understanding for each type of task and consider how this relates to their age

2. To ascertain which type of task (DCE or BWS) adolescents prefer and the reasons behind their preferences
3. To determine how many tasks adolescents feel they are able to comfortably undertake independently using an online format
4. To seek adolescents' suggestions on how the survey content could be adapted to make it more user-friendly

6.4 Method

To address this aim, a computer-based survey was designed for secondary school pupils to complete, containing both DCE and BWS tasks comprised of health states surrounding the impacts of dental caries. Children and young people were asked to 'think aloud' whilst they completed the survey, which was followed by a one-to-one semi-structured qualitative interview. The reporting of this qualitative stage was undertaken in accordance with the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist (Tong et al., 2007).

6.4.1 Survey design

A survey was designed with SurveyEngine (SurveyEngine GmbH, Berlin, Germany) (SurveyEngine (SurveyEngine GmbH, Berlin, Germany) GmbH, Berlin, Germany) for participants to complete independently using a tablet computer. The survey (appendix O) contained basic demographic questions surrounding age and gender. These were followed by the general health question used in the World Health Survey Individual Questionnaire, which asks 'in general, how would you rate your health today?' The standard response levels for this question ('very good', 'good', 'moderate', 'bad' and 'very bad') were adapted for a younger population, with 'moderate' being exchanged for 'OK'. This modification was approved by members of the Young People Panel. The global question from CARIES-QC followed this to determine self-reported overall dental health, which enquires as to how much of a problem the participants teeth are for them, with the response options of 'not at all', 'a bit' and 'a lot'.

Five questions from CARIES-QC (Table 6-1) relating to the items selected to form the classification system were included as a warm-up exercise, with the same three response levels (‘not at all’, ‘a bit’ and ‘a lot’). These were intended to familiarise children and young people with the wording used in the tasks, and to engage them in thinking about teeth. The survey asked participants to think about their teeth ‘today’ when answering these questions. Whilst a period of recall (such as ‘today’, ‘in the past week’, ‘this month’ etc.) was not specified in the original CARIES-QC measure, it was acknowledged that use of CARIES-QC as a utility measure would require a clear time point to be defined so that changes in utility before and after an intervention can be detected. Of the potential periods of recall, ‘today’ was considered to be the most straightforward to use, and the most applicable for a utility measure.

Basic information about tooth decay was provided, accompanied by a photograph of a decayed tooth. The latter had previously been selected by PPI representatives for an unrelated study, as it was considered to clearly show the condition in question, but without causing distress to younger children.

This was followed by five BWS tasks and five DCE tasks; the order of which was presented to the participant first was randomised to minimise ordering effects. Each type of task was preceded by a walkthrough demonstrating how to answer, and a practice question to complete. The practice question could be completed multiple times as required until the participant felt happy to commence the main tasks.

The tasks incorporated the health states from the classification system identified in the previous chapter (Table 6-1). As this was not intended to be a valuation survey, a specific design (e.g. full factorial) was not used to select the tasks for each participant to complete; instead each participant completed the same ten tasks, though the ordering of these was randomised. The tasks were chosen to include more straightforward choices, as well as some that required more challenging decision-making.

Table 6-1: Questions from CARIES-QC included in survey as warm-up exercise, with related health state descriptors from the classification system

Questions from CARIES-QC	Health state descriptors
How much do your teeth hurt you?	My teeth do not hurt me at all

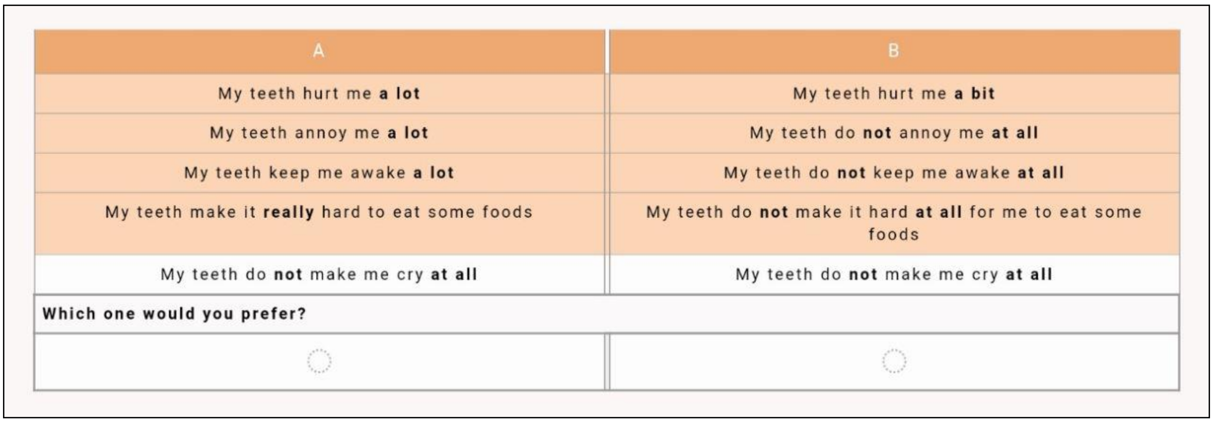
	My teeth hurt me a bit My teeth hurt me a lot
How much do your teeth annoy you?	My teeth do not annoy me at all My teeth annoy me a bit My teeth annoy me a lot
How much do you get kept awake by your teeth?	My teeth do not keep me awake at all My teeth keep me awake a bit My teeth keep me awake a lot
Do your teeth make it hard to eat some foods?	My teeth do not make it hard at all for me to eat some foods My teeth make it a bit hard to eat some foods My teeth make it really hard to eat some foods
How much have you cried because of your teeth?	My teeth do not make me cry at all My teeth make me cry a bit My teeth make me cry a lot

Error! Reference source not found. shows an example of a DCE task from the survey. As explained in section 6.2, the pairwise design provides the participant with two alternative hypothetical health state profiles from which they select their preferred option. **Error! Reference source not found.** provides an example of a BWS task from the survey. The profile case design provides the participant with just one health state profile, from which they first select the best feature, and then the worst feature, as previously discussed (section 6.2). Both of these examples demonstrate what could be considered as a dominance test. For the DCE shown in **Error! Reference source not found.**, it is clear that health state profile B would be preferable to profile A since profile B is better than or no worse for each item than profile A, and hence this is an example of a dominance test. It is then assumed that participants selecting profile A have either not understood the task, or have made an irrational decision (Tervonen et al., 2018). The use of a dominance test in BWS is less clear. In the example provided (**Error! Reference source not found.**), there are two potential ‘best’ options; ‘*my teeth do not make it hard at all for me to eat some foods’* and *‘my teeth do not make me cry at all’*. Furthermore, with just three response levels to choose from, it would be possible for a participant to select a middle option (‘a bit’) as being their ‘worst’ choice, without this necessarily indicating a lack of understanding or rationality.

Nonetheless, dominance tests for both the DCE and BWS were randomised within the survey accordingly.

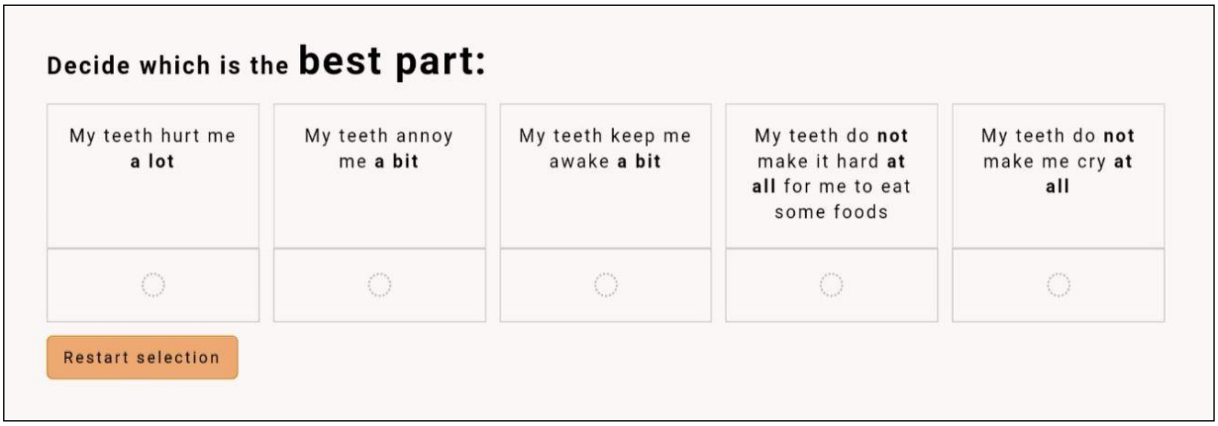
Figure 6-1: Example DCE task from the survey

Figure 6-2: Example BWS task from the survey



A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**
Which one would you prefer?	
○	○

The survey concluded with two summary questions regarding participants' difficulty of understanding and difficulty of choice, each with three response options ('easy', 'difficult' and 'somewhere in the middle') that were determined with involvement of PPI representatives (see section 6.4.8). The final question contained screenshots of the two types of task and asked participants to choose which one they preferred.



Decide which is the best part:

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○

The font and colour scheme for the survey (as seen in Figure 6-1 and Figure 6-2) were chosen in accordance with national guidance to aid participants with specific learning difficulties and visual impairments (UK Association for Accessible Formats, 2012).

The survey was designed to capture all responses automatically and recorded the length of time taken from commencement to completion of the survey. All data were downloaded into a Microsoft® Excel® 2016 (Microsoft Corporation, Redmond, USA) spreadsheet on completion of the interviews.

6.4.2 Setting

A secondary school in South Yorkshire, UK, was invited to participate in the study, primarily based upon the profile of the pupils with the school having above average proportion of pupils eligible for free school meals and ethnic diversity, according to the report from a previous inspection by the Office for Standards in Education, Children's Services and Skills (Ofsted, 2020). This was important to facilitate a sample of participants with varying sociodemographic profiles. The participating school was given £1000 to cover administrative and organisational costs.

6.4.3 Participants

Participants were identified according to the inclusion and exclusion criteria detailed below.

Inclusion criteria

- Children aged 11- to 16- years
- Children who are able to understand spoken English, i.e. able to understand and undertake the research with support

Exclusion criteria

- Children who are unable to understand and undertake the research even with support

6.4.4 Sample size and sampling methods

All students whom returned consent forms that had been completed by their parent were invited to participate providing they met the inclusion criteria, hence forming a convenience sample. Based upon previous research, approximately 30 to 35 participants were estimated to be required to reach data saturation, whereby no new codes or themes emerge, with at least five pupils from each year group to enable adequate representation (Stevens, 2015).

6.4.5 Participant recruitment

The following recruitment procedures were followed, in line with advice from the Research Ethics Committee (see section 6.4.9).

Posters explaining the study in simple terms were placed around the school to inform the school pupils about the study. One class (comprising between 20 and 30 children) was selected at random from each year group from Year 7 to Year 11, encompassing pupils aged 11 to 16 years. Children in these classes were give a copy of the parent information sheet and an ‘opt-in’ consent form (appendix A) to give to their parents, which was the usual method of written communication for this school.

Parents were given two weeks to return this form, after which time it was assumed that they would not wish their child to participate. As with previous studies using this ‘opt-in’ approach, a low return of signed parental consent forms was anticipated. If insufficient ‘opt-in’ forms were returned to meet the sampling requirements, a second, or third class in each year group were invited to participate in the same way.

All students for whom completed parental consent forms had been returned were given an age-appropriate young person’s participant information sheet (appendix Q). The study was explained to these potential participants in simple terms by HJR during their registration period (held first thing in the morning, and immediately after lunchtime) to ensure lessons were not disrupted.

The researcher (HJR) ensured the potential participant was aware of the purpose of the study and read through the information sheets with them, to ensure no pressure was placed on those with difficulty reading. Potential participants were informed that they would receive a £5 gift voucher on completion of the survey and interview, in acknowledgement of their time and commitment.

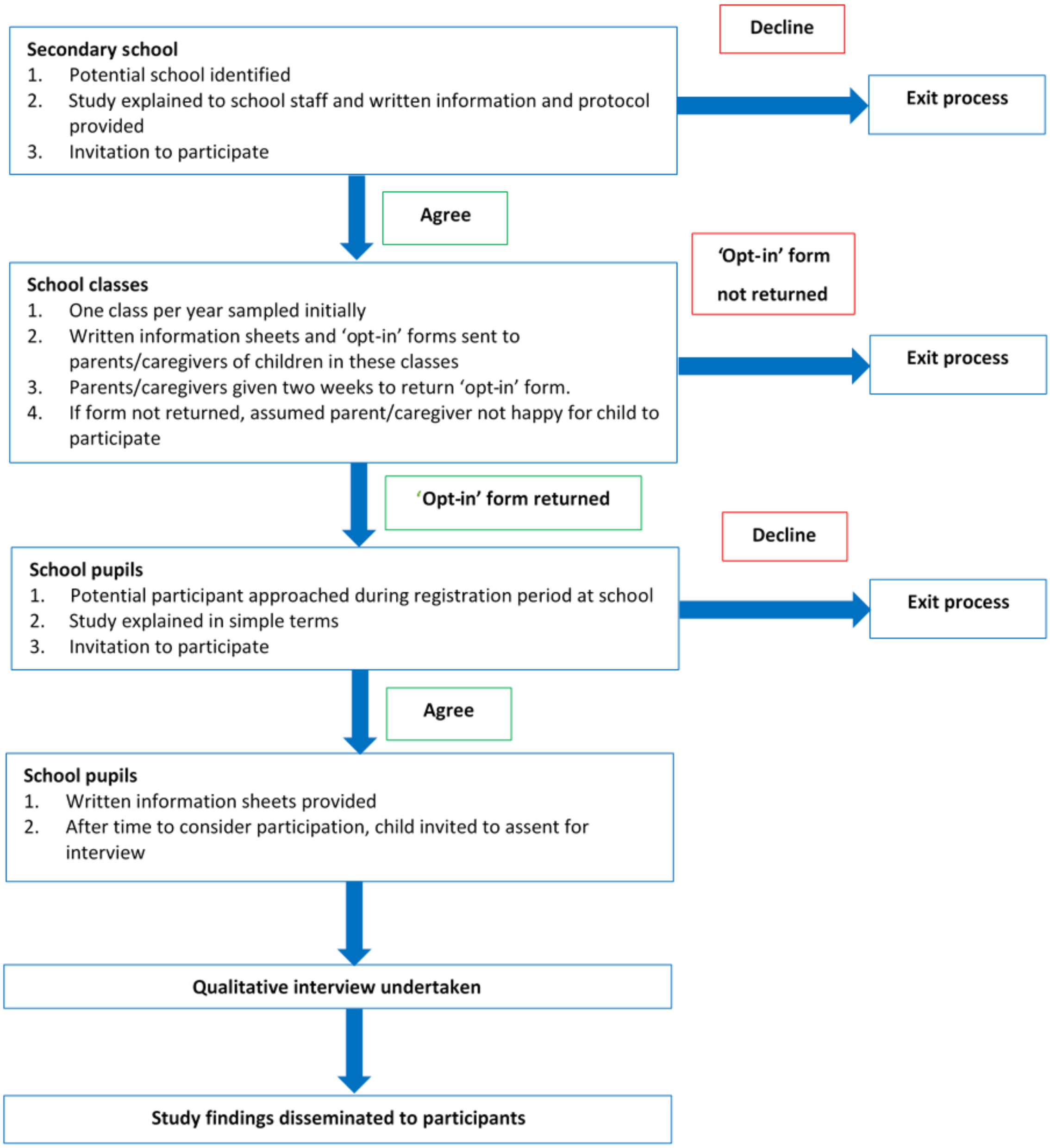


Figure 6-3: Flowchart displaying participant recruitment process for qualitative interviews

Where children agreed to participate, a suitable time was arranged to complete the survey and interview, utilising break times, registration periods and 'free' periods to ensure lessons were not affected. Older participants (from Year 10 and Year 11) were scheduled to have their interviews undertaken first, to avoid any interference with their revision for their upcoming school examinations. Children were asked to complete a young person assent form (appendix A) prior to commencing the survey and interview. A flowchart displaying the recruitment process can be seen in **Error! Reference source not found.**.

6.4.6 Data collection

Data collection for this part of the study comprised both quantitative and qualitative components; the survey captured quantitative data, whilst a combination of ‘think aloud’ and semi-structured interviews captured qualitative data.

Participants completed the survey in a quiet room within the school, away from communal areas and classrooms. The survey was presented on a touchscreen Fire™ HD tablet (Amazon, London, UK), with a larger 10” high definition screen to aid visual clarity. An ergonomic case was fitted to the tablet to allow it to stand upright, enabling the interviewer (HJR) to view the screen whilst participants completed the survey. The case also prevented damage to the tablet and allowed it to be wiped clean between participants.

Participants were asked to work their own way through the survey whilst ‘thinking aloud’. This approach (as previously discussed in section 5.2) allowed the researcher (HJR) to determine the participant’s thought process as they made each decision, facilitating an insight into their level of understanding (Willis, 2005). Participants were advised that they could ask for support with reading the survey, and could stop at any point without any repercussions, in accordance with the ethical considerations discussed in section 6.4.9.

A short semi-structured interview based upon a topic guide was then conducted, following completion of the survey. The topic guide (appendix R) was developed iteratively, with involvement of child and parent study representatives, and was designed to explore the reasons behind participants’ preference for one type of task over the other, their perspectives on how many tasks they felt they would be able to manage, and any suggestions they had on how the survey could be improved.

‘Think aloud’ and semi-structured interviews were conducted by HJR, with assistance from a junior colleague. As in the previous stage, participants were not made aware of HJR’s clinical background, to avoid this having an influence on their responses (Harden et al., 2000). Participants were advised they could stop the interview at any time, and that there would be no consequences if they chose to do so.

All ‘think aloud’ comments and interviews were recorded using an Olympus WS-853 digital voice recorder. Field notes were taken where required, particularly to document non-verbal observations. The duration of each interview, including the ‘think aloud’ component, was also recorded.

The experience and training of the researcher (HJR) in qualitative interviewing was outlined previously in section 5.5.5.

6.4.7 Analysis and interpretation

Simple descriptive statistics including were conducted on all sociodemographic data surrounding participant age, gender and self-reported general and dental health using Microsoft® Excel® 2016 (Microsoft Corporation, Redmond, USA). Data from the summary questions surrounding how easy or difficult participants found the tasks to understand and complete were also analysed in this way, along with the participant’s stated preferences between the two tasks.

All interview recordings taken during the ‘think aloud’ cognitive interviews and the semi-structured interviews were transcribed verbatim. Transcriptions were completed by an external company (©Dictate2us Ltd., Manchester, UK). All transcripts produced externally were checked by one HJR alongside the original digital voice recordings for accuracy. Data from each participant were analysed contemporaneously. Qualitative data were organised using a Framework approach in NVivo 12 software (©QSR International Pty Ltd., Chadstone, Australia) and analysed thematically by HJR and ZM independently, followed by a discussion to confirm the key themes (Ritchie et al., 2014). The prior experience of these researchers is summarised in section 5.5.6. Analysis focused on determining participants’ level of understanding for each task, the reasons why they preferred one type of task over another. Other aspects for analysis included the number of tasks they thought they could manage to complete, and the reasons why this might differ for each type of task, alongside their general thoughts on the survey and suggestions for improvement. The topic guide (appendix R) was modified as required to guide the interviews in eliciting further information surrounding emerging themes.

The experience and training of the researcher (HJR) in qualitative analysis was outlined previously in section 5.5.6.

6.4.8 Patient and public involvement

All participant-facing materials were developed with involvement of the Young People Panel as members of the steering group for the overall study. This ensured that participant information sheets and assent forms used age-appropriate terminology and explanations. Given that all participants in this stage of the research came within the adolescent group, it was felt that just one version of these documents would be suitable.

The Young People Panel were also involved in the development of the survey itself, to ensure that the structure made sense, the instructions on how to answer the tasks were clear enough, and that the colour scheme, text size and font were easy to read. Similarly, they were asked whether the questions within the topic guide were phrased appropriately, to avoid confusion.

Parent representatives were also involved, predominantly in the development of the parent/carer information sheets to ensure that it addressed any potential concerns that this group may have regarding their child’s participation, using appropriate language.

6.4.9 Ethical approval

As mentioned in section 5.5.8, ethical approval for the study was obtained from Yorkshire and the Humber Research Ethics Committee (Reference: 18/YH/0148).

Further to the original application for ethical approval, a substantial amendment (appendix S) was submitted to the committee with a draft of the survey, to seek their approval for its use in this stage. A favourable opinion was received, which can be seen in appendix S.

Due to the similarities of this stage with the body of work that was detailed in the previous chapter, many of the considerations mentioned in section 5.5.8 were applicable here. This included the project-specific safeguarding protocol (appendix A), though the researcher (HJR) was also aware of the schools own safeguarding procedures. Further to these, the following specific ethical issues were given attention as detailed below.

Researcher safety

The safety of researchers involved in this part of the study was also acknowledged to be paramount, particularly when undertaking field work alone. The University of Sheffield guidance on lone working was adhered to, ensuring that a member of the supervisory team was aware of the whereabouts and travel arrangements of the researcher (HJR) when conducting interviews.

Confidentiality

To ensure participant confidentiality, all interviews were conducted in a quiet room away from other classrooms and communal areas of the school, so they were not overheard by others.

Participants were anonymised through the assignment of a participant identifier. No participant identifiable information was recorded during the interviews. Data were only accessible by the direct research team. Data from the digital voice recorder were held on a secure network, accessed via a password-protected computer in a secure room in the dental school. Data in paper format, including field notes, were stored in a locked drawer, in a locked room within the dental school. The website which captured survey data was hosted on a UK-based web server in line with SurveyEngine (SurveyEngine GmbH, Berlin, Germany) protocols.

Pseudonyms or participant numbers were assigned whenever direct quotes from participants were used during dissemination of the study findings.

As in the previous stage, parents were not given access to the responses that their child gave during the qualitative interviews, to prevent breaching participant’s confidentiality, or damaging their trust in researchers.

6.5 Results

6.5.1 Participants

A total of 33 children (12 male, 21 female) took part in the survey and interviews were conducted between March and June 2019. Parent response rates were low, with just six to seven pupils from each class sampled (with each class comprising a minimum of 20 pupils) returning parental consent forms, with the exception of the Year 8 class, which returned no completed consent forms. A second class was sampled in Year 8, which allowed sufficient parental consent forms to be returned. All children who returned completed parental consent forms assented to participate in the study.

Table 6-2: Summary of participant characteristics

Participant characteristics	**(*n*=33)**
Age (years)	mean=14; range=11-16
Gender	
Male	12 (36.4%)
Female	21 (63.6%)
Self-reported general health: in general, how would you rate your health today?	
Very good	8 (24.2%)
Good	18 (54.5%)
OK	5 (15.2%)
Bad	2 (6.1%)
Very bad	0 (0%)
Self-reported dental health: how much of a problem are your teeth for you?	
Not at all	27 (81.8%)
A bit	4 (12.1%)
A lot	2 (6.1%)

Participants' ages ranged from 11 to 16 years, with a mean of 14 years (SD 1.55). The majority of participants reported themselves to be in 'good' general health (54%, n=18) and to have no dental problems (82%, n=27). Participant characteristics can be seen in Table Table 6-2.

6.5.2 Self-reported difficulty of understanding, difficulty of choice and preferred task

The time taken for participants to complete the survey and interview ranged from 10 to 20 minutes.

Participant’s responses to the summary questions from the survey are shown in

Table 6-3. The majority of participants found the tasks either easy to understand (n=18, 55%) or somewhere in the middle (n=14, 42%), with only one participant (3%) reporting the tasks as difficult to understand. There was no clear association between participants' age and their self-reported ability to understand the tasks.

Only 8 participants (24%) found it easy to choose an answer when completing the tasks, whilst 2 participants (6%) found it difficult to choose an answer. The rest of the participants (n=23, 70%) thought it was somewhere in the middle. Again, there was no obvious association between participants' age and their ability to select a response to the tasks.

The final question in the survey asked children to express a preference between the two types of task (DCE and BWS), which were displayed to them in the form of screenshots. The results were very close, with 14 participants (42%) preferring the DCE tasks, and 19 participants (58%) preferring the BWS tasks. As seen in

Table 6-3, preferences for each type of task are spread evenly for each age group, with the exception of Year 8 (comprising children aged 12-13 years), where all but one participant expressed a preference for BWS.

6.5.3 Dominance test

The dominance test which was randomised within the DCE tasks was failed by three participants (9%). However, more participants (n=5, 15%) ‘failed’ the BWS dominance test, taking into account the aforementioned issues with this approach. Of these, two participants correctly selected the ‘best’ option, but chose an incorrect ‘worst’ option (i.e. an attribute with a severity of *‘a bit’*, rather than *‘a lot’*). Conversely, one participant correctly identified the ‘worst’ option, but selected an incorrect ‘best’ option. The remaining two participants who failed the BWS dominance test did not correctly identify the correct ‘best’ or ‘worst’ options. Interestingly, there appeared to be no clear link between the failure of a dominance test and the age of participants.

Table 6-3: Self-reported difficulty of understanding, difficulty of choice and preferred type of task, by participant school year group

	Participant year (Y) groups (total n=33)					
	Y7 (age 11-12) n=7	Y8 (age 12-13) n=6	Y9 (age 13-14) n=7	Y10 (age 14-15) n=6	Y11 (age 15-16) n=7	**Total (%)**
Did you find the tasks:						
Easy to understand	2	4	5	3	4	18 (55)
Somewhere in the middle	5	2	1	3	3	14 (42)
Difficult to understand	0	0	1	0	0	1 (3)
Did you find it:						
Easy to choose an answer	2	2	1	2	1	8 (24)
Somewhere in the middle	5	3	5	4	6	23 (70)
Difficult to choose an answer	0	1	1	0	0	2 (6)
Stated preferred task:						
DCE	3	1	3	3	4	14 (42)
BWS	4	5	4	3	3	19 (58)

6.5.4 Qualitative findings

The qualitative findings arising from the 'think aloud' completion of the survey and subsequent semi-structured interviews are described below. The results covered three main areas; general findings relating to participants' understanding, more specific comments regarding the type of choice task, and their suggestions concerning the survey design.

General findings relating to understanding

General findings surrounded adolescents' ability to work their way through the online survey, to recall a single point in time, and their difficulties in understanding how to respond to the tasks.

Field note analysis revealed that participants were able to navigate the online survey with ease and did not request or require assistance in this regard. Participants had a tendency to address difficulties in understanding the tasks or choosing an answer by going back and reading the instructions again.

Participants were prompted to 'think aloud' when completing the warm-up tasks, which comprised five CARIES-QC questions surrounding the impacts and experiences from their teeth *today.* Nonetheless, participants did not consider their teeth during that day only, and instead recalled their dental experiences from different time points through their lives:

'I've never really thought about my teeth that much apart from like when they were wobbly' (Participant 26, 14 years old)

'I'm thinking in like general because like my teeth don't hurt anymore. They used to hurt a bit when I had braces on' (Participant 23, 15 years old)

These past experiences influenced some participant's responses to these questions:

'A few years ago I did have a hole in my tooth and I had to get it taken out and that really hurt a lot' (Participant 33, 13 years old)

Interviewer: *'When you were answering that question, were you thinking about how your teeth were today or how they've been...'*

Participant: *'Before…like the time they were pulling my tooth out. I was about like seven…I was crying because I was hurt and I'd had enough'* (Participant 11, 12 years old)

Further to this, participants expressed signs of misunderstanding when completing the tasks themselves. Whilst the instructions provided for the DCE tasks asked participants to express a preference between the two hypothetical states, instead participants had a tendency to select the option that most closely represented their own mouth and the dental impacts (or lack of) that they were experiencing. This suggested that they were unclear about how to respond to this style of task:

"(I choose) B because my teeth are fine..." (Participant 1, 16 years old)

“(I chose that) because I’ve actually never had a problem with my teeth” (Participant 7, 12 years old)

Similar findings were noted when some participants completed the BWS tasks, whereby they based their responses on their own dentition. Nonetheless, this was far less common:

“…So I’m thinking about like, my own teeth...” (Participant 20, 16 years old)

Participants also struggled to understand or complete the tasks in other ways:

“I think I’m going to go for B, because it has more bad things...” (Participant 15, 13 years old)

To summarise, participants were able to navigate the survey with ease. They struggled to limit their recall of dental impacts to one specific day only, which may have influenced their responses to questions. Participants had difficulty understanding the DCE tasks, selecting the option that most closely represented their own mouth, rather than the option they would have preferred.

Specific comments regarding the type of choice task

Children and young people highlighted a number of specific comments that related to the perceived differences in complexities of the two tasks and the ease with which they were felt to be understood. Further comments related to the extent that each type of task required the participant to make compromises, the number of alternative options to choose from, and the layout of the tasks.

Children had differing views regarding the DCE and BWS tasks. Children who stated a preference for the DCE tasks viewed their relative complexity in a positive light:

“*I think it gave me a bit more perspective on things…*” (Participant 2, 16 years old)

“*You can weigh up lots of things at once*” (Participant 15, 13 years old)

"*You have to think more about which one you'd rather have*" (Participant 29, 12 years old)

"*Well, personally I love reading so…everyone's different but to be honest I actually like the A and B ones more than this multiple choice*" (Participant 20, 16 years old)

Other children disagreed, and highlighted complexity as a key issue surrounding the DCE:

"…it was a lot of information to, like, read and process at the same time" (Participant 25, 14 years old)

"*…(it) wasn't really easy to understand...*" (Participant 33, 13 years old)

Conversely, children who expressed a preference for the BWS tasks valued their simplicity:

"*…(they were) a little bit easier to understand because you had like just less to think about*" (Participant 26, 14 years old)

These children also acknowledge the reduced need for compromise required to complete the BWS tasks:

"You don't have to have like all the other bits which you might not sort of wanted like" (Participant 17, 12 years old)

However, not all children viewed this relative simplicity favourably:

"*I think it was more vague than the A or B ones*" (Participant 20, 16 years old)

Furthermore, the increased number of options provided by the BWS tasks was considered a negative feature by some:

"*I find it easier to choose just between two rather than five*" (Participant 10, 11 years old)

Having seen examples of BWS in both a horizontal and vertical layout (though all BWS questions that participants answered used the horizontal format), children expressed a preference for the horizontal format, as they found it easier to read and valued its originality:

"Horizontal is better because then you can read across" (Participant 10, 11 years old)

"(I prefer) horizontal just because it was like, not like, every other survey that I've taken so…it was just different" (Participant 20, 16 years old)

To summarise, participants who preferred the DCE tasks valued the relative complexity that they offered. The BWS tasks were perceived to be the easiest to complete and required less compromise to be made when choosing a response, compared to the DCE, and a horizontal layout was considered as both novel and practical.

Suggestions relating to the survey design

Participants made a number of suggestions surrounding the number of tasks they could complete, the inclusion of practice questions, the amount of information that was provided about tooth decay and how this information could be delivered.

Participants proposed a large range in the number of tasks they could manage:

"Probably about five, so it's probably about the right amount before I start losing concentration" (Participant 19, 14 years old)

"Well I could…I could do loads…and like I can probably do about 30 or 40, but I know a lot of people wouldn't want to do more than 15 or 20" (Participant 21, 16 years old)

Generally between 8 and 10 tasks were suggested as being acceptable to themselves and other children they knew:

"I dunno [don't know]*, probably about 10"* (Participant 9, 14 years old)

"I don't know. I think I'd get through, like seven or eight of them and then… [start losing concentration]*"* (Participant 24, 14 years old)

Participants thought they could complete more BWS tasks than DCE tasks due to their ease of processing and there being less to read:

Participant: *"Probably more of those"* [BWS]

Interviewer: *"More? What makes you say that?"*

Participant: *"…because like I have a bit more choice and it's not as difficult because you have read them in the previous ones, you know, like do I want to cry more, do I want to cry less"* (Participant 29, 12 years old)

Participant: *"I could've probably done more of those ones I think"* [BWS]

Interviewer: *"Yeah? What makes you say that?"*

Participant: *"I suppose there was just less factors to like think about all at once"* (Participant 26, 14 years old)

"I think it wasn't the boredom that was the problem; it was just a lot of reading to do" [talking about the DCE tasks] (Participant 16, 14 years old)

Participants offered some suggestions on how the survey could be improved. Adolescents thought the walkthrough and practice question were useful, and that just one practice question would suffice.

Younger adolescents thought some more information about tooth decay initially would be useful, whilst older adolescents felt it was unnecessary:

"…you could have added a little bit more information of why it happens and what you can do to prevent it" (Participant 6, 11 years old)

"I think I kind of already knew that sort of stuff... because we've learnt about it before" (Participant 9, 14 years old)

Participants thought they forgot that they were thinking about tooth decay towards the end of survey and made suggestions on how to counteract this:

"If there'd been a sort of reminder in the middle of the quiz…" (Participant 14, 14 years old)

To summarise, adolescents perceived they would be able to manage between 8 and 10 tasks, though more of the BWS tasks than the DCE tasks due to the simplicity of the former. Adolescents had contrasting views regarding the amount of information that was provided to them about tooth decay, and admitted that towards the end of the survey they had started to forget that the topic was tooth decay.

6.6 Discussion

This section considers the main findings from the work and aspects of the study design in more detail. The first part will consider the key findings, and how they compare with or refute findings from other bodies of work. Consideration will then be given to the socio-demographic characteristics of the participants, before reflecting on ethical issues and recruitment. Strengths and acknowledged limitations of the research will be presented, before finally making recommendations for future research priorities.

6.6.1 Reflections on key findings

Preferences between DCE and BWS tasks

Adolescents in this study were asked to select their preferred type of task from screenshots of the DCE and BWS tasks. The results were similar, but slightly more (n=19) participants preferred the BWS than the DCE (n=14). Previous research suggested that BWS could be considered to be simpler tasks, as children as young as 10 years old demonstrated an ability to complete them (Stevens, 2015). It could therefore be presumed that children in younger year groups would prefer the BWS as they may have been able to comfortably complete them, as opposed to the DCE that they may have struggled with. Nonetheless, whilst the Year 8 participants expressed a clear preference for BWS tasks, there were still a number of Year 7 participants who preferred the DCE. This suggests that adolescents do not necessarily prefer the ‘easier’ option. It is important to acknowledge, however, that meaningful conclusions cannot be drawn from these quantitative data in light of the small sample size in this study.

This purported preference for BWS was elaborated by the qualitative findings obtained when participants were asked to explain their reasons for preferring one type of task over the other. Some, though importantly not all, preferred the complexity of the DCE tasks, in that it made them think more and gave them more to read, which they enjoyed. Nonetheless, more children appreciated the simplicity of the BWS tasks, in that there was less requirement for them to make compromises, and there was less information to both read and retain. Interestingly, there is some evidence to suggest that having more ‘difficult-to-answer’ questions within a survey may lead to an increase in participant drop out (Manfreda and Vehovar, 2002). As such, whilst some participants may like to complete a task perceived to be more complex (in this case, the DCE), there may be negative implications from using this approach within a valuation survey.

Participant preferences in terms of the type of task used, is often not considered to be a priority when developing a valuation survey, as other factors taken precedent. For example, a researcher may be more likely to use a preference-elicitation task that they are familiar with, or that has been used in many valuation surveys before. Nonetheless, it would seem logical that a participant is more likely to start, and importantly, finish, a survey that they engage with, can complete with ease and ultimately enjoy completing.

Suitability of tasks in relation to age

Whilst adolescents' preferences regarding the type of task are important, their understanding and ability to complete the tasks correctly is key. These factors were primarily determined by analyses of the descriptions provided by participants as they 'thought aloud' whilst completing the tasks. This was complemented by data collected from field notes and the qualitative interviews, alongside participants' ability to complete the dominance tests.

The qualitative findings from this study suggest that children and young people have a better understanding of and ability to complete BWS tasks than DCE tasks. This relates to the relative cognitive simplicity of BWS tasks, given that there is less compromise required in completing these tasks, and a reduced quantity of text to read, comparatively. This finding is supported by the aforementioned pilot study which found that younger children aged 10 to 13 years were able to complete BWS tasks but struggled with DCE tasks (Stevens, 2015).

The predominant issue affecting adolescents' ability to understand and complete the DCE tasks was also noted in the pilot study by Stevens. This relates to participants' tendency to select health states that most closely represented their own health. Stevens outlined three steps to the valuation process with children; firstly, imagining the hypothetical health state, then retaining the information and finally making a choice (Stevens, 2015). She identified that younger children (aged 7 to 9 years) struggled with this first stage, and instead made reference to their own health state. Whilst Stevens found this to affect young children's ability to complete both BWS and DCE, this finding was predominantly observed in relation to the DCE tasks in the present study. Similarly, this relationship with age was not identified in the present study, as adolescents of all ages demonstrated a predilection to choose the option that was most like their own teeth.

Interestingly, a recent study by Barber and coworkers (2019) also encountered difficulties in using DCE with adolescents. During their development of a DCE survey for hypodontia (the developmental absence of one or more teeth), adolescents found it difficult to understand the concept of trading off attributes. In keeping with findings from the present study, no relationship was observed between participant age and their ability to understand the tasks.

Indirectly, this body of work may have highlighted a potential issue surrounding the use of dominance tests for the DCE. This sample reported few problems with their teeth, and hence through completing these tasks by choosing the profile that was most similar to their own health state, young people were subconsciously still choosing the best option and hence passing the dominance test. This is an area that may benefit from further research to determine whether this also occurs in other adolescent groups and potentially adult populations. A number of issues have been highlighted in recent literature surrounding the use of dominance tests, particularly the lack of a consensus on how to account for those who fail the test during the analysis (Tervonen et al., 2018). Some studies use only the data from participants who pass the dominance test for analyses, though the present study would suggest that the assumption of a participants' rationality or their understanding of the tasks based upon this alone may be inappropriate (Tervonen et al., 2018).

On the surface, the results from the dominance tests suggested that more children understood and were able to complete the DCE tasks than the BWS tasks. Nonetheless, it is important to put this into context. The use and interpretation of dominance tests in DCE surveys is well established, yet this is not the case for BWS surveys, particularly where respondents are comparing across items and not across profiles, where there can be a truly dominant option.

As mentioned previously, it is not possible to design a dominance test for BWS tasks when used to elicit preferences for health states. This is because when an individual decides which attribute they consider to be the worst, this is informed not only by the severity of the item (*'not at all'*, *'a bit'*, *'a lot'*) but also how the participant perceives the impact of the attribute to be on their quality of life. This means that whilst it is possible to determine whether respondents have correctly determined the best attribute (i.e. with the severity level of *'not at all'*), it is not always possible to determine whether the choice of the worst attribute is irrational. Two of the participants in the present study 'failed' the dominance test by incorrectly identifying the worst attribute. Nonetheless, they did manage to successfully identify the best attribute. As such, it is difficult to determine whether these participants truly failed to understand the tasks or made irrational decisions, or whether they actually selected the worst attribute based upon their own perceptions of its impact on quality of life. If participants were considered to have 'passed' the dominance test providing the best attribute was correctly identified, there would have been an equal number of participants who had failed the DCE and BWS dominance tests. These issues surrounding the BWS dominance test limited the extent to which the present study was able to quantitatively compare how well the DCE and BWS tasks were understood.

Number of choice tasks

Adolescents in the present study felt that they could manage between eight and ten tasks before they lost concentration. Despite this, it is important to acknowledge that the approach used to determine how many tasks children and young people could complete in this study was particularly subjective. Children had little experience of completing these tasks upon which they could base their response, other than the survey they had just undertaken. It is important to acknowledge, however, that in this study these tasks were completed in a face-to-face setting, and hence it is possible that adolescents would be able to manage a different number of tasks in an online setting. Furthermore, in this study the tasks were completed alongside cognitive interviewing, and followed by a semi-structured interview, which together would have posed an additional burden to participants, potentially limiting the number of tasks that they felt that they could manage. Moreover, their perceived endpoints varied and were subjective too. Some children thought they would lose concentration after completing a certain number of tasks, others thought they would get bored, whilst some simply felt they wouldn't be able to complete any more tasks after a certain threshold was reached. This may account for the large range in answers provided. Nonetheless, children were aware that they had just completed a survey comprising 12 tasks; five DCE, five BWS and two practice questions (one for each type of task). This knowledge may have provided a basis for participant's responses to this question.

As discussed in section 0, findings such as this can help to determine how many choice tasks to present to participants in a valuation survey. Within the literature it is clear there is no consenus on this matter, as previous studies have presented a varying number of choice tasks to participants, ranging between 7 and 16 (Marshall et al., 2010, de Bekker-Grob et al., 2012, de Bekker-Grob et al., 2015, Vanniyasingam et al., 2016).

Typically, the adolescent population has been presented with slightly fewer tasks to complete than their adult counterparts, presumably in acknowledgement of their stage of cognitive development. The adolescent valuation of the CHU9D asked participants to complete ten BWS tasks, each with nine attributes for comparison, whilst the adolescent valuation of the EQ-5D-Y presented participants with 13 or 14 (depending on block allocation) tasks, each comprising five attributes (Ratcliffe et al., 2016a, Chen et al., 2019, Dalziel et al., 2020). In the dental field, a DCE survey on hypodontia (the developmental absence of one or more teeth) allocated eight tasks to participants, each with six attributes. Whilst these figures are not dissimilar to those perceived to be acceptable by participants in the present study, it would appear as though most studies select the number of tasks for pragmatic reasons related to sample size or block size, rather than what is most suitable for participants.

As would be expected, the type of task (DCE or BWS) influenced the number of tasks that participants felt they could manage. Adolescents thought they would be able to complete a larger number of BWS tasks than DCE tasks. This was due to the perceived ease of the BWS tasks and the reduced need for compromise. Furthermore, this was associated with participants expressing an overall preference for this type of task (BWS). To investigate this more objectively, however, a different study design would be required, using a mixed-methods approach.

Lastly, it is important to acknowledge that the number of tasks to complete is only one factor contributing to the burden on the participants; the complexity of the underlying classification system is also key. A system with many attributes and levels will undoubtedly result in more complex choice tasks and more content for the participant to read and process. The classification system for the present study, involving five attributes and three levels, is not considered to be particularly complex and so should minimise the burden in this respect.

These findings may provide some guidance regarding the number of choice tasks to present to adolescents in valuation surveys, in the absence of any clear evidence from the literature. Nonetheless, it is important to note that these findings relate specifically to the perspectives of a small sample of adolescents on the CARIES-QC-U classification system and hence cannot necessarily be applied more generally, or to more complex classification systems.

Participant recall

Indirectly, this study has raised the possibility of a much wider issue in asking children and young people to self-report their own health. PROMs of health-related quality of life are administered at multiple time points predominantly before, during and after delivery of an intervention. It is the difference in utility assigned to each health state experienced by an individual at these time points that can be used to determine the QALYs gained or lost. The present study found children were unable to select their responses considering their teeth in relation to 'today', and instead were referring to other impacts and experiences surrounding their dentition at other time points in their lives, particularly where they had suffered dental problems. This potential inability to focus on the present day when self-reporting their own health could affect the quality of data gained from children and young people to determine QALYs. Whilst it is possible that this was observed here because the sample was identified from the general population and participants were perhaps not knowingly experiencing problems with their dental health, it is possible that similar findings would be observed if the sample were reporting on their health in general.

Self-reported understanding and decision-making

The concluding part of the survey asked participants to report the ease with which they understood the tasks, and the ease with which they were able to select an answer. The majority of participants responded with 'somewhere in the middle', suggesting that there was a fair degree of cognitive processing involved, but that this did not approach their upper threshold. This high proportion of mid-point responses is a common finding in surveys due to a suspected social desirability bias (Nadler et al., 2015). This is thought to occur as respondents subconsciously attempt to avoid the bipolar endpoints, for fear that this may be viewed by others as an undesirable response. In the present study, with the researcher observing whilst the adolescent completed the survey, it is very possible that participants selected a neutral response to avoid judgement from the researcher. Similarly, it has been proposed that participants who are fatigued have a tendency to choose a mid-point response, when if pushed, they could select a more directional response (Krosnick, 1991). Nonetheless, it is important to acknowledge that it is entirely reasonable to select a mid-point response, and the respondent may legitimately hold a neutral opinion. Removal of the mid-point option can force these neutral respondents to make a selection that does not reflect their true opinion (Sturgis et al., 2014). The researchers considered that this would place undue pressure on the participants, particularly in light of their age. Furthermore, the PPI representatives for the study felt it was important to have a mid-point response option, and preferred the wording of this to be ‘somewhere in the middle’.

A potential limitation of this study relates to the survey design, in that these self-reported questions surrounding difficulty of understanding and difficulty of choice encompassed both types of tasks. As such, it is not possible to determine whether one type of task caused more difficulty for a respondent rather than another. Furthermore, it should be acknowledged that the sample size for this study was based upon the requirements of the qualitative study, and hence quantitative findings such as these should be interpreted with caution.

Mode of administration and survey design

Participants in the present study were able to work through the online survey easily, without the need for assistance. This fits with our expectations of this population, known as Generation Z, as digital natives (McKnight, 2018). Participants were methodical in their approaches to address any difficulties they encountered, even though these were not always successful. This demonstrates their ability to complete the online survey independently. Nonetheless, concerns exist surrounding the quality of data that can be obtained with an online survey. Without the presence of an interviewer, it is not possible to determine whether the participant has understood or engaged with the survey (Rowen et al., 2016).

The literature has reported varying outcomes when comparing the preferences elicited from online surveys to those elicited using other modes of administration, such as face-to-face computer-assisted personal interviews (CAPI) and mail surveys, with an equal number of studies reporting similar responses to those reporting differing responses (Damschroder et al., 2004, Mulhern et al., 2013a, Rowen et al., 2016, Watson et al., 2019). Whilst the present study did not intend to compare modes of administration, but rather establish whether adolescents are able to use the survey and comprehend the content independently (i.e. without asking for assistance), it is important to acknowledge that the very presence of a researcher is likely to have altered participants' behaviour significantly.

Two other concerns regarding the use of online valuation surveys surround a tendency for lower participation rates than mail surveys or CAPI, and a lack of representativeness within the samples in terms of educational attainment and other socioeconomic characteristics (Mulhern et al., 2013a, Rowen et al., 2016, Watson et al., 2019). Whilst these concerns can partly be addressed through setting recruitment quotas for the survey administrator to meet, the diversity of a sample for an online survey will always be restricted to those who have access to a computer.

Whilst participants were able to use the survey, they did make a number of practical suggestions regarding the design that would be important to incorporate into a valuation survey. These related to the task walkthrough, the use of practice questions, the information on tooth decay provided, and the format of the BWS questions.

Interestingly, participants acknowledged that towards the end of the survey they were starting to forget that it was about tooth decay. Adult valuation surveys for condition-specific PBMs are increasingly undertaken without the participant being aware of the condition in question. This is recommended to prevent participant's own knowledge or preconceptions about the condition from influencing their preference weights (Rowen et al., 2012). Nonetheless, for valuation with adolescents, it was considered important to provide some context about the condition. Furthermore, it would be difficult to conceal the condition entirely due to there being frequent mentions of teeth throughout the survey. Whilst there would be no negative implications if the tasks (some, or all) were to have been completed without the participant having full knowledge of the condition, adolescents were keen to ensure that this did not occur, through the survey containing a reminder of the condition halfway through.

6.6.2 Participants

Response rates and diversity

A low response rate was anticipated due to the use of the ‘opt-in’ parental consent approach. The challenges of school-based research are well acknowledged in the literature, and a reliance upon gaining consent from parents and carers is known to be a particular barrier (Pokorny et al., 2001, Esbensen et al., 2008, Schilpzand et al., 2015). Furthermore, children who do have parental permission to participate in school-based research are more likely to have a two-parent family, with parents having higher educational attainment, and less likely to be from minority ethnic backgrounds (Kearney et al., 1983, Pokorny et al., 2001, Esbensen et al., 2008). This alone can limit diversity; a feature that is particularly important in qualitative research.

School-based research studies have the potential to be misunderstood by parents, so strategies to enhance communication with parents have been proposed to help improve response rates. These include provision of prior notice of the research plans, engagement with parents at school events such as parents’ evenings, written reminders with personal touches (e.g. hand-signed), and regular updates of recruitment progress within the school (Esbensen et al., 2008, Schilpzand et al., 2015). Whilst adoption of these strategies may have helped improve response rates for the present study, a pragmatic approach was necessary, given the relatively small sample size required.

Only one year group (Year 8) required a second class to be sampled, as no parental consent forms were returned from the class that was initially sampled. On exploring the reasons for this, it became apparent that the tutor for this particular group was not keen for their pupils to take part in the study. As such, the pupils had not been encouraged to return their parental consent forms. In order to facilitate research in schools, it is crucial to have the support and understanding of the school staff. A range of techniques to increase buy-in from staff members have been proposed in the literature, such as involving staff members to aid recruitment and providing incentives for individual classes, or the school as a whole, when recruitment targets are met (Esbensen et al., 2008, Blom-Hoffman et al., 2009).

Nonetheless, the use of incentives in this way could encourage staff to pressure pupils and their parents to take part. On reflection, it may have been beneficial for the researcher (HJR) to have spoken to, or written to all staff members whose classes were being sampled prior to the commencement of recruitment, rather than relying upon other staff members and written correspondence to relay the key messages (Esbensen et al., 2008, Schilpzand et al., 2015). Improved communication here may have prevented the need for a second class to be sampled.

Fortunately, the number of returned consent forms was sufficient to meet the requirements of this study though this was ultimately reliant on pupils with parental permission also assenting to participate. All pupils who had parental consent agreed to take part in this study, indicating that adolescents were keen to be involved. This positive response from adolescents is not uncommon in school-based research, with nonparticipation amongst this population being notably rare (Esbensen et al., 2008). Possible reasons for this occurring in the present study may be that they considered the topic to be interesting, or were keen to express their personal views on the subject. Furthermore, the voucher given to participants on completion of interviews may have provided some financial motivation, though it was purely intended to thank participants for their time. This level of participation also suggests that this study posed few barriers to their participation, such as a burden on their time, or interference with their school day.

Participant postcodes and ethnicity data were not collected as part of this survey. Postcode data are typically collected to identify the level of deprivation of the areas in which participants live, through the use of the Index of Multiple Deprivation tool, as seen in the previous stage (Index of Multiple Deprivation, 2015). This indicator of a participant's background is often used in caries research due to the acknowledged association between deprivation and caries experience (Pitts et al., 2015). Furthermore, ethnic disparities in caries experience are known to exist (Rouxel and Chandola, 2018). Whilst collecting this data may have facilitated a more diverse sample if using a purposive sampling strategy, this would have had little bearing on the convenience sampling approach that was ultimately used, in line with recommendations from the research ethics committee. Furthermore, the school itself was chosen as it had a higher proportion of pupils from minority ethnic backgrounds, and an above average proportion of pupils eligible for free school meals; the latter being considered as a proxy indicator of deprivation, as in other UK-based studies (Pitts et al., 2015, GOV.UK, 2019, Ofsted, 2020). Whilst this approach ensured that the study sample was identified from a diverse population, it did not allow confirmation that the sample itself was diverse, in terms of ethnicity and deprivation.

Data were collected, however, regarding participant gender. This showed that female participants comprised two thirds of the sample in this part of the study. Typically, a disproportionate gender ratio such as this could be attributed to response bias, though as participation in this study was dependent upon gaining parental consent, there are likely to be more factors involved here. The school population for the 2018/2019 school year did have a slightly higher proportion of females than males (52.9% to 47.1% respectively) which may have been reflected in this study (GOV.UK, 2019).

Despite the issues surrounding diversity in terms of ethnicity, deprivation and gender, the sampling process did retain an element of purposiveness in that it ensured participants were sampled from each school year group. This was important to allow adolescent's level of understanding to be related to their age.

Self-reported general and dental health

Over half of the sample in this study reported having 'good' general health, with a quarter of participants reporting 'very good' general health. It is difficult to ascertain how this relates to the general UK adolescent population, as health-related data are typically collected for a much narrower remit, such as childhood obesity, or the number of children receiving vaccinations, rather than overall self-reported health. It is, however, possible to relate these findings to the population of the school from which participants were recruited from, alongside available data from mainstream schools in England. The school was known to have a smaller proportion of pupils requiring Special Educational Needs (SEN) support, or a SEN Education, Health and Care Plan (EHCP) compared to the national average (10.8% and 1.7% of pupils in mainstream schools in England respectively) (GOV.UK, 2019). Only 6.1% (n=2) of participants in the present study reported themselves to have 'bad' general health, which is slightly higher than the proportion of pupils with SEN support and EHCP within the wider school, though less than the national average. Nonetheless, it is important to note that SEN support and EHCPs only apply to pupils with health and social care needs that affect their education, and is likely to exclude many children and young people with other health-related difficulties.

Furthermore, this finding is heavily dependent on how participants interpreted the term ‘health’. A study of 8 to 11-year-olds in Hungary found that children had a good understanding of health and disease at this age, though many of their views were consistent with the biomedical model of health, in that they did not consider the psychological, social and environmental influences (Piko and Bak, 2006). Nonetheless, there is evidence to suggest that older children, such as those involved in the present study, hold a more holistic view of health (Žaloudíková, 2010). It is also acknowledged that children’s perceptions of health are strongly influenced by their own experiences and those of their family, as well as their culture and environment (Renslow and Maupin, 2018). As such, there is likely to be significant variation in how the term is interpreted amongst this age group in particular.

The majority of participants in this study stated that their teeth were not a problem for them, with just one fifth reporting that their teeth were ‘a bit’ or ‘a lot’ of a problem. This was an interesting finding, given our knowledge on the prevalence of dental caries. The 2013 Child Dental Health Survey found that almost a half (46%) of 15-year-olds and a third (34%) of 12-year-olds sampled in England, Wales and Northern Ireland, had ‘obvious decay experience’ (Pitts et al., 2015). As such, it is very possible that a proportion of participants in the present study whom reported no problems from their teeth, actually had dental caries. This would align with the aforementioned research by Tickle and coworkers, which demonstrates that many children with caries experience no symptoms (Tickle et al., 2002). Whilst the above discussion relates to caries, it is important to note that the question from which this data were derived (the global question from CARIES-QC) was vague, in that the problems reported could have been caused by a whole range of dental conditions, such as molar-incisor hypomineralisation or malocclusion. Furthermore, it not possible to draw any meaningful conclusions from these quantitative data given that the sample size was so small.

It was considered important to recruit a sample that was derived from the general population, rather than a group of patients with caries (the reasons for this are outlined in section 2.6.1). This helped to recreate the population that would complete the final valuation survey, and gave insight to the researchers on any difficulties this population may have experienced in completing the survey, due to their supposed lack of disease experience. Nonetheless, it is clear that due to the prevalence of dental caries within the population, it is very likely that any population-based sample will actually contain large numbers of 'patients' with the condition, whether they are aware of having the condition or not. Despite this, the aforementioned issues that can occur through use of a patient sample are unlikely to be produced by those within the population who are simply unaware of their disease status.

6.6.3 Ethical and governance considerations

The process for gaining ethical approval for this stage resulted in significant changes being made to the study design.

Consent

Originally, the protocol was submitted for ethical approval using a recruitment process whereby parents were asked to opt their child out of the project if required. The advantage of this approach was that it would allow the adolescents themselves to consent to participate, which acknowledges their ability to make decisions in this regard. Furthermore, it would overcome issues surrounding a low return-rate of completed parental consent forms (this is discussed further in section 6.6.2). Whilst this passive approach has been successfully adopted in other UK school-based studies when recruiting from this age group (Marshman et al., 2019), the committee did not allow it in this case. Instead, the ethics committee required parents to opt their child into the project.

Sampling approach

The original application for ethical approval sought to use a purposive sampling approach, to ensure a range of views were obtained from participants with differing backgrounds, in terms of gender and ethnicity. However, the committee expressed concern that potential participants whom had returned completed parental consent forms may be disappointed if they were not approached to take part. To address this, they advised that every pupil whom returned a completed parental consent form should be invited to participate.

In order to meet this requirement, it was no longer possible to maintain a purposive sampling strategy. To do so would have likely involved the recruitment of more adolescents than would have been necessary to reach data saturation. In light of this, a convenience sample was utilised instead. The implications of this approach are discussed further in section 6.6.2.

6.6.4 Strengths

Child-centred approach

This study has researched an important methodological area that is particularly relevant given the increasing interest in health state valuation with children. A notable strength of this study is the extent of involvement of children and young people, not only as active participants, but also as PPI representatives. The involvement of these representatives in the development of all participant-facing materials, including the survey and topic guide ensured the wording used was appropriate and understood by this age group. Furthermore, their views on design and formatting were crucial in the production of a survey that engaged participants, and was straightforward for them to work through.

Novel research

This is the first study to use a qualitative approach to compare the use of DCE and BWS to elicit health state preferences relating to dental caries from UK adolescents, using a computer-based survey. The only other study to have qualitatively investigated the use of these tasks with this age group was conducted with children and young people in New Zealand, using paper-based tasks derived from the CHU9D classification system (Stevens, 2015).

The use of ‘think aloud’ cognitive interviewing facilitated an important insight surrounding the extent to which participants understood these tasks, whilst the semi-structured interviews allowed further exploration of their views on the tasks, how many they thought they could complete, and which type they preferred. A further strength to this study was provided by the researcher (HJR) having formal training and experience in qualitative techniques, alongside expertise in communicating with children and young people.

6.6.5 Limitations

Reliability of qualitative findings

Two researchers (HJR and ZM) undertook the analysis and interpretation of the qualitative findings in this study independently, with discussion to agree upon the themes. Nonetheless, it remains best practice to ask a researcher independent from the research team to identify the predominant themes from a small sample of transcripts, to determine how they would compare to those selected by the study team. This may have helped to improve the validity and applicability of the findings from this study (Burnard, 1991, Appleton, 1995).

The credibility of the qualitative findings from this study could also be questioned as the interpretation of the data was not relayed back to, and discussed with, the original study participants. This did not allow confirmation that the data had been interpreted as intended (Guba and Lincoln, 1981).

Application in other areas of healthcare

A further limitation relates to the condition-specific nature of the classification system. Whilst directly relevant for this body of work, which focusses on the development of a preference-based measure for children with dental caries, it is not known whether this methodology would produce similar findings if repeated in the context of general health, or other specific conditions. Further research is required with different populations, using different classification systems.

6.6.6 Further research priorities

Short-term priorities

As discussed in section 6.6.5, the findings of this study relate specifically to the population described, and to the classification system used. It would be of particular interest to determine how the results may differ with use of a more widely-used paediatric PBM, such as the EQ-5D-Y (Ravens-Sieberer et al., 2010).

Alongside this, as the use of BWS increases, there is a need for further clarity on the use and interpretation of a dominance test for these tasks.

Longer-term goals

This study, somewhat unintentionally, has identified a potential issue regarding adolescents' ability to recall a specific time point, and the impacts that this could have on the values that they provide. As this was not the main focus of the present study, further qualitative research should investigate adolescents' completion of PBMs, to determine whether their past experiences of a condition impact on their selection of a response. If this were to occur, there could be significant ramifications for the use of adolescent-completed paediatric PBMs and PROMs. Likewise, the potential exists for this to occur in the adult population, and hence further research is recommended.

In a similar vein, this research has highlighted a possible problem surrounding the use of a dominance test for DCE tasks with adolescent populations. The present study identified their ability to 'pass' a dominance test through selecting the option that most closely represented their own health state. This contradicts the purpose of the dominance test in determining understanding and rationality. In light of this, it is important for future research to determine whether the dominance test for DCE in this population, and even adult populations, is serving its intended purpose.

6.7 Conclusion

This study found that adolescents aged 11 to 16 years old are able to complete an online valuation survey independently. The qualitative results from this study indicated that adolescents had a better understanding of BWS tasks compared to DCE tasks, suggesting that the former would be a more appropriate task to use with this population in a valuation survey. This can be justified further as adolescents also expressed a preference for completing BWS tasks. Participants in this study perceived that they would be able to complete between eight and ten BWS tasks.

Whilst acknowledging the aforementioned concerns raised by researchers regarding the use of BWS to determine utilities, the necessity of gaining high quality data from participants, through the use of a valuation task that they can better understand, and prefer to complete, should be seen to outweigh these issues. As such, the results from this stage were used to directly inform the final part of this study, which sought to gain preference-weights from adolescents of the same age for the CARIES-QC classification system previously described.

7 Valuation of a child-centred caries-specific preference-based measure

7.1 Background

The body of work outlined in the previous chapter suggested that best-worst scaling tasks (BWS) would be more suitable for preference elicitation with adolescents than discrete choice experiments (DCE). Whilst it was considered important to acknowledge these findings through the use of BWS tasks in a valuation survey with this population, it was also necessary to review the potential difficulties surrounding their use in this way.

Nonetheless, whilst the issues surrounding BWS should not be ignored, a degree of pragmatism is required. The use of a task that adolescents are able to understand allows the researcher to access values that more accurately reflect participants' views. The need to obtain reliable values must be considered paramount, in light of the impact that inaccurate data could have on the estimation of utilities.

Aside from this, there is one significant practical issue affecting the use of BWS, and that relates to the difficulty in applying the values to the 1-0 full health to death scale required to enable the determination of utilities, in this case QALYs. This is explored further below.

7.1.1 Background to methodology

Whilst the terms are sometimes used interchangeably, it is important to note the distinctions between the concepts of values and utilities. Technically, it has been argued that in order to generate utilities, the task used must allow participants to state a preference for risk, in order to meet the criteria underlying expected utility theory (Feeny and Torrance, 1989). This theory, also known as the von Neumann-Morgenstern approach, relates to decision-making in situations where the outcome is not known with certainty (Feeny and Torrance, 1989). The standard gamble (SG) is considered to be the only task that satisfies this theory, as an individual's preferences can be revealed through the choices they make in a series of gambles between differing health outcomes (Feeny and Torrance, 1989, Mehrez and Gafni, 1993).

Even the time trade-off (TTO) does not satisfy this criteria, as a comparison between life in a particular health state for a specified number of years, and being in a state of perfect health for a specified number of years does not incorporate uncertainty (Feeny and Torrance, 1989, Mehrez and Gafni, 1993). As healthcare interventions do not occur in a world of certainty, these tasks can only provide values. Nonetheless, a simple correction for the effects of uncertainty can be applied to values elicited using TTO in order to determine utilities (Mehrez and Gafni, 1993, Keeney and Raiffa, 1993).

Cardinal techniques such as these have been widely used in health economics, though in recent years the TTO has gained more prominence, in part due to its use in the valuation protocol of the EQ-5D. However, as mentioned previously, the mention of 'death' in both SG and TTO precludes their use for adolescent populations as it could cause distress. As such, the use of ordinal tasks is particularly beneficial in this population.

Ordinal tasks such as BWS and DCE are based upon a different theory of decision-making, known as Random Utility Theory (RUT) (Thurstone, 1927). This proposes that an individual has a latent 'utility' for each choice alternative, which is comprised of two parts; an explainable component, and an unexplainable component, with the latter incorporating all the unidentified factors that influence choices (Louviere et al., 2010). As such, researchers can predict the probability of an individual selecting a particular alternative, but not the exact alternative that the individual will choose. Crucially, the raw health state values from BWS and DCE tasks sit on an ordinal scale rather than on the 1-0 full health to death QALY scale (Coast et al., 2008b).

Ordinal values can be initially anchored to the least valued attribute level, and a linear transformation can be undertaken to place the best health state (in the present study, the level 'not at all' for each of the five attributes, which can be expressed as 00000) at one on the QALY (1-0) scale, and the PITS health state (the worst health state defined by the classification system, in this case, 22222) at zero (Coast et al., 2008b, Ratcliffe et al., 2015). This is the approach that Coast and colleagues used with values obtained using BWS for the ICECAP index of capability for older people, whereby the PITS, a state of having no capability, was given a value of zero (Coast et al., 2008a). Nonetheless, in order to generate QALYs, zero must represent death, not the PITS.

Rowen and colleagues (2015) explored the different ways in which ordinal values could be translated onto the QALY scale so that zero represents death. One way to achieve this is through anchoring, either using the coefficient for 'death', or the PITS health state, that have been obtained using cardinal tasks.

The latter of these approaches was used by Ratcliffe and colleagues (2016a) to determine an adolescent value set for CHU9D in Australia, using TTO data from young adults for anchoring purposes. Obtaining the TTO values from an adult population precludes the need to involve adolescents in completing a task that refers to death; a practice that is not recommended for a number of reasons that have been previously discussed. A similar method was also adopted during the adolescent valuation of the same instrument in a Chinese population (Chen et al., 2019).

The aforementioned study by Rowen and colleagues recommended two alternative approaches to translate ordinal scores onto the 1-0 scale. The hybrid model combines ordinal and cardinal data, and uses individual-level data, meaning it is particularly useful when there are fewer data points. Whilst this approach has been demonstrated to have good performance, the use of noisier individual-level data could limit the accuracy of this approach, and concerns have been raised surrounding the use of a hybrid model in the UK valuation of the EQ-5D-5L (Hernandez Alava et al., 2020). Similar results have been found using mapping; an approach that uses mean-level data (Rowen et al., 2015).

Mapping estimates cardinal utilities for all health states, or a selection of them, from the ordinal scores. This approach is considered to produce greater accuracy than anchoring using the PITS health state alone (Rowen et al., 2015). The adolescent valuation of the CHU9D in both Australia and China used this approach in addition to anchoring for comparative purposes, finding greater accuracy using mapping (Ratcliffe et al., 2011, Chen et al., 2019). Given that mapping has been shown to perform well in similar studies and is considered more straightforward than the hybrid approach, it was considered the most appropriate way to translate the ordinal BWS scores onto the cardinal scale in the present study (Rowen et al., 2015). To the authors' knowledge, all published studies that have used this mapping approach to anchor BWS data, have used cardinal data obtained using TTO tasks (Netten et al., 2012, Ratcliffe et al., 2016a, Chen et al., 2019). Interestingly, no previous published studies have used cardinal utilities elicited using the DCE_{TTO}.

Despite the lack of their use previously in this way, for the purposes of the present study it was considered to be most appropriate to use the DCE_{TTO} to obtain cardinal utilities that could be used to map adolescent scores onto the 1-0 QALY scale. Primarily, this was due to the wealth of evidence demonstrating successful delivery of DCE_{TTO} in the format of an online survey, allowing the researcher to collect data rapidly for a large sample (Norman et al., 2016, Mulhern et al., 2018). This was considered to be a key advantage over a small sample interviewer-led TTO survey as it provided the flexibility to also enable an adult value set to be generated. Furthermore, the DCE tasks that were used in the previous survey, which had been developed with involvement of patient and public involvement representatives, would require minimal adaptation to incorporate a duration attribute.

7.2 Aim

The aim of this stage was to use BWS methods to develop an adolescent value set for CARIES-QC-U. The specific objectives required to meet this aim were as follows:

- To undertake a population-based valuation survey with adolescents, using BWS tasks to determine their preferences;
- To undertake a population-based valuation survey with adults, using DCE_{TTO} tasks to determine their preferences;
- To model the adolescent BWS data to generate latent utility values for all health states;
- To model the adult DCE_{TTO} data to generate cardinal utility values for all health states;
- To rescale adolescent ordinal values onto the 1-0 full health-dead scale using cardinal values derived from adults to produce a value set that provides utility values for each health state defined by the CARIES-QC-U classification system

7.3 Method

The final stage of the overall study consisted of two online valuation surveys for CARIES-QC-U; a BWS survey for completion by adolescents, and a DCE_{TTO} survey for completion by adults. The purpose of the latter was to determine cardinal utilities to which the BWS values could be mapped.

7.3.1 Survey framing and presentation

Two surveys (appendices T and U) were developed with SurveyEngine (SurveyEngine GmbH, Berlin, Germany) (SurveyEngine (SurveyEngine GmbH, Berlin, Germany) GmbH, Berlin, Germany), based upon the same framing and presentation that was used in the previous chapter. The survey for adult participants was intended to be as similar as possible to the adolescent survey, with the obvious exception of the task itself, to minimise any differences surrounding the context in which the tasks were completed.

The first part of the survey contained basic demographic questions surrounding age, gender and ethnicity. Postcodes were requested to determine the geographical spread of participants amongst the devolved nations of England, Wales, Scotland and Northern Ireland, and also to determine levels of deprivation in accordance with the most recent indices of multiple deprivation tools for the respective nations (NISRA, 2017, Ministry of Housing, 2019, Stats Wales, 2019, Scottish Government, 2020). The adult survey also enquired into participants' marital and employment status to determine whether the sample was nationally representative for these factors.

This section was followed by three questions regarding health and dental health. The first was the general health question used in the World Health Survey Individual Questionnaire, which asks 'in general, how would you rate your health today?' The standard response levels for this question ('very good', 'good', 'moderate', 'bad' and 'very bad') were adapted to improve comprehensibility for a younger population, with 'moderate' being exchanged for 'OK', in line with suggestions from the Young People Panel. The global question from CARIES-QC ('how much of a problem are your teeth for you?') was also included to determine self-reported dental health, with the response options of 'not at all', 'a bit' and 'a lot'. The final question in this section sought to evaluate participants' caries history through enquiring as to whether they had ever had a filling or a tooth taken out because it had a hole or cavity.

Five questions from CARIES-QC (shown in Table 6-1 previously) relating to the items selected to form the classification system were included as a warm-up exercise, with the same three response levels (‘not at all’, ‘a bit’ and ‘a lot’). These were intended to familiarise participants with the wording used in the tasks, and to engage them in thinking about teeth. The survey asked participants to think about their teeth ‘today’ when answering these questions, as in the previous stage. The word ‘today’ was inserted using capital letters, with a reminder for the participant to think about their teeth ‘today’ included with every question. This was in acknowledgement of adolescents’ difficulty in recalling impacts from the present day only, as noted in the previous stage, whereby participants had a tendency to base their responses to these questions upon their previous experiences. This addition was recommended by the Young People Panel.

The same basic information about tooth decay was provided in these surveys as was provided in the previous survey, accompanied by the photograph of a decayed tooth to illustrate the condition. Again, it should be noted that the photograph was selected by child patient and public involvement (PPI) representatives for an unrelated study, as they felt it clearly showed the condition in question, but did not cause distress.

A ‘walkthrough’ was incorporated into each survey, to demonstrate how the task should be answered, followed by a practice question. For the adolescent survey, this practice question also acted as a dominance test. Participants’ choice from the practice question was reaffirmed, with a description of the alternative choice(s) that they did not select. Participants were asked to confirm they were happy to proceed to the main questions, otherwise they were able to do the practice question a second time.

Next, based upon the findings from the previous stage, adolescent respondents were allocated eight tasks to complete, whilst adult participants were allocated nine tasks. Screenshots of these tasks can be seen in Figure 7-1 and Figure 7-2.

3. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

My teeth hurt me **a lot**	My teeth annoy me **a lot**	My teeth keep me awake **a bit**	My teeth make it **really** hard to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○

Figure 7-1: Screenshot of the BWS survey for adolescent participants

3. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a bit**	My teeth do **not** hurt me **at all**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a lot**	My teeth keep me awake **a bit**
My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth make me cry **a lot**	My teeth do **not** make me cry **at all**
You will live for **2 years** in this state and then you will die	You will live for **2 years** in this state and then you will die
Which one would you prefer?	
○	○

Figure 7-2: Screenshot of the DCE$_{TTO}$ survey for adult participants

Adult participants were not asked to take any perspective other than their own when completing the tasks. The debate surrounding perspectives was outlined previously in section 2.6.1. The experimental design for each survey is described in the next section.

The surveys both concluded with two summary questions, regarding the participants' difficulty of understanding and difficulty of choice. Three response options were provided, as in the previous stage, following the recommendations of the PPI representatives ('easy', 'difficult' and 'somewhere in the middle'). The full surveys, following adjustments after piloting, are shown in appendices T and U.

7.3.2 Selection of health state profiles for valuation

Prior to valuation, it is necessary to select the health state profiles to be valued. The methods used to select health states differ depending on whether the classification system has independent dimensions, or whether the classification system consists of items that form a unidimensional component. For example, EQ-5D is considered as having independent dimensions, whereas CORE-6D, a condition-specific measure for common mental health problems, was considered as having a unidimensional component since five of the six items moved together (Mavranezouli et al., 2013). A unidimensional component occurs where items do not occur independently, for example, where for most people having problems with one item means they will also have the same level of problems with another item since they are capturing closely related symptoms that co-occur. The factor and Rasch analyses conducted in chapter 5 indicated that the CARIES-QC-U classification system had a single factor that covered four of the items in the classification system ('*annoy*' was the other item). Whilst this reflects unidimensionality in the underlying concept of OHRQoL, it does not indicate that the items co-occur. Instead, this was established using the original CARIES-QC validation dataset (the dataset that Rasch and psychometric analyses were conducted on in chapter 5), where cross tabulations of the responses to the items in the classification system were generated, and the health states defined by the measure were further generated. These analyses found that across the 200 respondents 81 different states were reported and across these there was a large variety across response option combinations across the items (e.g. 02001, 11021, 20011, 22001) demonstrating that the items do not co-occur. It is therefore appropriate to use approaches for selecting health state profiles for valuation that assume independence between the items.

Best-worst scaling tasks

As shown in Table 7-1, the classification system for CARIES-QC-U, contains five attributes (*'hurt'*, *'annoy'*, *'hard to eat'*, *'kept awake'* and *'cried'*). Each attribute can take one of three levels ('not at all', 'a bit' and 'a lot'). The maximum number of possible health states defined by the classification system is the product of the number of levels across all attributes. In this case it is possible to define 243 health states (3^5).

For the BWS survey a full factorial design was used, comprising all 243 health states, so that every health state was valued directly. This was possible due to the large sample size. This approach was also used in the adolescent BWS valuation of EQ-5D-Y; a generic paediatric PBM that has the same number of attributes and levels in each attribute (Dalziel et al., 2020). For each respondent, health states were randomly selected from the 243 possible health states. This ensured that each health state was valued an approximately equal number of times.

Discrete choice experiment with duration tasks

In-keeping with the literature, the DCE_{TTO} duration attributes included four levels; one year, four years, seven years and ten years (Rowen et al., 2018a, Rowen et al., 2018b, Norman et al., 2019, Mulhern et al., 2020). Each health state within the DCE_{TTO} was simply labelled A or B, to prevent potential heuristics due to label content (de Bekker-Grob et al., 2010).

There is a requirement to use at least as many choice sets as there are coefficients to be estimated. Including duration, there are six three-level attributes for the DCE_{TTO}, which gives 43 coefficients when all main effects and interactions between the dimensions and duration are included. This can be calculated using the formula below, which indicates the degrees of freedom:

number of parameters = (levels-1)*(duration levels)*number of dimensions+(number of levels for dimension-1)

Table 7-1: The CARIES-QC-U classification system with relevant variable names

Dimensions	Level	Variables	Health state descriptors
Hurt	0	Hurt0	My teeth do not hurt me at all
	1	Hurt1	My teeth hurt me a bit
	2	Hurt2	My teeth hurt me a lot
Annoy	0	Annoy0	My teeth do not annoy me at all
	1	Annoy1	My teeth annoy me a bit
	2	Annoy2	My teeth annoy me a lot
Kept awake	0	Awake0	My teeth do not keep me awake at all
	1	Awake1	My teeth keep me awake a bit
	2	Awake2	My teeth keep me awake a lot
Hard to eat	0	Eat0	My teeth do not make it hard to eat some foods
	1	Eat1	My teeth make it a bit hard to eat some foods
	2	Eat2	My teeth make it really hard to eat some foods
Cried	0	Cry0	My teeth do not make me cry at all
	1	Cry1	My teeth make me cry a bit
	2	Cry2	My teeth make me cry a lot
Duration* in Life Years	**Soft launch variable**		**Main survey variable**
	LY1 (1 year)		LY6m (6 months)
	LY4 (4 years)		LY12m (1 year)
	LY7 (7 years)		LY18m (1 year 6 months)
	LY10 (10 years)		LY24m (2 years)

Notes: *attribute included in the DCE$_{TTO}$ survey only

If we add in the relevant numbers for the present study this would be:

(3-1)*(4)*5+(4-1) = 43

To minimise participant fatigue (and resultant errors), and maximise completion rates, each participant was presented with nine tasks. To allow for this, the number of choice tasks chosen was 120 and the design was blocked into groups of nine. The design was generated using the d-create command; an add-in on Stata (StataCorp LLC, Texas, USA) developed by Professor Arne Risa Hole at the University of Sheffield (Risa Hole, 2015). The dcreate command generates a D-efficient design and uses the modified Federov algorithm (Cook and Nachtrheim, 1980, Zwerina et al., 1996, Sandor and Wedel, 2001, Carlsson and Martinsson, 2003).

The lead researcher (HJR) attended a course on Discrete Choice Experiments (Health Economics Research Unit, The University of Aberdeen) to gain training in DCE design and was supported by a health economist (DR) with expertise in DCE methodologies. Further to this, direct communication was made with the d-create command developer via email to confirm the best approach.

7.3.3 Survey administration

Surveys were administered online, as the previous stage demonstrated that adolescents were able to complete surveys independently using this format. The surveys were available for completion on desktop computers, tablet computers or mobile phones. The surveys were hosted by SurveyEngine (SurveyEngine GmbH, Berlin, Germany) and promoted on a number of online survey platforms. These platforms have large numbers of subscribers and members, known as panels, who regularly undertake surveys in the field of market and opinion research. Panel members are typically notified by email when a relevant survey becomes available. Alternatively, members are able to view the live surveys by visiting the respective websites. The member can then decide whether to commence the survey, providing they meet the stated criteria. Participation in each individual survey is entirely voluntary; there is no obligation for panel members to complete any survey.

Participants for the valuation surveys were recruited from across the United Kingdom (UK), though a quota for recruitment from each of the devolved nations was not defined.

7.3.4 Participants

Participants who met the following inclusion and exclusion criteria were invited to take part in the surveys.

Inclusion criteria

- Member (or child of a member) of a survey platform managed by a market research agency and commissioned by SurveyEngine (SurveyEngine GmbH, Berlin, Germany)
- Resident in the UK
- Aged 11-16 (adolescent survey)
- Aged over 18 years (adult survey)

Exclusion criteria

- Unable to read English language

7.3.5 Sample size and sampling methods

A total of 700 participants was considered to be sufficient to produce stable data for either a DCE_{TTO} or BWS survey. This was in line with previous studies of this type (Coast et al., 2008a, Norman et al., 2016), and was confirmed via personal correspondence with Dr Terry Flynn; an expert in this field. Furthermore, this would enable a soft launch to be conducted to allow the data to be reviewed and any necessary alterations to be made to the survey before completion by the remaining participants.

As such, a sample size of 700 adolescent participants and 700 adult participants were recruited for the surveys. The first 100 adolescents and 100 adults that were recruited formed the sample for the soft launch, whilst the remaining 600 adolescents and 600 adults formed the main survey sample. The final versions of both surveys can be seen in appendices T and U.

Quotas were assigned to ensure a minimum number of participants of each age (11, 12, 13, 14, 15 and 16 year-olds) were recruited for the BWS survey. Similar quotas were set for age bands for the adult DCE_{TTO} survey. Recruitment was due to continue until these quotas were met. It was acknowledged that slightly more participants could be recruited over the quota given the nature of an online survey.

7.3.6 Participant recruitment

Potential adult participants were able to access the adult DCE_{TTO} survey through their regular survey platform. On viewing the survey, they were able to read the participant information, from which they could decide whether to take part. The survey then asked these potential participants to consent electronically. For those who declined to consent, the survey terminated and screened out.

Adult panel members with children aged 11 to 16 years were also invited to view the study information relating to the adolescent survey. In the adolescent survey, parents were asked to provide consent for their child to participate. The adolescent could then take over the survey immediately, or if they were not available to take part at that time, they were provided with a link that they could use to start again from that point at a more convenient time. When adolescents commenced the survey they were asked to confirm their age to clarify that they met the inclusion criteria, and were provided with age-appropriate information about the study before being asked to assent. If a parent or child declined to consent or assent respectively, the survey terminated and screened out.

Participants were thanked for their participation in line with the standard practice of each individual survey platform. For the majority of platforms, these incentives consisted of points that accumulate as more surveys are completed, and can be exchanged for cash, vouchers or a donation to charity. The amount and value of these points varied from platform to platform.

7.3.7 Data collection

All survey data were stored by SurveyEngine (SurveyEngine GmbH, Berlin, Germany) on a UK-based server. Once the survey quotas had been met, the surveys were closed. The anonymised data from both surveys were accessed securely by the lead researcher (HJR) and downloaded from a secure area of the SurveyEngine (SurveyEngine GmbH, Berlin, Germany) webpages. More details on the protection of this data and the approaches to ensure confidentiality was maintained are provided in section 7.3.13.

7.3.8 Analysis

Soft launch

All data were checked and analysed using Stata/MP 16.0 (StataCorp LLC, Texas, USA) following the soft launch. This was predominantly to ensure that the surveys were functioning as intended. The marginal choice frequencies of BWS survey data were checked to ensure that options that had been selected as ‘best’ predominantly consisted of descriptors of lower severity levels, and vice versa. Similarly, regressions and anchored values were estimated for DCE_{TTO} data, and reviewed to identify any potential discrepancies. Any required changes to the surveys were made at this point before the main surveys were launched. The final versions of both surveys can be seen in appendices T and U.

Main survey

The main survey data were analysed using Stata/MP 16.0 (StataCorp LLC, Texas, USA). Completion rates were determined for each survey by comparing the number of participants who provided consent (or assent for the adolescent survey) with the number of participants who answered the final question. It was not possible to determine response rates for these surveys as they were promoted on a number of platforms and it could not be established how many people had seen them.

The types of devices on which participants undertook the surveys were analysed descriptively. Participants were required to complete the survey in one sitting. The median and range were calculated for the length of time taken by adolescents and adults to complete their respective surveys in their entirety. It was not possible to determine the length of time it took participants to complete the tasks alone.

Sociodemographic and health characteristics

Sociodemographic data were analysed using descriptive statistics to determine the diversity of the samples, which were compared against data from the UK 2011 Census (ONS, 2011). Postcode data, where provided, were used to determine geographical spread of participants across the devolved nations, and the level of deprivation of the area in which participants' resided. For the latter, the respective Index of Multiple Deprivation tools were employed for each devolved nation (NISRA, 2017, Stats Wales, 2019, Ministry of Housing, 2019, Scottish Government, 2020). The level of deprivation was determined in deciles for consistency between the respective tools, and recorded in a Microsoft® Excel® 2016 spreadsheet (Microsoft Corporation, Redmond, USA).

Participants' responses to questions surrounding their general health, dental health and receipt of previous fillings or extractions due to caries were analysed descriptively using Stata/MP 16.0 (StataCorp LLC, Texas, USA), and compared between adolescents and adults. Likewise, participant responses to the five questions from CARIES-QC relating to the five items within the classification system for CARIES-QC-U were described and compared between the two populations.

Difficulty and understanding

Participants' self-reported difficulty of understanding and difficulty of choice were analysed descriptively in Stata/MP 16.0 (StataCorp LLC, Texas, USA). In the absence of convention, the dominance test for the BWS task was determined to have been passed if the participant had correctly identified both the 'best' and 'worst' feature of the health state. The degree of concordance as to the 'best' feature alone was also reported, to acknowledge the participants' own perspective regarding the 'worst' feature. The dominance test for the DCE survey was determined to have been passed if the participant selected alternative 'B', the dominant health state over 'A' for this task. The proportion of participants meeting this criterion was reported.

Marginal frequencies

Marginal choice frequencies were determined to summarise the BWS data, using Microsoft® Excel® 2016 (Microsoft Corporation, Redmond, USA). These were calculated by dividing the number of times a domain level was selected as being 'best' or 'worst', by the number of times that domain level was available to be chosen within the survey (Dalziel et al., 2020). A graphical representation of the choice distribution was generated.

7.3.9 Modelling BWS data

Values for the BWS data were estimated in Stata/MP 16.0 (StataCorp LLC, Texas, USA) using multinomial (conditional) logistic regression in a random utility framework, in line with previous research (Coast et al., 2008a, Ratcliffe et al., 2016a, Huynh et al., 2017).

The equation to be estimated is specified as:

$$U_{id} = X_{dl}\beta'_{dl} + \varepsilon_{dl}$$

where U_{id} represents the utility that individual *i* derives from choosing dimension *d* and X_{dl} represents a vector of CARIES-QC-U attribute levels where *d* represents the 5 dimensions and *l*=0,1,2 represents the 3 severity levels, β'_{dl} is the vector of coefficients where, for example, β_{11} represents the coefficient for attribute 1 (hurt) level 1 ('a bit') and ε_{dl} is the random component (Ratcliffe et al., 2016b).

The conditional logit model considers all choice options as attribute levels, rather than complete states. The 'worst' choice data can then be appended to the 'best' choice data for each health state scenario, to form a best-worst pair (Flynn et al., 2007, Huynh et al., 2017). The model will then generate attribute level values on a latent scale (note this is not anchored on the full-health 1 to 0 scale required to generate QALYs) (Coast et al., 2008a). The dimension with the highest marginal frequency for 'best' at level 0 ('not at all') was selected as the reference for the model.

Logical consistency

The sign and magnitude of the coefficients were reviewed for significance at the 5% level, as well as logical consistency, whereby there is an expectation that the utility value decreases (or at least stays the same) as oral health-related quality of life (OHRQoL) deteriorates (e.g. level 1 coefficients should be smaller in absolute values than level 2 coefficients)(Rowen et al., 2018b). This is a requirement should there be any intention for the measure to be used to inform policy (Brazier et al., 2002, Rowen et al., 2018b). Any inconsistent adjacent levels were merged to produce a single utility decrement that can be applied when the dimension is at either of the respective levels, in line with previous studies of this type, before re-estimating a fully consistent model (Rowen et al., 2018b).

A significant, negative, and logically consistent coefficient for each domain level was considered ideal, as it would demonstrate that the attributes are functioning as intended, and that participants were able to distinguish between the levels in terms of severity.

Heterogeneity of coefficients

Heterogeneity is the part of the inherent variability between participants that can be attributed to specific sociodemographic or health-related characteristics of those participants (Ramaekers et al., 2013). The reduced sample approach was used to explore the presence of heterogeneity relating to gender, age, self-reported general and dental health, and previous caries experience. The coefficients for each re-estimated reduced sample model were then reviewed to establish the impact of each characteristic on the modelled results. This was an appropriate approach given the large size of the full sample.

Model robustness

It was then necessary to demonstrate that the regression model was robust, in that the results produced would not have been dissimilar had certain participant groups been excluded. In order to determine robustness, the model was re-estimated using a reduced sample approach to exclude participants who failed the dominance test, those who found it difficult to choose an answer and those who found the tasks difficult to understand (Bansback et al., 2012). Further models were estimated to determine the effect of excluding participants who completed the survey very quickly (less than three minutes), and those who took a very long time (over thirty minutes) to complete the survey. These parameters were determined by viewing a histogram of the time taken for participants to complete the survey. The extent to which the coefficients were affected by the exclusion of these participants was reviewed by the lead researcher (HJR) and checked by a second (DR) and a decision was made on whether to proceed with or without these participants (Lancsar and Louviere, 2006, Bansback et al., 2012).

All modelling was undertaken using Stata/MP 16.0 (StataCorp LLC, Texas, USA).

At the time that this study was undertaken, there were no UK-based, international or online courses available for the lead researcher (HJR) to attend in order to gain formal training in BWS analysis. Initial attempts to seek bespoke training from experts in this field were unsuccessful, though an external consultant (JS) with expertise in BWS, identified through Surveyengine (SurveyEngine GmbH, Berlin, Germany), agreed to provide support for this stage of the analysis.

7.3.10 Modelling DCE$_{\text{TTO}}$ data

Regressions were estimated for the DCE$_{\text{TTO}}$ data using the conditional logit model, in line with previous research of this type (McFadden, 1973, Bansback et al., 2012). The formula for this model is:

$$\mu_{ij} = \alpha_i + \beta_1 t_{ij} + \beta'_2 \mathbf{x}_{ij} t_{ij} + \varepsilon_{ij}$$

where μ_{ij} represents the utility of individual i for health state profile *j*, α_i is an individual specific constant term, ε_{ij} represents the error term, β_1 is the coefficient for duration in life years t and β'_2 represents the coefficients on the 10 interaction terms of duration and attribute variables composed of levels 1 and 2 of each quality of life attribute (where level zero is the baseline).

The values were converted from being on a latent scale, to the 1-0 full health to death QALY scale using the marginal rate of substitution, whereby each coefficient is divided by the coefficient for duration (Bansback et al., 2012): $\frac{\beta_{2ij}}{\beta_1}$. This produces a utility weight for each level of a dimension. The sum of the utility weights for the relevant level of each dimension can be added to 1 in order to generate anchored utility values (Bansback et al., 2012).

Logical consistency

As with the BWS, the sign and magnitude of the coefficients were reviewed for significance at the 5% level, as well as logical consistency (Rowen et al., 2018b). A definition of logical consistency is provided in section 7.3.9. A significant, negative coefficient for each domain level was considered ideal, as it would demonstrate that the attributes are functioning as intended, and that participants were able to distinguish between the levels in terms of severity.

Any inconsistent adjacent levels were merged to produce a single utility decrement that can be applied when the dimension is at either of the respective levels, in line with previous studies of this type (Rowen et al., 2018b). The fully consistent model was then estimated for the main effects.

Test of linearity

The duration attribute was modelled as a linear and continuous variable, hence it was necessary to confirm this assumption was correct through conducting a test of linearity (Payne et al., 2011, Rowen et al., 2018a). Duration was modelled as a categorical variable and the coefficients were plotted using Microsoft® Excel® 2016 (Microsoft Corporation, Redmond, USA) and examined for linearity.

Heterogeneity of coefficients

Interaction terms were included to explore the presence and impact of heterogeneity with regard to gender, age, employment, marital status, self-reported general and dental health, and previous caries experience. The unanchored coefficients were then reviewed to determine the impact of these different characteristics on the results. Positive interactions indicate that there is a lower utility decrement to the attribute level, so the overall utility values for health states are higher (closer to 1). This approach was used rather than a reduced sample approach since it easily enabled the impact to be determined across a large number of different characteristics, in particular where there were not a large number of respondents. The reduced sample approach, as applied to the BWS data, was also examined and found the same results but has not been reported here.

Model robustness

As with the BWS data, the robustness of the regression model was explored using a reduced sample approach to exclude participants who failed the dominance test, those who found it difficult to choose an answer and those who found the tasks difficult to understand (Bansback et al., 2012). Further models were estimated to explore the effect of excluding participants who completed the survey very quickly (less than three minutes), and those who took a very long time (over thirty minutes) to complete the survey. These parameters were determined by viewing a histogram of the time taken for participants to complete the survey. The extent to which the coefficients were affected by the exclusion of these participants was reviewed by the lead researcher (HJR) and confirmed by a second (DR) before a decision was made on whether to proceed with or without these values (Lancsar and Louviere, 2006, Bansback et al., 2012).

All modelling was undertaken using Stata/MP 16.0 (StataCorp LLC, Texas, USA).

The aforementioned course on Discrete Choice Experiments (Health Economics Research Unit, The University of Aberdeen) provided the lead researcher (HJR) with initial training in DCE analysis to gain training in DCE design, and further support was provided by a senior health economist (DR) with expertise in DCE methodologies.

7.3.11 Anchoring adolescent BWS values onto the QALY scale

As explained in section 7.1.1, the mapping approach was used to estimate cardinal values (DCE_{TTO}) for the latent BWS values for all health states. The DCE_{TTO} and BWS regressions were then plotted and the relationship was reviewed for linearity.

Through the use of the conditional logit model as described above, values were generated from both the BWS and $\mathrm{DCE}_{\mathrm{TTO}}$ data for each of the 243 health states. In Microsoft® Excel® 2016 (Microsoft Corporation, Redmond, USA), the following formula was applied to these data:

$$DCETTO_j = f\left(BWS_j\right) + \varepsilon_j$$

where DCETTO_j represents the mean $\mathrm{DCE}_{\mathrm{TTO}}$ value of health state *j*, BWS represented the modelled latent utility value for health state *j*, and ε_j is the error term (Rowen et al., 2015). This assumes a linear approach with an intercept.

The mean health state values were transferred into Stata/MP 16.0 (StataCorp LLC, Texas, USA), and the relationship was plotted and reviewed for linearity. Ordinary least squares regressions were estimated to generate the mapping models mapping the BWS latent values onto $\mathrm{DCE}_{\mathrm{TTO}}$ values (Rowen et al., 2015). The inclusion of squared and cubic terms were explored to determine the most appropriate model specification (Rowen et al., 2015). The mapped utility predictions were then plotted and compared to the modelled BWS latent values and $\mathrm{DCE}_{\mathrm{TTO}}$ values.

7.3.12 Patient and public involvement

Adolescent members of the aforementioned Young People Panel were involved in trialing the adolescent BWS survey to determine its ease of use, as well as the suitability of the wording and formatting. This included the participant information section at the start of the survey. Whilst much of this survey was similar to the one used in the previous stage, there were a few notable differences that required the Young People Panel’s attention. Similarly, the two parent representatives for the wider study were involved in the development of the parent information for this survey, as well as the consent page. The parent representatives also trialed the initial drafts of the adult DCE survey and gave suggestions for improvements, the majority of which related to formatting details, which were incorporated accordingly.

Alongside their involvement in the survey development, the two parent PPI representatives and four other adult PPI representatives who agreed to assist with this stage were involved in a review of the duration attribute following the pilot stage (soft launch). Further details of this can be found in section 7.4.1.

7.3.13 Ethical approval

Two further substantial amendments (substantial amendments 2 and 3) were submitted to the NHS Research Ethics Committee for this final stage of the study. A favourable opinion was received for substantial amendment 2 (appendix V), which sought approval for the use of the two surveys, which were somewhat different to that which had been used in the previous stage. A favourable opinion was also received for substantial amendment 3 (appendix W), which sought approval to confirm the sample size.

Whilst this stage was not expected to raise any significant ethical concerns, due consideration was given to the following aspects.

Data protection

In line with the European Union General Data Protection Regulation 2016/679 and the UK application of this regulation in the Data Protection Act 2018, participants were provided with specific details of how their data would be used and stored. The study sponsor, Sheffield Teaching Hospitals Foundation Trust, acted as the data controller for the study. The Data Processing Agreement held with SurveyEngine (SurveyEngine GmbH, Berlin, Germany) complied with GDPR and confirmed their responsibilities as the data processors.

Confidentiality

Survey data were collected by Suveyengine on a UK-based server. Data were anonymised, and each participant was allocated an identification number. No participant names or email addresses were stored, and the survey platform that the data originated from was not identifiable. Participant IP (Internet Protocol) addresses were masqueraded to produce pseudo-IP addresses.

The spreadsheets containing anonymised data were downloaded from the SurveyEngine (SurveyEngine GmbH, Berlin, Germany) webpages and stored to a secure drive, accessible only by the research team.

7.4 Results

This section will first consider the results from the soft launch and any necessary alterations that were made to the survey prior to its administration to the remaining participants. The main survey results will then be presented, followed by the results from modelling the BWS and DCE_{TTO} data. Finally, the anchoring of the BWS values onto the QALY scale using mapping is described.

7.4.1 Soft launch

The results from the first 99 adolescent and 101 adult participants were analysed initially following the soft launch.

The marginal frequencies for the adolescent BWS were reviewed (see

Table 7-2) and were as expected. As such, the survey continued without the need for further adjustments, and hence the soft launch sample was incorporated within the main sample. The results of these are reported together in section 7.4.2.

Nonetheless, issues were observed with the adult DCE$_{TTO}$ survey soft launch results. Primarily, the estimated utilities from the DCE$_{TTO}$ survey (seen in

Table 7-5) were somewhat greater than had been expected. Importantly, it seemed implausible that ‘my teeth hurt me a lot’ could give a loss of almost 0.6 on the 1-0 full health to death scale used to generate QALYs. Further analyses of the sample were undertaken to determine whether discrepancies amongst the sample characteristics may have contributed to these results. The sociodemographic and health characteristics, and self-reported understanding of the soft launch adult sample can be seen in

Table 7-3 and Table 7-4 below, though these did not reveal a cause for concern. National population norms derived from the 2011 UK census are provided for reference (ONS, 2011).

After discussing other possible reasons for this, the researchers considered that the durations attached to each attribute (one, four, seven or ten years) may have been too long. It was possible that such long durations had substantially influenced participant responses to the DCE_{TTO}, meaning that too much weight may have been placed on the duration attribute, and also that it may not be realistic to imagine experiencing these impacts for such a long duration.

Table 7-2: Marginal choice frequencies for data obtained from soft launch of adolescent BWS survey

Attribute	Descriptor	Best	Worst
Hurt0	My teeth do not hurt me at all	**0.602**	0.050
Hurt1	My teeth hurt me a bit	0.068	0.166
Hurt2	My teeth hurt me a lot	0.030	0.474
Annoy0	My teeth do not annoy me at all	0.479	0.114
Annoy1	My teeth annoy me a bit	0.183	0.087
Annoy2	My teeth annoy me a lot	0.090	0.195
Awake0	My teeth do not keep me awake at all	0.328	0.073
Awake1	My teeth keep me awake a bit	0.048	0.160
Awake2	My teeth keep me awake a lot	0.050	0.307
Eat0	My teeth do not make it hard to eat some foods	0.350	0.120
Eat1	My teeth make it a bit hard to eat some foods	0.171	0.159
Eat2	My teeth make it really hard to eat some foods	0.093	0.272
Cry0	My teeth do not make me cry at all	0.466	0.071
Cry1	My teeth make me cry a bit	0.027	0.224
Cry2	My teeth make me cry a lot	0.027	**0.521**

Notes: Highest best and worst frequencies are displayed in **bold.** Hurt0: my teeth do not hurt me at all; Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy0: my teeth do not annoy me at all; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake0: my teeth do not keep me awake at all; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat0: my teeth do not make it hard to eat some foods; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry0: my teeth do not make me cry at all; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot

A selection of DCE_{TTO} with differing durations were discussed with the two parent representatives for the study, and the four adult PPI representative. The selection of DCE_{TTO} incorporated some with the current durations in the survey (one, four, seven or ten years) as well as some with shorter durations (six months, one year, one year and six months, two years) The discussions centred on the length of the duration, and the extent to which it influenced their decision. All representatives felt that the shorter durations influenced their decision less than the longer durations. As such, the DCE_{TTO} survey was adjusted so that the duration attribute for the tasks contained shorter time periods (one, four, seven or ten years were exchanged for 6 months, one year, one year and six months, two years respectively).

Another issue with the adult DCE_{TTO} survey was noted, in that a dominance test had not been included. Whilst the previous stage identified possible issues with the dominance test for DCE_{TTO} tasks in particular, the researchers understood that the inclusion of a dominance test is considered to be conventional. As such, the levels for each attribute in the practice question were altered to form a dominance test, as in the BWS survey. Acknowledging the aforementioned issues, duration was kept the same for each alternative in the dominance question.

No changes were made to the BWS survey, so the BWS data from the soft launch were able to be analysed as part of the main survey. As the DCE_{TTO} survey had been adjusted, the data obtained during the soft launch were excluded from further analyses.

Table 7-3: Sociodemographic characteristics of adult participants in the soft launch of the DCE$_{TTO}$ survey

Sociodemographic and health characteristics	Adults n=99 (%)	Population norms %
Gender		
Male	58 (57.4)	49.1[a]
Female	43 (42.6)	50.9[a]
Other	0 (0.00)	-
Age		
18-24	1 (1.0)	12.0[b]
25-34	6 (6.0)	17.0[b]
35-44	14 (13.9)	17.7[b]
45-64	50 (49.5)	32.5[b]
65+	30 (29.7)	20.8[b]
Prefer not to say	-	
Ethnicity		
White	96 (95.1)	87.2[a]
Asian/Asian British	0 (0.0)	6.2[a]
Black/African/Caribbean/Black British	2 (2.0)	3.0[a]
Mixed/Multiple ethnic groups	2 (2.0)	-
Other ethnic group	0 (0.0)	2.9[a]
Prefer not to say	1 (1.0)	-
Main activity		
In employment or self-employment	45 (44.6)	61.7[c]
Retired	32 (31.7)	13.9[c]
Housework	7 (6.9)	4.3[c]
Student	2 (2.0)	9.3[c]
Seeking work/unemployed	6 (5.9)	4.4[c]
Long term sick	5 (5.0)	4.3[c]
Prefer not to say	1 (1.0)	-
Other	3 (3.0)	2.2[c]
Marital status		
Single	21 (20.8)	35.9[a]
Married/partner	68 (67.3)	47.0[a]
Separated/divorced	8 (7.9)	7.7[a]
Widowed	3 (3.0)	9.4[a]
Prefer not to say	1 (1.0)	-

Notes: [a]=proportion of total UK population [b]=proportion of UK adult population (aged over 18 years) [c]=proportion of English adult population (aged over 16 years)

Table 7-4: Health-related characteristics of adult participants in the soft launch of the DCE_{TTO} survey

Health characteristics	**Adults n=99 (%)**
Self-reported general health: in general, how would you rate your health today?	
Very good	12 (11.9)
Good	53 (52.5)
OK	29 (28.7)
Bad	6 (5.9)
Very bad	1 (1.0)
Self-reported dental health: how much of a problem are your teeth for you today?	
Not at all	53 (52.5)
A bit	37 (36.6)
A lot	11 (10.9)
Self-reported caries experience: have you ever had a filling or a tooth taken out because it had a hole or cavity?	
Yes	85 (84.2)
No	16 (15.8)
Participant understanding	
Did you find the tasks:	
Easy to understand	68 (67.3)
Somewhere in the middle	2 (2.0)
Difficult to understand	31 (30.7)
Did you find it:	
Easy to choose an answer	43 (42.6)
Somewhere in the middle	18 (17.8)
Difficult to choose an answer	40 (39.6)

Table 7-5: Regression results and anchored utility decrements using soft launch data of DCE_{TTO} survey for CARIES-QC-U

Variables	Estimated regressions	Anchored values
Hurt1_LY	-0.066***	-0.172
	(0.000)	
Hurt2_LY	-0.228***	-0.590
	(0.000)	
Annoy1_LY	0.001	0.002
	(0.978)	
Annoy2_LY	-0.049**	-0.126
	(0.016)	
Awake1_LY	-0.020	-0.051
	(0.314)	
Awake2_LY	-0.114***	-0.295
	(0.000)	
Eat1_LY	-0.026	-0.068
	(0.207)	
Eat2_LY	-0.072***	-0.188
	(0.001)	
Cry1_LY	-0.050**	-0.130
	(0.014)	
Cry2_LY	-0.146***	-0.378
	(0.000)	
LY	0.386***	
	(0.000)	
Observations	1,818	
Log likelihood	-500.6	
Rho-squared	0.205	

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

7.4.2 Main survey

Of the 858 adolescents who agreed to take part in the survey (including the soft launch), a total of 723 completed the survey; a completion rate of 84.3%. In contrast, 1220 adults consented to participate in the main survey, but only 626 completed the survey, giving a completion rate of 51.3%. The majority of participants, both adolescents and adults, completed the survey on their desktop computers (n=660; 91.3% and n=570; 91.1% respectively). The rest of the adolescents completed the survey on their tablet computers (n=63, 8.7%), whilst some adults used tablets or mobile phones (n=50, 78.0% and n=6, 1.0% respectively).

The time taken for adolescents to complete the BWS survey ranged from 2 minutes to 272 minutes, with a median time of 8 minutes (Figure 7-3). The time taken for adults to complete the DCE_{TTO} survey ranged between 2 minutes to 95 minutes, with a median time of 8 minutes (Figure 7-4). It was not possible to determine the proportion of time that participants spent on completing the valuation tasks.

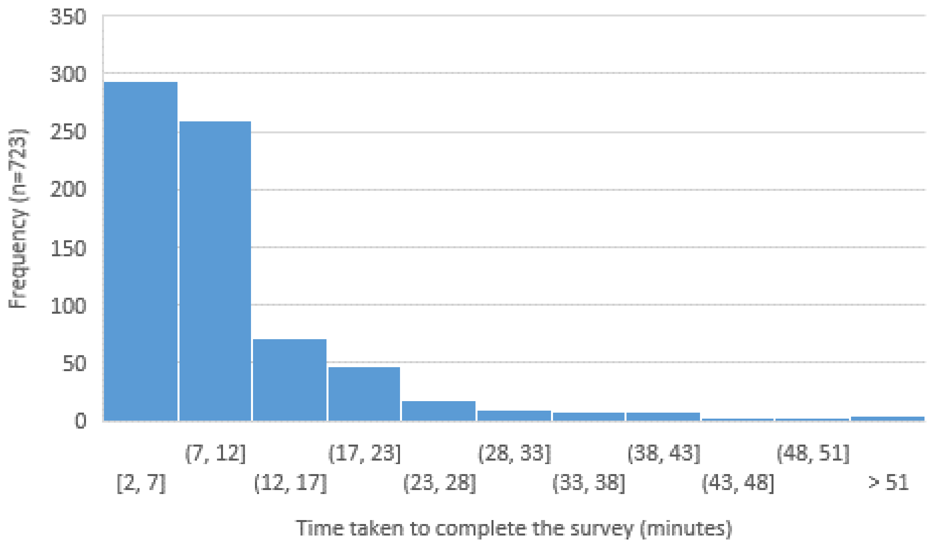


Figure 7-3: Histogram displaying the variation in time taken to complete the adolescent BWS survey

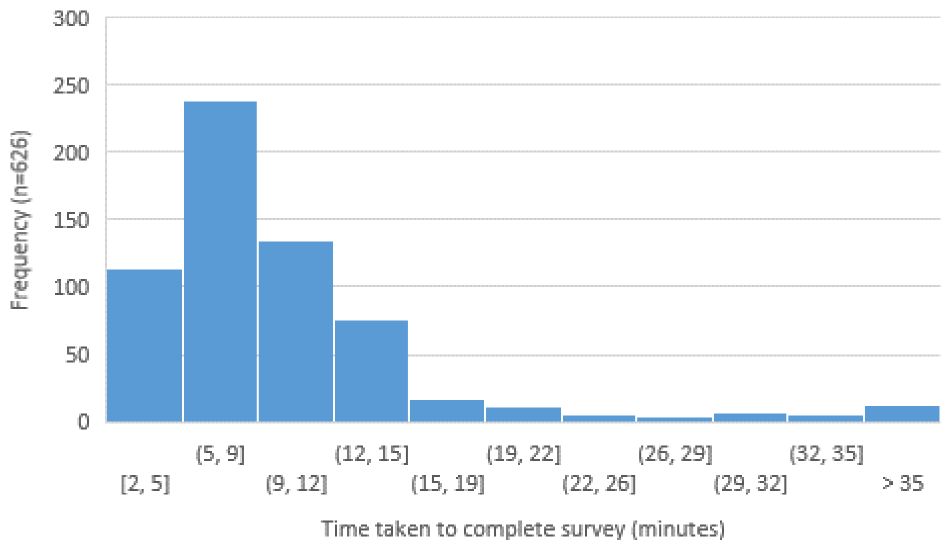


Figure 7-4: Histogram displaying the variation in time taken to complete the adult DCE_{TTO} survey

Sociodemographic and health characteristics

Sociodemographic and health characteristics are provided in

Table 7-6, alongside population norms derived from the 2011 UK Census for reference (ONS, 2011).

A total of 723 adolescents completed the BWS valuation survey, with similar proportions of male (n=387; 53.5%) and female (n=333; 46.1%) participants, as shown in

Table 7-6. This slightly higher proportion of males is not reflected in the wider population, whereby there are more females than males (ONS, 2011). Participants resided in each of the devolved nations, though the majority of participants (78.7%) lived in England, in line with the population norms. Relatively few participants resided in Scotland compared to the Census data (4.7% compared to 8.2%). It was not possible to locate a small proportion of participants, as 60 (8.3%) did not know their postcode, and 19 (2.6%) provided a postcode that was not recognised.

Participants were aged between 11 and 16 years, with a comparable number of participants of each age within this range. The sample comprised a higher proportion of 13-year-olds and a lesser proportion of 16-year-olds than would be reflected in the wider population (21.0% compared to 16.6% and 12.5% compared to 17.1% respectively). A variety of ethnicities were represented within the sample, though the majority of participants (84.2%) identified as being White. Population data were unavailable for mixed or multiple ethnic groups as this classification of ethnicity did not align with the Census data, though it would appear that other ethnic groups were underrepresented in this sample (0.6% compared to population norm of 2.9%). Whilst there was a range in the levels of deprivation within the sample, as determined by the most recent Indices of Multiple Deprivation for each devolved nation, almost half (48.8%) resided in the most deprived five deciles of the UK (NISRA, 2017, Stats Wales, 2019, Ministry of Housing, 2019, Scottish Government, 2020).

In the main adult DCE_{TTO} survey there were a total of 626 participants, also with similar proportions of males and females (n=288; 46.0% and n=336; 53.7%). Whilst this sample reflected the slightly higher proportion of females in the general population, it is possible that there is an under-representation of males (46.0% in the sample compared to 49.1% in the general population). Again, the majority of the adult sample resided in England (n=519; 82.9%), though there was representation from each of the devolved nations. The sample comprised similar proportions of participants from Scotland and Wales (5.8% and 5.3% respectively), which does not reflect the higher proportion of the UK population residing in Scotland (8.2% in Scotland compared to 4.7% in Wales). It was not possible to locate 28 adult participants (4.5%) as they either did not provide their postcode (n=17; 2.7%) or provided a postcode that was not recognised (n=11; 1.8%).

Most participants were within the 45-64 year-old age bracket (n=191; 30.5%), whilst only 11.5% (n=72) of the sample was comprised of younger adults aged 24-35 years. These proportions broadly reflect the wider UK population. In terms of ethnicity, the adult sample was less diverse than the adolescent sample, with over 90% (n=659) of participants describing themselves as White. Black, African, Caribbean and Black British ethnic groups were under-represented, forming just 1.4% of the sample, but comprising 3% of the UK population. As with the adolescent survey, the ethnicity categories used did not align with the UK Census data, though other ethnic groups appear to be under-represented, comprising just 0.6% of the sample, but 2.9% of the population.

Table 7-6: Sociodemographic characteristics of participants who completed the valuation surveys and population norms

Sociodemographic characteristics	Adolescents n=723 (%)	Adults n=626 (%)	Population norms %
Gender			
Male	387 (53.5)	288 (46.0)	49.1[a]
Female	333 (46.1)	336 (53.7)	50.9[a]
Other	3 (0.4)	2 (0.3)	-
Country of residence			
England	588 (78.7)	519 (82.9)	84.3[a]
Scotland	34 (4.7)	36 (5.8)	8.2[a]
Wales	28 (3.9)	33 (5.3)	4.7[a]
Northern Ireland	13 (1.8)	10 (1.6)	2.8[a]
Unknown	79 (10.9)	28 (4.5)	-
Age			
11	106 (14.7)	-	15.9[b]
12	124 (17.2)	-	16.3[b]
13	152 (21.0)	-	16.6[b]
14	126 (17.4)	-	16.9[b]
15	123 (17.0)	-	17.1[b]
16	90 (12.5)	-	17.1[b]
18-24	-	72 (11.5)	12.0[c]
25-34	-	124 (19.8)	17.0[c]
35-44	-	112 (17.9)	17.7[c]
45-64	-	191 (30.5)	32.5[c]
65+	-	127 (20.3)	20.8[c]
Prefer not to say	2 (0.3)	-	
Ethnicity			
White	609 (84.2)	563 (89.9)	87.2[a]
Asian/Asian British	62 (8.6)	42 (6.7)	6.2[a]
Black/African/Caribbean/Black British	20 (2.8)	9 (1.4)	3.0[a]
Mixed/Multiple ethnic groups	27 (3.7)	7 (1.1)	-
Other ethnic group	4 (0.6)	4 (0.6)	2.9[a]
Prefer not to say	1 (0.1)	1 (0.2)	-
Main activity			
In employment or self-employment	-	343 (54.8)	61.7[d]

Main activity continued	**Adolescents n=723 (%)**	**Adults n=626 (%)**	**Population norms %**
Retired	-	133 (21.3)	13.9[d]
Housework	-	43 (6.9)	4.3[d]
Student	-	47 (7.5)	9.3[d]
Seeking work/unemployed	-	30 (4.8)	4.4[d]
Long term sick	-	25 (4.0)	4.3[d]
Prefer not to say	-	1 (0.2)	-
Other	-	4 (0.6)	2.2[d]
Marital status			
Single	-	178 (28.4)	35.9[a]
Married/partner	-	363 (58.0)	47.0[a]
Separated/divorced	-	57 (9.1)	7.7[a]
Widowed	-	25 (4.0)	9.4[a]
Prefer not to say	-	3 (0.5)	-
Deprivation deciles (IMD)			
1 (most deprived)	80 (11.1)	51 (8.2)	-
2	69 (9.5)	61 (9.7)	-
3	72 (10.0)	63 (10.1)	-
4	81 (11.2)	73 (11.7)	-
5	51 (7.1)	61 (9.7)	-
6	62 (8.6)	54 (8.6)	-
7	65 (9.0)	56 (9.0)	-
8	54 (7.5)	67 (10.7)	-
9	58 (8.0)	57 (9.1)	-
10 (least deprived)	52 (7.2)	55 (8.8)	-
Postcode not provided	60 (8.3)	17 (2.7)	-
Postcode not recognised	19 (2.6)	11 (1.8)	-

Note: [a]=proportion of total UK population [b]=proportion of UK adolescents aged 11-16 [c]=proportion of UK adult population (aged over 18 years) [d]=proportion of English adult population (aged over 16 years)

The majority of the sample were employed or retired (n=343; 54.8% and n=133; 21.3% respectively), with smaller proportions of participants either unemployed (n=30; 4.8%) or with long-term sickness (n=25; 4.0%). The sample had a higher proportion of participants who described their main activity as retired or housework than is reflected in the wider population (21.3% compared to 13.9%, and 6.9% compared to 4.3% respectively). Regarding marital status, most adult participants were married (n=363; 58.0%), with the second highest proportion of participants reporting themselves to be single (n=178; 28.4%). In the wider population, these proportions are more similar than is reflected in this sample (47.0% married and 35.9% single). As with the adolescent survey sample, almost half of the sample (49.4%) resided in the most deprived five deciles, as determined by the Indices of Multiple Deprivation for the respective devolved nations (NISRA, 2017, Ministry of Housing, 2019, Stats Wales, 2019, Scottish Government, 2020).

The self-reported general and dental health of adolescents and adults that participated in these surveys can be seen in

Table 7-7. Over half of the adolescents in the sample (n=383) reported their general health to be very good, whilst no adolescent participants reported their health to be very bad. Almost two thirds of adolescents (n=450) reported no problems with their teeth, whilst the remainder felt their teeth were ‘a bit’ (n=249; 34.4%) or ‘a lot’ (n=24; 3.3%) of a problem. Approximately half of the adolescents in this study (n=350) reported previous experience of caries, through having a filling or a tooth taken out.

Adult participants reported poorer general health than the adolescent sample, with only a fifth (n=126) describing it to be very good. A small proportion reported their health to be very bad (n=8; 1.3%). Similarly, half of the adult participants reported problems with their teeth (n=271; 43.3% ‘a bit’ and n=39; 6.2% ‘a lot). Most adults in this sample reported having a filling, or a tooth removed due to caries (n=498; 79.6%).

The results from the warm-up CARIES-QC questions can be seen in

Table 7-8. Whilst no impacts were reported by the majority of adolescents, over a third of participants experienced feeling annoyed about their teeth (n=220; 30.4% ‘a bit’ and n=24; 3.3% ‘a lot’), making this the most widely experienced impact in this sample.

CARIES-QC impacts were more commonly observed in the adult sample than the adolescent sample. Similar to the adolescent sample, feeling annoyed was also the most widely reported impact (n=239; 38.2 ‘a bit’ and n=47; 7.5% ‘a lot’), closely followed by finding some foods hard to eat (n=214; 34.2% ‘a bit’ and n=38; 6.1% ‘a lot’).

Table 7-7: Self reported general health and dental health of participants in the adolescent BWS and adult DCE_{TTO} surveys

Health-related characteristics	Adolescents n=723 (%)	Adults n=626 (%)
Self-reported general health: in general, how would you rate your health today?		
Very good	383 (53.0)	126 (20.1)
Good	272 (37.6)	297 (47.4)
OK	59 (8.2)	154 (24.6)
Bad	9 (1.2)	41 (6.6)
Very bad	0 (0.0)	8 (1.3)
Self-reported dental health: how much of a problem are your teeth for you today?		
Not at all	450 (62.2)	316 (50.5)
A bit	249 (34.4)	271 (43.3)
A lot	24 (3.3)	39 (6.2)
Self-reported caries experience: have you ever had a filling or a tooth taken out because it had a hole or cavity?		
Yes	350 (48.4)	498 (79.6)
No	373 (51.6)	128 (20.5)

Table 7-8: Participant responses to CARIES-QC warm-up questions

Responses to CARIES-QC	Adolescents n=723 (%)	Adults n=626 (%)
How much do your teeth hurt you?		
Not at all	573 (79.3)	452 (72.2)
A bit	134 (18.5)	163 (26.0)
A lot	16 (2.2)	11 (1.8)
How much do your teeth annoy you?		
Not at all	479 (66.3)	340 (54.3)
A bit	220 (30.4)	239 (38.2)
A lot	24 (3.3)	47 (7.5)
How much do you get kept awake by your teeth?		
Not at all	617 (85.3)	562 (89.8)
A bit	90 (12.5)	56 (9.0)
A lot	16 (2.2)	8 (1.3)
Do your teeth make it hard to eat some foods		
Not at all	550 (76.1)	374 (59.7)
A bit	151 (20.9)	214 (34.2)
A lot	22 (3.0)	38 (6.1)
How much have you cried about your teeth?		
Not at all	604 (83.5)	546 (87.2)
A bit	106 (14.7)	67 (10.7)
A lot	13 (1.8)	13 (2.1)

Difficulty and understanding

Most adolescent participants (n=501; 69.3%) found the tasks easy to understand, with only 51 (7.1%) participants reporting them to be difficult. The majority of adolescents found it easy to choose an answer, though a small proportion found it difficult (n=68; 9.4%). The dominance task was deemed to have been passed if the participant identified both the 'best' and the 'worst' attributes correctly. Against these criteria, just over half of the sample (n=389; 53.8%) passed the test. This rose to over four-fifths of the sample (n=604; 83.5%), however, when considering those participants who correctly identified only the best attribute. As the dominance test also functioned as the practice question, participants were able to repeat it a second time before proceeding with the main tasks. Of the 38 participants that repeated the dominance test a second time, 34 (89.5%) were able to correctly identify the 'best' attribute, whilst 12 (31.6%) were able to correctly identify both the 'best' and 'worst' attributes. None of these 12 participants had passed the test in their previous attempt. It is important to note that the identification of the 'worst' attribute is of lesser importance compared to the 'best' attribute. Whilst the selection of the 'best' attribute is based upon the principle that experiencing no impacts related to a condition is better than experiencing impacts, the selection of the 'worst' attribute may be related to an individual's own preferences. This was discussed in greater detail in section 6.6.1.

Table 7-9: Participants' self-reported difficulty of understanding, difficulty of choice and ability to pass the dominance test

Participant understanding	**Adolescents n=723 (%)**	**Adults n=626 (%)**
Did you find the tasks:		
Easy to understand	501 (69.3)	451 (72.0)
Somewhere in the middle	171 (23.7)	39 (6.2)
Difficult to understand	51 (7.1)	136 (21.9)
Did you find it:		
Easy to choose an answer	440 (60.9)	277 (44.3)
Somewhere in the middle	215 (29.7)	104 (16.6)
Difficult to choose an answer	68 (9.4)	245 (39.1)
Dominance test		
Pass*	389 (53.8)	578 (92.3)
Correctly identified 'best' attribute	604 (83.5)	-
Second attempt requested	**n=38**	**n=26**
Pass*	12 (31.6) of which 0 (0.0) had failed the first attempt	23 (92.0) of which 18 (78.3) had failed the first attempt
Correctly identified 'best' attribute	34 (89.5)	-

Notes: *both 'best' and 'worst' attributes needed to be identified correctly for adolescents to pass the BWS dominance test

Whilst the vast majority of adults (n=451; 72.0%) found the tasks easy to understand, over one-fifth found them difficult to understand (n=136; 21.9%). Interestingly, over a third of adults reported difficulty in choosing an answer (n=245; 39.1%). Most adults passed the dominance test (n=578; 92.3%). As the dominance test also functioned as a practice question, participants had the opportunity to take it a second time before proceeding to the main questions. A total of 26 participants took the practice question/dominance test a second time. Of these, five participants (19.2%) had passed the dominance test first time, and repeated the test a second time and passed again. The majority of those who repeated the test had initially failed the test, but managed to pass the second time (n=18; 78.3%).

Marginal frequencies for BWS

Marginal frequencies for the BWS survey, seen in Table 7-10, show that the dimension most consistently rated as best by adolescents in this sample, was ‘my teeth do not hurt me at all’ (rated best 61.4% of the times it was presented). The dimension most consistently rated as worst by adolescents was ‘My teeth make me cry a lot’, which was rated as worst 50.2% of the times it was presented. The best and worst choice frequencies are plotted on the graph in Figure 7-5, showing a decreasing trend between the ratings. This demonstrates that the level with no impacts from caries (‘not at all’) are more often chosen as ‘best’, whilst the level with significant impacts from caries (‘a lot’) was more often chosen as ‘worst’.

Table 7-10: Marginal choice frequencies for data obtained from adolescent BWS survey

Variables	Descriptor	Best	Worst
Hurt0	My teeth do not hurt me at all	**0.614**	0.047
Hurt1	My teeth hurt me a bit	0.087	0.166
Hurt2	My teeth hurt me a lot	0.040	0.489
Annoy0	My teeth do not annoy me at all	0.404	0.074
Annoy1	My teeth annoy me a bit	0.171	0.121
Annoy2	My teeth annoy me a lot	0.079	0.206
Awake0	My teeth do not keep me awake at all	0.341	0.081
Awake1	My teeth keep me awake a bit	0.074	0.141
Awake2	My teeth keep me awake a lot	0.056	0.305
Eat0	My teeth do not make it hard to eat some foods	0.346	0.107
Eat1	My teeth make it a bit hard to eat some foods	0.156	0.168
Eat2	My teeth make it really hard to eat some foods	0.077	0.277
Cry0	My teeth do not make me cry at all	0.455	0.070
Cry1	My teeth make me cry a bit	0.052	0.247
Cry2	My teeth make me cry a lot	0.044	**0.502**

Notes: highest best and worst frequencies are displayed in **bold**. Hurt0: my teeth do not hurt me at all; Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy0: my teeth do not annoy me at all; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake0: my teeth do not keep me awake at all; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat0: my teeth do not make it hard to eat some foods; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry0: my teeth do not make me cry at all; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot

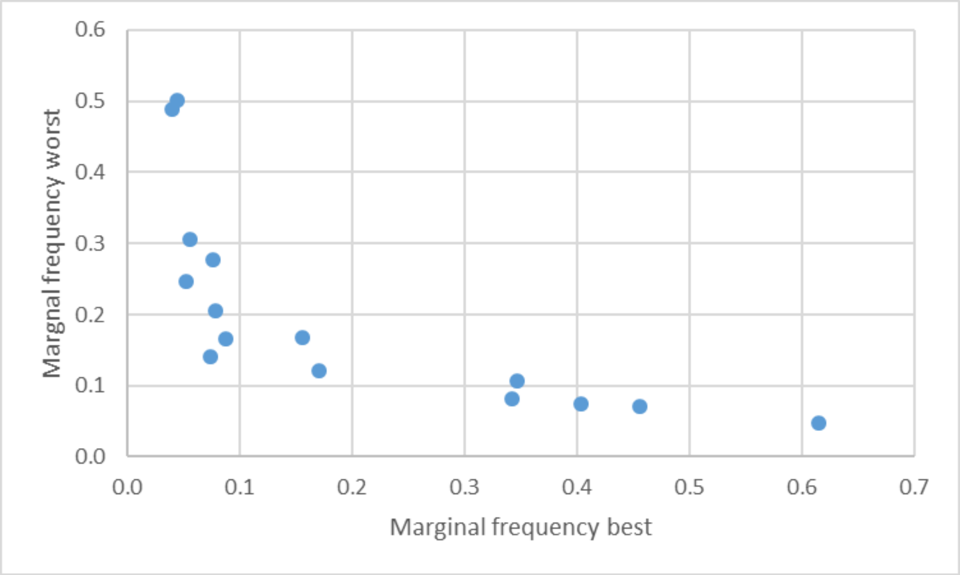


Figure 7-5: Plot of best and worst choice frequencies from adolescent BWS survey

7.4.3 Modelling BWS

The estimated regressions from the conditional logit model can be seen in Table 7-11. Hurt0 ('not at all') was chosen as the reference level for the model as it had the highest marginal frequency for 'best'. The coefficients are all seen to be negative and significant.

Heterogeneity of coefficients

Heterogeneity was explored using the reduced sample approach in relation to participant age, gender, self-reported general and dental health, and previous caries experience. The direction and significance of these coefficients (appendix X) were reviewed for differences. The coefficients remained similar for each model and all values remained negative and significant, with the exception of the coefficients for Cry0. These were no longer found to be significant for models including only 11-, 12-, and 14-year-old adolescents. Similarly, the Cry0 coefficients were no longer significant for models including only participants who reported themselves as having bad or very bad general health, current dental problems or previous experience of dental caries. There were a number of anomalies present in the model of participants with bad or very bad health, though this may have been impacted by the small sample of participants with this health characteristic.

Model robustness

Regressions were estimated for seven further models, each with a reduced sample. These excluded participants that failed the dominance test, reported difficulty understanding the tasks or difficulty in choosing an answer within the tasks, and combinations of these. These also explored robustness when excluding those that completed the survey in less than 3 minutes, and more than 30 minutes. The regressions can be found in appendix Y. Minimal differences were seen between the models, with all coefficients remaining negative and significant throughout.

Logical consistency

The model coefficients in Table 7-11 were reviewed for logical consistency. It can be seen that each worsening level for each attribute has a lower value than the previous level, demonstrating that the model is fully consistent in its current state.

The largest decrements at the lower severity levels can be seen for Awake1, and Cry1, suggesting that the movement from having no problems in these dimensions to some problems has a larger impact for these dimensions than for the other dimensions. In contrast, the largest decrements for the most severe level are observed for Cry2, closely followed by Hurt2, indicating that these have the largest impact on utility when at the most severe level and hence have the greatest relative importance across the dimensions.

Table 7-11: Estimated regressions from the conditional logit model using data from the BWS survey for CARIES-QC-U

Variables	**Standard model**
Hurt0	-
	-
Hurt 1	-2.235***
	(0.000)
Hurt2	-3.406***
	(0.000)
Annoy0	-0.959***
	(0.000)
Annoy1	-1.989***
	(0.000)
Annoy2	-2.720***
	(0.000)
Awake0	-0.866***
	(0.000)
Awake1	-2.322***
	(0.000)
Awake2	-2.827***
	(0.000)
Eat0	-0.949***
	(0.000)
Eat1	-1.874***
	(0.000)
Eat2	-2.543***
	(0.000)
Cry0	-0.266***
	(0.000)
Cry1	-2.039***
	(0.000)
Cry2	-3.097***
	(0.000)
Observations	56,870
Log likelihood	-14362
Rho-squared	0.215

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot.

7.4.4 Modelling DCE_{TTO}

The results from the conditional logit model can be seen in Table 10-5. All attribute coefficients had the expected sign (negative) except Annoy1, which was positive. Similarly, all values in the initial model were significant ($p=\leq 0.05$), with the exception of Annoy1 ($p=0.820$). Whilst Annoy2 was significant ($p=0.000$), it is possible that participants did not have a significantly different preference between Annoy1 and Annoy0.

Logical consistency

It is important that the utility decrement is larger for level 2 coefficients compared to level 1 coefficients within their respective dimension. The results showed that the coefficient for Annoy1 was slightly positive. Although this was not statistically significant, it should be negative in order to be logically consistent. Chi squared tests were undertaken to determine whether levels 1 and 2 of each attribute are the same. All were found to be significantly different, including for Annoy ($\chi^2=58.23$; $p=0.000$). As such, levels 0 and 1 of Annoy were merged. The fully consistent model can also be seen in Table 7-12.

On reviewing the anchored coefficients for the fully consistent model, the largest utility decrement can be seen for Hurt2, suggesting this item has the greatest relative importance. Conversely, Eat2 had the smallest utility decrement, suggesting this item has the least relative importance. Amongst the level 1 coefficients Hurt1 also has the largest utility decrement, demonstrating the large relative impact on utility from the Hurt dimension.

Table 7-12: Regression results and anchored utility decrements for the standard model and the fully consistent model using data from the DCE$_{TTO}$ survey for CARIES-QC-U

Variables	Standard model	Anchored model	Fully consistent model	Anchored fully consistent model
Hurt1_LY	-0.373***	-0.173	-0.374***	-0.173
	(0.000)		(0.000)	
Hurt2_LY	-1.217***	-0.564	-1.217***	-0.562
	(0.000)		(0.000)	
Annoy1_LY	0.009	0.004	-	
	(0.820)		-	
Annoy2_LY	-0.262***	-0.121	-0.266***	-0.123
	(0.000)		(0.000)	
Awake1_LY	-0.209***	-0.097	-0.210***	-0.097
	(0.000)		(0.000)	
Awake2_LY	-0.634***	-0.293	-0.634***	-0.293
	(0.000)		(0.000)	
Eat1_LY	-0.126***	-0.058	-0.126***	-0.058
	(0.000)		(0.000)	
Eat2_LY	-0.354***	-0.164	-0.355***	-0.164
	(0.000)		(0.000)	
Cry1_LY	-0.215***	-0.099	-0.215***	-0.099
	(0.000)		(0.000)	
Cry2_LY	-0.565***	-0.262	-0.565***	-0.261
	(0.000)		(0.000)	
LY	2.160***	-	2.166***	-
	(0.000)		(0.000)	
Observations	13,086		13,086	
Log likelihood	-3468		-3468	
Rho-squared	0.235		0.235	

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

Linearity of life years variable

The assumption that the duration attribute was linear was confirmed through a test of linearity where the duration variables were entered into the regression as dummy variables. As shown in the plot in Figure 7-6, the life years coefficients for the dummy variables can be seen to form a straight line. The coefficients can be found in appendix Z.

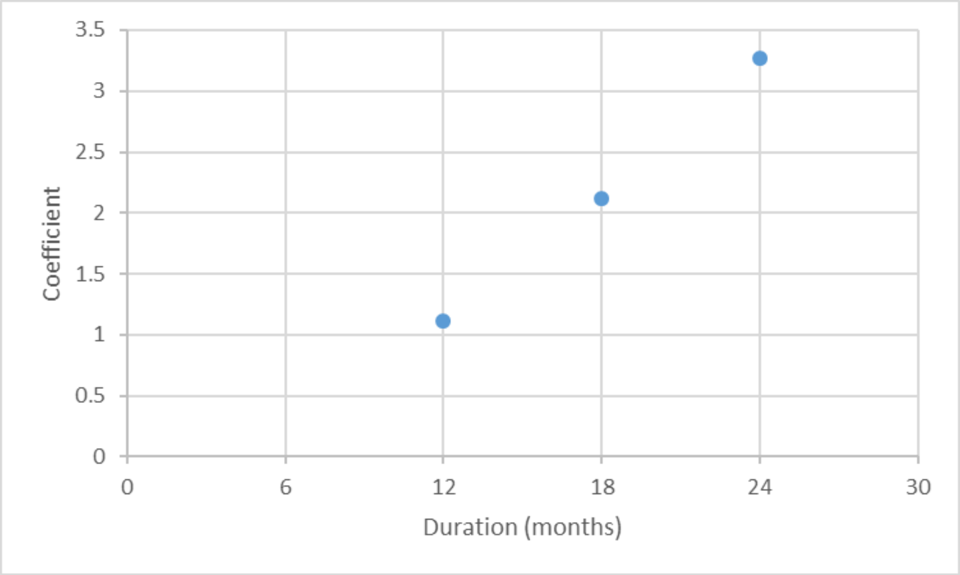


Figure 7-6: Plot demonstrating linearity of the duration attribute in the DCE$_{TTO}$ tasks

Heterogeneity of coefficients

The inclusion of interaction terms were explored to determine whether gender, age employment status, marital status, general health, the presence of existing dental problems and previous caries experience had an impact on the preferences provided. Table 7-13 shows the impacts of these sociodemographic and health-related interaction terms. Positive attributes were identified for almost all dimensions for participants with self-reported current dental problems, demonstrating that these participants gave higher values than those without current dental problems. Conversely, a number of negative interactions were seen for participants over the age of 65 years, suggesting older participants gave lower values for these dimension levels. The coefficients relating to these interaction effects can be found in appendix AA.

Table 7-13: The impact of sociodemographic and health-related interaction terms in the DCE_{TTO} survey results for CARIES-QC-U

	Male participants	Self-reported current dental problems	Previous caries experience	Employed participants	Married participants	Participants with bad or very bad general health	Participants aged 65 and over	Participants aged 18-24
Hurt1_LY		+			+	-		
Hurt2_LY	+	+			+	-	-	
Annoy1_LY		+						
Annoy2_LY		+					-	
Awake1_LY								
Awake2_LY	+	+						
Eat1_LY		+						
Eat2_LY		+		+				-
Cry1_LY		+					-	
Cry2_LY		+		+			-	

Notes: + a positive, statistically significant ($p \leq 0.05$) coefficient – a negative, statistically significant ($p \leq 0.05$) coefficient. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

Model robustness

Regressions were estimated for seven further models, each with a reduced sample. These excluded participants that failed the dominance test, reported difficulty understanding the tasks or difficulty in choosing an answer within the tasks, and combinations of these. These also explored the robustness of the standard model to exclusion of participants that completed the survey very quickly, or slowly. The regressions can be found in appendix BB. The models were seen to produce minimal changes in the regressions estimated, suggesting the baseline model was robust, with the exception of Annoy1 in which the coefficient changed from being positive to negative. Whilst this beneficial change was observed in most of the additional models (robustness models 1 to 5), it was not significant. Interestingly this change in the sign of the Annoy1 coefficient was not seen when participants that completed the survey very quickly or very slowly were excluded (robustness models 6 and 7).

7.4.5 Anchoring adolescent BWS values onto the QALY scale

The estimated regressions mapping the BWS values onto DCE_{TTO} values, and the models exploring the inclusion of squared (quadratic) and cubic terms, can be seen in

Table 7-14:

Table 7-14. The absolute error between the observed and predicted values was determined and the mean of these for each model are shown in the same table. A large proportion of the predictions were greater than 0.05 or 0.1 from the observed values, suggesting a relatively large error between the two. Nonetheless, out of the three models, the quadratic model can be seen to have the fewest predictions with error greater than 0.05 or 0.1.

The constants shown in

Table 7-14 are greater than 1. This is due to the presence of predictions that were greater than 1. Utilities cannot be greater than 1, so these were capped at 1 and the errors recalculated. Three predictions were capped in the linear model, whilst the quadratic and cubed models each had only one prediction above 1 that was capped. This had no effect on the absolute error for the any of the models but slightly reduced the number of predictions >0.05 and >0.1 from the observed DCE_{TTO} for the linear model (180 and 123 respectively).

Table 7-14: Mapping models used to generate health state utility values using adolescent BWS preferences for CARIES-QC-U

	Anchored DCE_{TTO} utilities		
Variables	**Linear**	**Squared**	**Cubed**
Modelled BWS value linear	0.119***	1.047*	0.079
	(0.000)	(0.070)	(0.471)
Modelled BWS value squared	-	-0.004***	-0.000
	-	(0.006)	(0.992)
Modelled BWS value cubed	-	-	0.000
	-	-	(0.763)
Constant	1.503***	1.194***	1.280***
	(0.000)	(0.000)	(0.000)
Mean absolute difference	0.113	0.110	0.110
Number of predictions >0.05 from observed DCE_{TTO}	181	177	180
Number of predictions >0.1 from observed DCE_{TTO}	124	114	115
Observations	243	243	243
R-squared	0.788	0.795	0.795

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1

The plot shown in Figure 7-7 illustrates how closely the estimated utilities compare to the observed values for the regression model, whilst Figure 7-8 and Figure 7-9 show the plots for the quadratic and cubic models respectively. All three models tend to underestimate at the higher end of the scale and overestimate at the lower end of the scale, though the quadratic model appears to produce utility values that most closely follow the pattern of observed DCE_{TTO} values. When considering the amount of error within the models alongside the plots, the quadratic model would appear to be preferable. The mapped estimates from the quadratic model shown in appendix CC can be used directly to score CARIES-QC-U health states in cost-utility analyses, where it should be noted that the state with a prediction above 1 has been capped at 1.

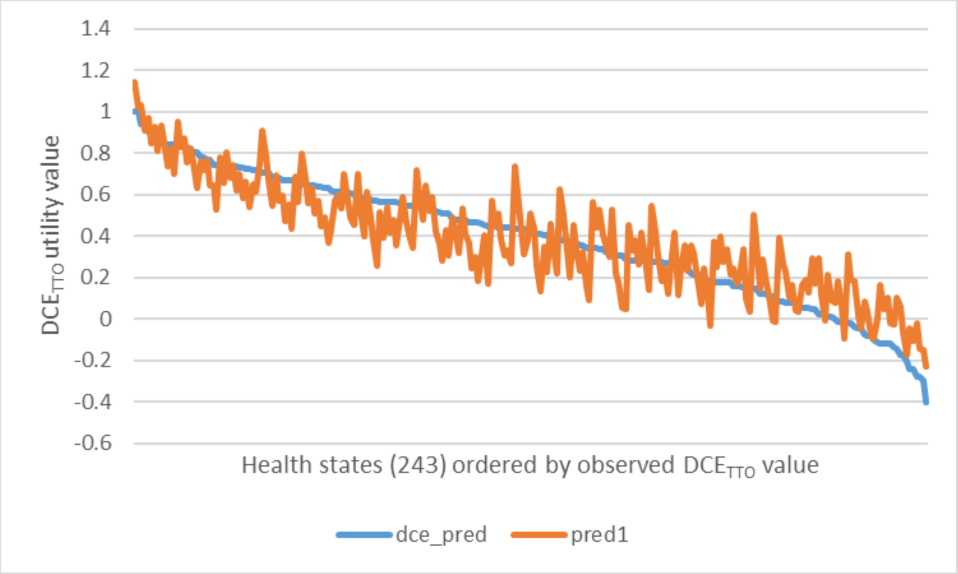


Figure 7-7: Plot showing the relationship between the estimated and observed values for the linear model

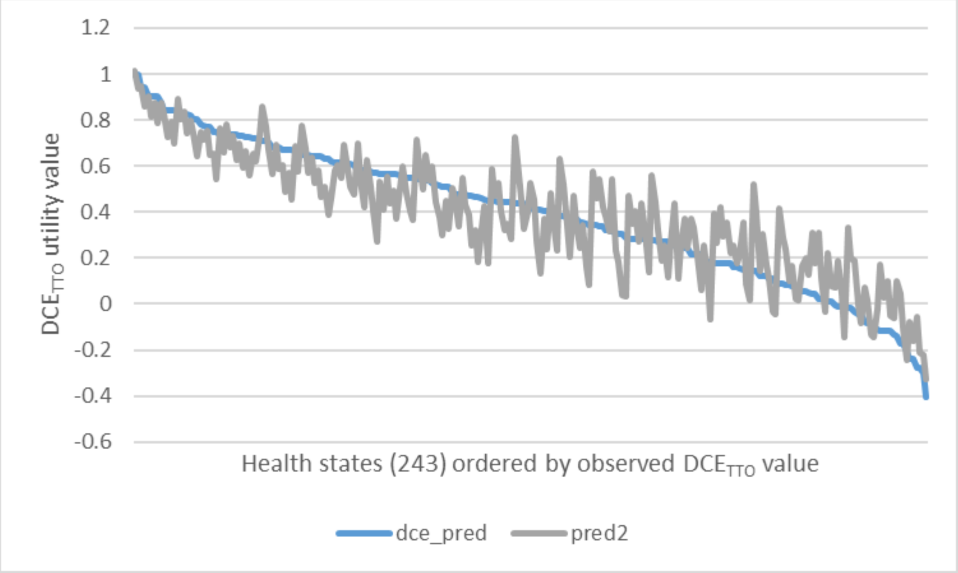


Figure 7-8: Plot showing the relationship between the estimated and observed values using the quadratic model

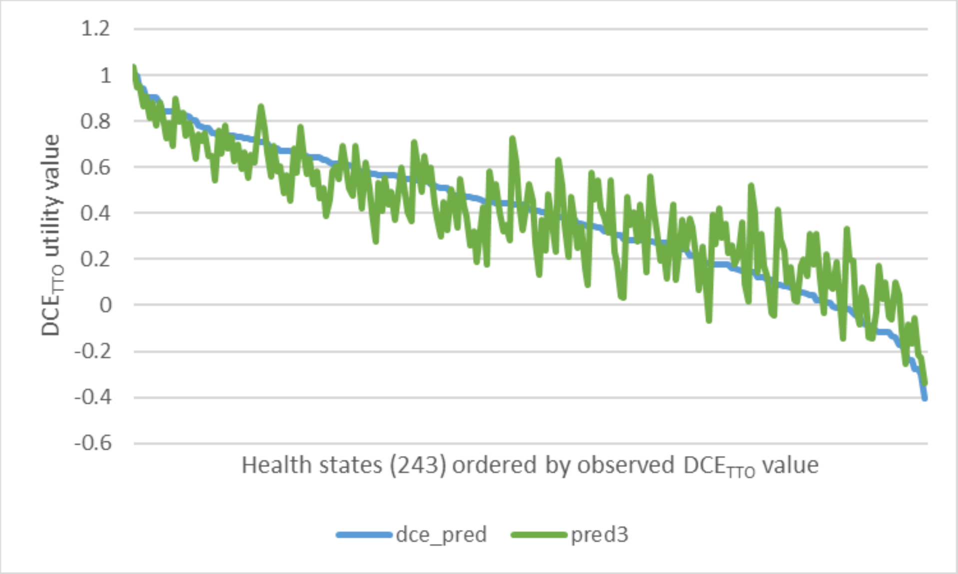


Figure 7-9: Plot showing the relationship between the estimated and observed values using the cubic model

7.5 Discussion

This chapter has described the valuation of CARIES-QC-U; a child-centred preference-based measure specific to caries. The section below firstly considers the main findings from this stage of the study in greater detail, discussing how they compare with or refute findings from other bodies of work in this field. Consideration will then be given to the socio-demographic and health-related characteristics of participants alongside the diversity and representiveness of the sample, before reflecting on ethical issues and recruitment. Strengths and acknowledged limitations of the research will be presented, including an appraisal of the study design used, before finally making recommendations for future research priorities.

7.5.1 Reflections on key findings

BWS coefficients

The results suggest that adolescents felt *'hurt'* to be the most important attribute in CARIES-QC-U. The majority of participants tended to place more weight on the attribute relating to their teeth hurting a lot (Hurt2). This dimension level was found to have the lowest coefficient in this sample. Conversely, participants valued the no impairment level of this attribute (Hurt0) most highly.

The second lowest coefficient related to participants crying about their teeth a lot (Cry2). The importance of this attribute, which was tentatively categorised as an emotional impact in section 5.7.4, is not surprising, particularly in the context of an adolescent population. The adolescent valuation of CHU9D found that adolescents placed far greater importance on what the authors termed 'mental health attributes' (a dimension comprising the attributes *'worried'*, *'sad'* and *'annoyed'*) than adults (Ratcliffe et al., 2012b, Ratcliffe et al., 2016a, Ratcliffe et al., 2016b). In line with the CHU9D findings, the present study found that adults placed greater emphasis on the physical impacts of caries, primarily the *'hurt'* dimension.

Previous BWS studies have used approaches to scale the coefficients, to allow the PITS state (in this case, the state with the lowest OHRQoL specific to caries: 22222) to represent 0 and the state with no impacts (00000) to be placed at 1 (Coast et al., 2008a). This was not considered necessary for the present study due to the use of mapping techniques instead. Nonetheless, this approach was explored by subtracting one fifth of the index value of state 22222 from all attributes and then dividing by the index value for state 00000, in line with the method described by Coast and colleagues. The results obtained using this scaling method, however, appeared implausible, with some coefficients having an incorrect sign, hence they were not utilised for any further analyses.

DCE$_{TTO}$ utilities

The utility decrement from the adult survey for a lot of dental pain (Hurt2) was particularly large; more so than would be expected. A disutility of 0.56 is similar to what other studies have reported for health states in considerably more severe, systemic and life-threatening conditions such as cancer (Lloyd et al., 2006, Paracha et al., 2018).

Toothache is known to produce a particularly debilitating pain, though of course there is a degree of subjectivity to the measurement of pain (Renton, 2011). Yet despite the acknowledged severity of the pain, toothache related to caries is likely to be temporary. Even in the absence of treatment, the inflamed pulpal tissue will ultimately die due to the prolonged bacterial insult, and the symptoms will subside (Renton, 2011). This may provide an explanation for the notably high utility decrement related to pain for adult participants, as described below.

Almost 80% of the adult sample reported receipt of treatment for dental caries; either a filling or an extraction. Almost half of the adult participants reported some degree of current problems relating to their dentition. As such, it is highly likely that a substantial proportion of adult participants had previously experienced toothache, or were experiencing it at the time they completed the survey. For these participants, the thought of experiencing the same severity of pain for the durations stated in the tasks was understandably likely to be considered extremely unpleasant.

Concern relating to the length of these durations first arose following the soft launch of the survey, whereby the very large utility decrement for Hurt2 was initially noted. Through adopting shorter durations for the main survey, in line with recommendations from adult PPI representatives, the utility decrement was slightly reduced, though not to the extent that had been anticipated.

Whilst the shorter durations were felt to be more appropriate for a condition such as dental caries, it is important to acknowledge that these are not clinically realistic. For example, it would not be feasible for severe dental pain related to dental caries to last for even the shortest duration of six months. The use of much shorter durations, such as a week, may be more clinically realistic, but when combined with the statement regarding death (‘and then you will die’) the reduced life span would almost certainly influence the results.

Logical consistency

Whilst the adolescent BWS preferences produced a logically consistent model, the adult DCE_{TTO} values did not. The intermediate level for the *‘annoy’* dimension was not logically consistent, suggesting that adult participants did not have a preference between Annoy1 and Annoy0. The process of collapsing levels, so that two levels are merged and presented as a single combined level, is not novel; this approach was used in the valuation of the SF-6D and the adolescent valuation of the CHU9D, amongst many other measures. It is considered to be an accurate and reliable approach to dealing with inconsistencies such as these.

It is important to consider why these two levels (Annoy0 and Annoy1) were logically consistent and distinct in the adolescent survey results, but not in the adult survey. The earlier stages of this study (chapter 5) discussed the views of PPI representatives regarding the classification system, whereby it was apparent that the terms *'annoy'* and *'hurt'* were considered to be very different concepts by children and young people, though adults had concerns that the terms were very similar. Furthermore, children had a tendency to view *'annoy'* as a physical impact, similar to a niggling sensation, rather than akin to feeling irritated or frustrated as an adult might interpret it. This differing interpretation of the term may have led to the discrepancies in the valuation of this attribute.

Dominance test, difficulty and understanding

There is no clear convention for the use of a dominance test in a BWS valuation survey, which may explain why one was not included for the adolescent valuation of CHU9D led by Ratcliffe and coworkers (2016a). The present study did include a dominance test, which doubled as a practice question, and reported both the findings for adolescents that were able to correctly identify the 'best' attribute, as well as those that were able to correctly identify both the 'best' and 'worst' attributes. It was unsurprising that fewer adolescents were able to perform the latter. This acknowledges that, whilst there can be little dispute over which attribute is 'best' (since having no problems must be better than having problems), there is scope for personal judgement as to which attribute is the worst, based upon the participants' own views and experiences. Further discussion on this topic can be found in section 6.6.1.

The overwhelming majority of adult participants were able to pass the DCE_{TTO} dominance test. It should be noted that the findings from the previous stage of this study suggested that adolescents may have managed to pass the DCE_{TTO} dominance test simply by selecting the option that most closely represented their own dental impacts or experiences. This approach often led them to choose the better alternative. It is possible that adults may also respond in this way; an absence of evidence to demonstrate this effect is not proof that this does not occur.

For the DCE_{TTO} in particular, a strong argument has been made to exclude participants who failed a dominance test, on the basis that they may not be making rational decisions, or may not understand how to complete the tasks (Devlin et al., 2003). Nonetheless, concerns have been raised that this approach could risk the exclusion of valid preferences (Lancsar and Louviere, 2006). A cautious approach was taken in the present study, by exploring the robustness of the results to the exclusion of those who failed the test or reported difficulties of choice or understanding, so that the impact on the utilities could be assessed. For the DCE_{TTO} data, the exclusion of these participants (robustness models 1 to 5 in appendix BB) did produce a negative Annoy1 coefficient, which can be considered advantageous in comparison to the standard model, though these were still not significant. As such, the present study considered the standard models for both the DCE_{TTO} and the BWS surveys to be robust to the inclusion of these participants.

There was no dominance test included for the soft launch of the adult survey, though this was included for the rest of the sample in order to comply with convention. Unfortunately, this meant that there were no data available regarding the pilot participants' ability to pass the dominance test.

Additionally, the dominance tests for both surveys had a dual function as practice questions. The implications of this are discussed further in 7.5.5.

Survey completion times

Further to the exploration of robustness in relation to the dominance test, difficulty of understanding and difficulty of choice, the impact of the length of time taken to complete the surveys was also determined. Previous research has excluded participants that completed the survey very quickly on the basis that it would not be possible to complete the survey in such a short time whilst applying the necessary cognitive processes to undertake the tasks correctly (Bansback et al., 2012). Similarly, participants who have taken a very long time to complete a survey have been excluded, as it has been assumed that they either did not understand, or did not give the survey their full attention.

On viewing the histograms, the researchers agreed that participants completing the survey in under three minutes and over 30 minutes could be considered as outliers. The adult and adolescent coefficients were re-estimated with the exclusion of these outliers. The magnitude of the coefficients for the adolescent BWS data were seen to differ minimally when compared to the standard model and the sign of the coefficients remained negative throughout, suggesting that the standard model was robust to the inclusion of these participants. Interestingly, the exclusion of these participants was the only robustness test that did not produce a negative coefficient for Annoy1 for the $\mathrm{DCE}_{\mathrm{TTO}}$ data, suggesting that the impact of this was not as substantial as the exclusion of participants for other reasons.

Further to this, there was not a strong argument to demonstrate how long it should take a participant to complete the surveys in the present study, and hence it could not be determined whether this approach would result in the exclusion of the correct participants. Moreover, through the use of a more inclusive approach, the preferences of participants with reading difficulties or the need for frequent breaks can be encompassed.

Heterogeneity

The model exploring heterogeneity in the responses from participants with bad or very bad general health show a number of anomalies, though these can be explained due to the very low number of observations in this model. With the exception of this model, there was little heterogeneity in the BWS data.

For the adult $\mathrm{DCE}_{\mathrm{TTO}}$ data, significant interactions were identified relating to heterogeneity in the values from participants with self-reported current dental problems. Participants with existing dental problems provided higher values, indicating that they did not perceive the health states to be as bad as their counterparts without dental problems did. This fits with the existing literature regarding patient values in healthcare, which suggests that those with a condition tend to value it more highly, implying that they have adapted to the impacts of their condition (Brazier et al., 2005). Whilst a meta-analysis concluded that there were no differences in the values from patients and those from the wider population, there were apparent differences when reviewing the individual studies (Dolders et al., 2006). This has led to calls for patient values to be adopted in place of those from the wider population (Versteegh and Brouwer, 2016). Whilst there remains a resistance to the use of patient values, as caries is such a prevalent condition, any attempts to avoid patient values entirely would be futile since any general population sample will contain a proportion of people with experience of the condition.

Another area of heterogeneity in the $\mathrm{DCE}_{\mathrm{TTO}}$ data relates to participant age. Significant negative interactions were noted for older participants, suggesting they tended to provide lower values. There are a range of reasons why there could be variations in the values provided by participants of different ages. Older people may place greater emphasis on duration than younger adults, or may view any loss of health more severely due to any existing impairment (Dolan et al., 1996). The literature reports variations in valuations provided by adults of different ages, notably a sharp decrease in valuations provided by those over 70 years of age (Dolan and Roberts, 2002).

Other researchers have considered the influence of participant age on the values provided. Ratcliffe and colleagues used the preferences of young adults, rather than all ages of adults, to rescale ordinal adolescent values (Ratcliffe et al., 2015, Ratcliffe et al., 2016a). The use of a young adult sample was considered to provide the most similar values to adolescents, given that it was not possible to obtain cardinal values directly from the latter group. However, to the authors' knowledge, there is no evidence to support this view. Furthermore, it should be noted that adolescence is a very unique period of life, and any attempts to obtain similar values from a different age group are unlikely to succeed.

Anchoring adolescent BWS preferences onto the QALY scale

A substantial amount of error was seen between the observed and predicted DCE_{TTO} values when anchoring the BWS preferences onto the 1-0 scale using mapping, with differences of 0.1 utilities for over half of the health states. This is likely to be explained by the fact that these values were obtained from two different populations, using two different preference-elicitation tasks.

To address the fact that some of the predictions were over 1, the highest possible value at full health, the utilities were capped at 1 for the small number of states where this occurred. An alternative way to address the predictions over 1 would have been to estimate further models, such as a tobit model (Mukuria et al., 2019). Nonetheless, there was only one prediction over 1 in the selected quadratic model for the present study, plus there has been some debate in the literature as to whether the tobit model offers much improvement (Mukuria et al., 2019).

Whilst the addition of squared and cubic terms has not been shown to improve model performance in other research, the quadratic model did appear to be more favourable in the present study when reviewing the mean absolute error and the model plots (Rowen et al., 2015). All three plots demonstrated a degree of underprediction at the higher end of the scale and overprediction at the lower end of the scale, which has been seen in other mapping studies, though notably with different types of data (Rowen et al., 2009).

7.5.2 Participants

Diversity of samples

Valuation surveys do not usually collect postcode data, and tend to focus on establishing the socio-economic status of individual participants rather than the level of deprivation of the area in which they reside. Typically, the questions required to determine socio-economic status are not applicable for adolescents and hence would not have been appropriate to include in an adolescent survey. In contrast, most adolescents are aware of their postcode. Furthermore, caries research frequently collects postcode data to allow further exploration of the aforementioned relationship between caries and deprivation. For the purposes of aligning the two surveys as much as possible, postcode data was requested from adolescent and adult participants. This allowed the participants location within the UK to be determined, alongside the deprivation level of the area in which they resided.

Both surveys had representation from each of the devolved nations, indicating a good geographical spread of participants. Nonetheless, there were slightly fewer adult and adolescent participants from Scotland than would have been anticipated in line with the population norms derived from UK Census data (ONS, 2011). For future surveys of this type, there may be some benefit to the setting of quotas for recruitment from each of the devolved nations to ensure the sample is fully representative in this way.

It was not possible to locate a number of participants from each survey from their postcode using the Indices of Multiple Deprivation for each devolved nation (NISRA, 2017, Stats Wales, 2019, Ministry of Housing, 2019, Scottish Government, 2020). There are a number of possible reasons for this. Firstly, it may be due to a simple typing error made when the participant entered the postcode. Whilst adolescent participants were given the option to state that they did not know their postcode, this response was not afforded to adult participants. It is possible that a proportion of adult participants were unaware of their postcode, but were pressed to enter one in order to proceed with the survey, which in turn may have led to errors. Moreover, whilst the Scottish Index of Multiple Deprivation had been published very recently, the indices from the other devolved nations were less current (NISRA, 2017, Stats Wales, 2019, Ministry of Housing, 2019, Scottish Government, 2020). Any new housing areas that had been built since publication may not have been included in the indices. This may provide a further possible explanation for the unrecognised postcodes.

Whilst the previous stage reported levels of deprivation by quintiles, deciles were the only standardised format used by the Indices of Multiple Deprivation for each devolved nation and hence were used to report the findings in this stage (NISRA, 2017, Stats Wales, 2019, Ministry of Housing, 2019, Scottish Government, 2020). In both surveys, almost half of the participants resided in more deprived areas of the UK.

Each survey had good representation from across the respective age ranges. Notably, there was a smaller proportion of 16-year-old participants in the BWS survey than in the general UK population, although the minimum quotas that were set during recruitment had been met. Recruiting from this age group is understandably difficult, as young people gain more independence and have increasing school pressures.

The results suggest that the samples in the present study were not as ethnically diverse as the population norms, though this is likely to be due to the use of different ethnicity categories for data collection than those used in the UK Census. The categories used within the survey are identical to those used within the National Health Service (NHS), whilst the Census uses fewer categories, with no separate category for mixed or multiple ethnic groups. Future surveys of this type would be advised to use categories that align with the UK Census for ease of analysis and interpretation.

The adult survey enquired into participants’ main activity, and found that the proportion of participants who described their main activity as either retired or doing housework were substantially higher than the population norms. This was anticipated as these participants may be more likely to spend more time at home, and hence may have more opportunity to undertake surveys than those who are employed. Interestingly, the sample also comprised a greater proportion of married participants and a smaller proportion of single participants than is reflected in the wider population.

The population norms from the most recent UK Census were discussed above to try to establish how representative the study samples were. Nonetheless, these figures were obtained approximately nine years ago, and hence the population is likely to have changed in the period of time between the Census and the commencement of these surveys.

It is not possible to report a response rate for these surveys. As recruitment was undertaken via Surveyengine (SurveyEngine GmbH, Berlin, Germany), using market research platforms, the research team were not given access to any details of individuals who started to view the survey but declined to consent to participate. Furthermore, the survey was passively advertised via multiple survey platforms so it would not be possible to identify how many potential participants viewed the advert but did not click on the link to the survey.

Health-related characteristics of samples

The self-reported general health of the adolescent sample was substantially greater than that of the adult sample. This is understandable as many conditions affecting health are more prevalent with increasing age.

Regarding dental health, almost half the adult sample reported their teeth to be a problem for them. Interestingly, however, the adult participants' responses to the warm-up questions from CARIES-QC did not fully reflect this, with the exception of their responses to the questions relating to feeling annoyed about their teeth, and having difficulty eating some foods. As CARIES-QC was not designed for use with adult participants, the items within it may not be relatable to this population. Furthermore, a high proportion of adults reported impacts from their teeth, suggesting that perhaps they were not thinking about the impacts from their teeth 'today', and were instead recalling previous experiences. Similarly, the impacts that they reported may not have been due to caries.

Whilst the reasons for adults self-reported dental problems may not have been specifically related to caries, almost 80% of the sample reported previously having a filling or a tooth removed due to decay. This is a reflection of the prevalence of dental caries within the population. Although the findings are over a decade old now, the results of the most recent Adult Dental Health Survey identified visible caries in almost a third of a representative sample of adults in England, Wales and Northern Ireland, noting participants had an average of just 17.9 sound teeth (White et al., 2012).

Similarly, over a third of adolescent participants reported current impacts from their teeth in their responses to the CARIES-QC questions, though these impacts may not have necessarily have been related to caries. It is also possible that participants were not thinking about their teeth ‘today’ when they completed these questions. Furthermore, almost half of adolescents reported having a previous filling or extraction due to dental decay. This fits with our current knowledge regarding the prevalence of caries in this population, whereby almost a third of 12-year-olds and half of 15-year-olds had obvious decay experience in a representative sample of England, Wales and Northern Ireland (Pitts et al., 2015).

Both adolescent and adult participants were provided with binary response options (yes or no) to the question regarding their previous experience of a filling or dental extraction, with no option to state that they were unsure. This may have been an oversight, in that anecdotally many individuals are unclear regarding the treatment they have received, or the reasons why a certain intervention was necessary. Furthermore, there may have been an element of self-selection bias surrounding the very high proportions of adults and adolescents reporting previous fillings or extractions; an individual may be more likely to volunteer to take part in a survey if they are interested in, or have had experience of the topic in question.

7.5.3 Ethical and governance considerations

Parental consent

In line with the feedback received from the NHS Research Ethics Committee for the previous stage of this study (chapter 6), it was clear that parental consent would be required once again for the present stage. As recruitment was undertaken by Surveyengine (SurveyEngine GmbH, Berlin, Germany), the requirement to obtain parental consent was not considered to adversely affect this process. Nonetheless, it does fail to acknowledge the ability of young people to make decisions in this regard.

NHS ethical approval

Although the samples for this stage were identified from the general population, the NHS Research Ethics Committee were keen to review the participant-facing materials for this part of the study and provide the necessary approvals. This was not originally anticipated, as studies such as this would not usually be considered to fall within the remit of this committee, though the research team were grateful for their feedback. Two further substantial amendments were submitted for this stage, to seek approval for the use of the surveys and the confirmed sample size.

Participant complaint

During the survey completion phase, the lead researcher (HJR) received a complaint from an adult participant. The email explained the participant's distress at the content of the DCE_{TTO} tasks, namely the wording of the duration attribute that culminated in 'and then you will die'. The participant felt unable to complete the rest of the survey and requested a response from the study team. This raises an important point with wider implications.

Researchers in this field are aware of the complexities surrounding discussing death with participants. This is one of the primary reasons why cardinal tasks are generally not considered to be suitable for use with children and adolescents. Ratcliffe and colleagues (2011) reported encountering difficulties obtaining ethical approval to use a TTO with young adolescents, and were required to make modifications to the task to exclude any reference to death; a modification that would render the TTO (and also the DCE_{TTO}) defunct for the purpose of facilitating the anchoring of ordinal values. The authors suggested that older adolescents may exhibit a stronger ability to tolerate and accept the concept of immediate death as part of a health state valuation exercise when compared to younger adolescents. Yet interestingly, there is no clear evidence to demonstrate that adults can tolerate and accept the concept of immediate death in these tasks, and perhaps more importantly, there is no evidence that these tasks do not cause adult participants distress. Nonetheless, valuation studies incorporating TTO and SG tasks with adults will routinely manage to obtain ethical approvals. Clearly, the inclusion of the concept of death within these tasks caused this particular participant distress to the point that they felt the need to raise their concern. However, it is entirely possible that other participants experienced similar distress but did not communicate this to the researchers.

Further to this, it should be noted that this participant did not complete the survey as a result of their distress. It is possible that this also played a role in other participants' decision to not complete the survey. There is a need for further research into the psychological and emotional impacts that participants may experience when completing these tasks, and the extent to which these impacts may influence their decision on whether or not to complete the survey.

Fortunately the careful and timely management of this particular complaint led to a positive outcome. The lead researcher (HJR) initially responded with a holding email to reassure the participant that their concerns were being taken seriously and that they would be sent a formal response as soon as possible. A more thorough response was sent subsequently, explaining the reasons why the questions were phrased in that particular way, as well as signposting the participant on where to access further support if required. The participant was grateful that their response had been addressed and also provided a suggestion on how the wording could be phrased in a more positive way; 'the rest of my life would be ___ years, all of them lived with this condition'. The participant provided consent for their suggestion to be published to raise awareness amongst other researchers in this field. The anonymised email exchange can be seen in appendix 0.

7.5.4 Strengths

Novel

Whilst this is not the first study to use BWS to access the preferences of adolescents, it is the first study, to the authors' knowledge, to anchor adolescent preferences using adult DCE_{TTO} values. The involvement of children and young people during the development of the survey ensured the study remained child-centred, whilst the views of adult representatives played a particularly important role in the determination of durations for the main DCE_{TTO} survey.

Experimental design

One strength of this study was the use of a full factorial design for the adolescent BWS survey. This meant that preferences were obtained directly from participants for every health state defined by the classification system. This approach was possible primarily due to the compactness of the CARIES-QC-U classification system, with only five dimensions and three levels. An identical approach was used to obtain adolescent preferences for the generic EQ-5D-Y instrument (Dalziel et al., 2020). Whilst a full factorial design was not possible for the DCE_{TTO}, efforts were made to optimise efficiency and balance the design (Risa Hole, 2015).

Importantly, for both the BWS and DCE_{TTO} surveys, not all participants completed a task that involved the PITS state (22222: Hurt2, Annoy2, Awake2, Eat2, Cry2). Whilst this could be considered a limitation of the study, it should be acknowledged that the choice sets were randomly allocated, and hence it is likely that all participants will have had the opportunity to value some severe CARIES-QC-U health states.

Accessibility of surveys

The colour scheme and font for both surveys was chosen in accordance with national guidance to aid participants with specific learning difficulties and visual impairments (UK Association for Accessible Formats, 2012). This was intended to improve accessibility for these individuals.

7.5.5 Limitations

Plausibility of health states

In section 7.5.1, the implications of the length of duration in the DCE_{TTO} tasks were discussed, noting that the possibility of having severe dental pain for over six months would be implausible. Further to this, it should be acknowledged that the combination of descriptors that formed the health states for both the DCE_{TTO} and BWS tasks could also be considered unrealistic. For example, a scenario whereby an individual's teeth do not hurt them at all, but their teeth keep them awake a lot would be unlikely to occur. Typically the pain itself would be the reason that a child is unable to sleep, though this could occur for other reasons, such as worry or discomfort. Whilst the scenarios are intended to be hypothetical, it is possible that some health states may have been implausible. Whilst this is a limitation of this PBM, it is a limitation of most PBMs where there are implausible combinations for some dimensions. However, this implausibility was not taken into account during health state selection and hence this could be considered a potential limitation of the study design.

Dual function of practice question

The practice question for each valuation survey had a dual function in providing the dominance test. The primary issue with this approach is that the participants undertook the dominance test before they had been given the opportunity to familiarise themselves with the task and response format. This may have led more participants to fail the dominance test than if it had been randomised amongst the main tasks.

A further concern is that participants were made aware that they were completing a practice question. This may have led them to believe that it was less important than the main questions, and hence potentially put less consideration into their response.

Lastly, participants were able to request a second opportunity to undertake the practice question. This led to difficulties in analysing the responses, in that a participant may have initially failed the dominance test, but passed the second time, or vice versa.

Choice of preference-elicitation tasks

The use of a BWS task to access the preferences of adolescents has been clearly justified in the previous chapter, though it is important to acknowledge the limitations of using this approach and the concerns of other researchers in the field. One concern that has been raised is that the profile BWS task can technically only produce values and not preferences, as the latter requires one aspect to be traded for another (Coast et al., 2008a, Krucien et al., 2017). As the participant is asked to decide from attributes *within* one health state, rather than *between* profiles, the premise of opportunity cost is not met (Krucien et al., 2017).

Further concerns relate to the level of concordance between BWS and DCE tasks. Whilst one study has reported similar patterns in social care preferences elicited with BWS and DCE tasks, and no significant differences in the values produced after rescaling, contrasting findings have also been reported (Potoglou et al., 2011, Krucien et al., 2017). Krucien and colleagues (2017) reported low levels of concordance between the two methods, as well as lower stability, monotonicity, and continuity for BWS data. The authors also stated that these findings suggested that the BWS tasks produced lower quality data.

Lastly, the choice to use a DCE_{TTO} task to elicit preferences from adults is also not without limitation. It has been suggested that the DCE_{TTO} task produces values that are lower than the TTO, though it is acknowledged that more research is needed in this area (Bansback et al., 2012, Rowen et al., 2018a).

Selection bias

As the samples for these surveys were identified from a survey platform, the participants are likely to regularly complete surveys such as this and may have developed skills and expertise in this process. Despite efforts to identify nationally representative samples, a degree of selection bias will exist, whereby these participants may not reflect the views of the wider population who do not regularly engage with surveys of this type.

Both surveys were restricted for completion on computer devices only, rather than mobile phones in an effort to maximise participant concentration during the tasks. There was an intention to open up both surveys for completion on mobile phones should there have been any difficulties in recruitment, though this was ultimately not necessary. Nonetheless, despite these restrictions, six participants (0.96%) did manage to undertake the survey on their mobile phones. This is thought to have occurred due to these participants using more advanced mobile phones that functioned in a similar way to a tablet computer. Overall, however, participation was limited to those who had access to a computer, which may have excluded individuals with low socio-economic status.

Parental Influence

Following receipt of parental consent, the survey indicated that the adolescent should take over for the rest of the questions. Nonetheless, there was no mechanism in place to prevent parents completing the survey with their child. As such, the extent of parental influence over the adolescent’s responses remains unclear. Whilst there is no available evidence in the literature describing the influence of parents on adolescent responses to a valuation survey, a study by Granville-Garcia and colleagues investigated this effect on young children completing a self-reported measure of OHRQoL (Granville-Garcia et al., 2016). The researchers reported no significant differences between the responses provided by five-year-old children to the SOHO-5 (Scale of Oral Health Outcomes for Five-Year-Old Children) when alone or when accompanied by a parent (Granville-Garcia et al., 2016). Nonetheless, these participants were much younger than those in the present study, hence the influence of parental presence or involvement in this population requires further investigation.

7.5.6 Further research priorities

Short term priorities

It is important for the valid concerns raised by the aforementioned participant in this study to be acknowledged and thoroughly investigated. The use of terminology surrounding the concept of death has long been a part of preference-elicitation tasks such as TTO and SG. These tasks are not considered appropriate for use with children and adolescents and previous studies have been denied ethical approval to use these tasks in their original form with younger populations (Ratcliffe et al., 2011). Nonetheless, there exists an assumption that these tasks are appropriate for adult participants to complete, and that the inclusion of terms related to death would not cause them harm. A qualitative approach to determine the impacts of the use of these tasks in adults should be prioritised, with efforts to develop alternative wording that could minimise these impacts whilst limiting any detrimental effects on the preference-elicitation processes.

Before the new UK adolescent value set for CARIES-QC-U can be widely used, it requires validation. It would seem prudent to apply the algorithm to existing CARIES-QC datasets, though excluding the original dataset that was used to identify the classification system (Gilchrist et al., 2018). This is discussed further in section 8.3.

Longer-term goals

In the longer term, it would be useful to compare the utilities determined using the adolescent value set for CARIES-QC-U with other generic instruments, such as the new adolescent scoring of EQ-5D-Y, perhaps alongside a clinical trial of a community-based oral health improvement intervention (Dalziel et al., 2020).

As CARIES-QC is used more widely in studies across the world, the opportunities to derive utilities for these study participants also increase (Arrow and Forrest, 2020). It would be necessary to validate and potentially adapt the classification system for the relevant population, using the same processes described in chapter 5. A value set can then be obtained from the population in the same way as detailed in the present chapter, allowing direct comparisons to be made between utilities for participants from each country.

Whilst the adolescent value set has been the primary focus of this body of work, it is important to acknowledge that this study has also generated an adult value set for CARIES-QC-U, and that there may be circumstances where values obtained from a population that can vote and contribute to taxes are preferable to those from adolescents.

7.6 Conclusion

This chapter has described the final stage in the development of a child-centred caries-specific PBM. This is the first utility measure specifically designed for application in dentistry, using BWS techniques to access the preferences of adolescents. The adolescent and adult value sets produced are able to provide a utility for every health state defined by the CARIES-QC-U classification system. After the validation of CARIES-QC-U, there are a wealth of potential applications for its use; some of which are explored in the following chapter, alongside the wider implications of using QALYs in dentistry.

8 Discussion

This thesis has presented a body of work comprising four studies, with each component informing the next, culminating in the development of CARIES-QC-U, a child-centred, caries-specific preference-based measure (PBM) suitable for use in economic evaluation of interventions and programmes to improve child oral health. As the key findings, ethical and governance issues, strengths, limitations and areas for future research have been discussed for each individual stage, this chapter will have a different focus. An initial summary of each stage will be provided, followed by a consideration of the wider issues relating to the generation of QALYs in caries research and the use of an adolescent value set. Finally the research and policy-level implications of CARIES-QC-U will be considered.

8.1 Overall summary

First, a systematic review was undertaken to explore the scope, methodological and reporting quality, and involvement of children in existing economic evaluations of child oral health interventions (chapter 3). This found that whilst the majority of economic evaluations were related to dental caries and prevention, most of these were cost-effectiveness studies. These had used a wide range of outcome measures, which would make it difficult to draw comparisons between interventions. Only one study reported outcomes in QALYs, which were generated using CHU9D; an instrument which has been shown to lack sensitivity to changing caries status. There was almost no involvement of children as active participants, plus a range of issues were observed regarding the methodology and reporting of the studies that were reviewed. This clarified the need for the development of an instrument that could generate QALYs, to acknowledge the impact of caries on OHRQoL, and facilitate the comparison of outcomes across interventions. It also highlighted the importance of developing such a tool with involvement of children, to ensure its relevance and suitability, and ensuring that children could complete the measure themselves, to engage them as active participants in economic evaluations.

The second step identified CARIES-QC as the basis of the future measure, predominantly due to its specificity for the condition and population in question, its satisfactory psychometric properties, and the involvement of children during its development. This stage focused on the selection of attributes from the original 12-item three-level CARIES-QC to produce a short version that could be used as a classification system for a PBM. The selection process identified five attributes (*'hurt'*, *'annoy'*, *'kept awake'*, *'carefully'* and *'cried'*) that performed well in Rasch analysis and classical psychometric testing, and received favourable views from child and adult study representatives, as well as the team that originally developed CARIES-QC. This preliminary classification system was validated with children and young people who had a current diagnosis of caries, using a qualitative, think aloud approach. Children thought the items selected were appropriate and covered a range of aspects of the condition, but felt an alternative item (*'hard to eat'*) would better encompass the impacts related to eating compared to the item originally selected (*'carefully'*). The classification system was modified accordingly and approved by the Young People Panel before it was considered valid for use in a valuation survey.

The next stage centred on determining the most appropriate task to use to elicit preferences from adolescents, as well as establishing the age range that were able to undertake these tasks independently, and the number of tasks they could complete. Adolescents at a local secondary school completed a survey comprising a series of BWS and DCE tasks based upon the aforementioned classification system. Think aloud and semi-structured interviews with adolescents at a local secondary school as they undertook the survey revealed a preference for, and greater comprehension of BWS tasks compared to DCE. Participants had a tendency to select a response for the DCE based upon how their own mouth felt. As such, participants were able to pass the dominance test for this task by selecting the option with the least impacts, similar to their own mouth, but in doing so they did not demonstrate understanding. No relationship was identified with age, with adolescents aged 11 to 16 years demonstrating an ability to undertake the BWS tasks independently, and participants felt they could manage 8-10 tasks comfortably.

These findings were used to inform the final stage of this study; a BWS valuation study to obtain preferences from adolescents for every health state defined by the aforementioned classification system. A concurrent DCE_{TTO} survey was completed by adult participants, with the aim of generating cardinal values that could be used to anchor the ordinal values from adolescents on the 1-0 full health-dead scale required to generate QALYs. The scaled BWS model and anchored DCE_{TTO} model demonstrated robustness and little heterogeneity, though it was necessary to collapse one level of the *'annoy'* attribute to produce a fully consistent model from the DCE_{TTO} data. The mapping approach was used to anchor the adolescent BWS values onto the 1-0 QALY scale, with the quadratic model displaying predictions closest to the observed DCE_{TTO} values.

8.1.1 Novel features of CARIES-QC-U

CARIES-QC-U is original in a number of ways. Primarily, this is the first condition-specific preference-based measure in the field of dentistry as a whole, and certainly the first in the much narrower field of child oral health. Health economics in dentistry has remained an area little explored until recently, whereby there appears to be growing interest in the literature both in the UK and worldwide. As CARIES-QC is increasingly used as a patient-reported outcome measure (PROM) in the UK, and undergoing translation and validation elsewhere, the opportunity for researchers to obtain concurrent utilities may prompt more economic evaluations to be conducted in this field. Moreover, the generation of QALYs will enable the data to be easily compared across interventions, allowing meta-analyses to be conducted with ease.

This is also the first PBM to use DCE_{TTO} to anchor BWS values. Previous PBMs have either not anchored the values onto the QALY scale, or have used values obtained using TTO tasks for anchoring purposes (Coast et al., 2008a, Ratcliffe et al., 2011). Further research is required to determine the potential differences in utilities that the DCE_{TTO} produces, when compared to the TTO (Bansback et al., 2012, Rowen et al., 2018a).

The development of this PBM is also truly interdisciplinary, contributing to the evidence base in three fields; health economics, paediatric dentistry and dental public health. Yet, whilst the originality of CARIES-QC-U offers much to the evidence base in these fields, it also has a number of wider implications that require consideration, as discussed below.

8.2 Generating QALYs

Whilst this thesis has presented the advantages of QALYs and the benefits of their use in dentistry, it is important to acknowledge that there have also been many concerns highlighted regarding the QALY system (Kahneman, 2008, Pettitt et al., 2016). A recent review highlighted a range of criticisms of the QALY system, categorising these into ethical concerns, methodological issues and theoretical assumptions, and context- or disease-specific considerations (Pettitt et al., 2016). These can be seen below in Table 8-1.

Table 8-1: A critique of QALYs, adapted from Pettitt et al., 2016

Category	Limitations
Ethical Limitations	Valuing life - valuation of one individual's life over another's Determining personhood - to measure quality of life, life must be present (debate arises in brain-dead patients, foetuses etc) Potentially reduces freedom of choice May set false limits on healthcare (e.g. restrict budgets) Overly utilitarian - all QALYs are considered equal regardless of individual or situational circumstances
Methodological issues and theoretical assumptions	Variation in underlying measurement methodology / technique Validity and reliability of measurements concerning utility values of health status Perceptive and evaluative differences in population assigning utility values (e.g. physicians vs. patients) Failure of utility scores to account for contextual factors (e.g. severity of initial health state, disease prevalence) Employment of league table comparisons (comparing heterogeneous population and time periods) Discrimination towards therapies requiring high initial investments but offering long-term benefits (e.g. regenerative medicines) Failure to acknowledge non-health related benefits (e.g. societal, return to work)
Condition-specific or contextual considerations	Failure to reflect needs of those with rare conditions Bias against disabled patients who may not be able to achieve substantially higher quality of life Issues with trade-offs between quality and quantity of life, particularly in acute/transient conditions Insufficient sensitivity to measure small but meaningful changes in health status

One key limitation relates to its overly utilitarian approach, in that all QALYs are considered equal, regardless of individual or situational circumstances (Pettitt et al., 2016). Furthermore, whilst the emphasis of QALYs is on quality of life (or OHRQoL in this case), they can fail to consider other factors relating to wellbeing, and typically exclude any process attributes. Importantly, QALYs are not used to inform policy-making in some countries, particularly those with primarily private healthcare systems.

8.2.1 Generating QALYs with CARIES-QC-U

Whilst the implications surrounding the use of condition-specific PBMs (CSPBMs) as opposed to generic PBMs (GPBMs) have already been discussed in section 2.5.1, it is important to understand how these relate to CARIES-QC-U.

As mentioned previously, there is an ongoing debate as to whether QALYs generated through use of a CSPBM can be considered equal to those generated using a GPBM (Cookson et al., 2009, Versteegh et al., 2012). The use of CSPBM can be indicated in situations where the use of a GPBM would be considered inappropriate (Versteegh et al., 2012, NICE, 2013, Rowen et al., 2017). As highlighted previously, the application of a paediatric GPBM (CHU9D) to child oral health research has been reported to have a lack of sensitivity to changes in caries status in trials conducted in New Zealand and Wales (Foster Page et al., 2015, Chestnutt et al., 2017). As such, the use of a caries-specific paediatric PBM, such as CARIES-QC-U, is justified.

Whilst CARIES-QC-U may appear to have a very narrow remit for a CSPBM, it is important to remember that caries is the most prevalent chronic disease to affect children globally, and also the most common reason for a child to have a general anaesthetic in this country (Kassebaum et al., 2015, NHS Digital, 2019).

One key concern regarding CSPBMs is that the attributes they contain may not be relevant for all of those with the condition (Rowen et al., 2017). In the case of CARIES-QC-U, the attributes, or items, were identified by children and young people with current experience of the condition in question (see chapter 5). The use of an appropriate population to identify these attributes offers a particular benefit to CARIES-QC-U, as the experiences of children and young people with caries may differ to those of adults, and may be expressed using alternative terms. Nonetheless, it must be acknowledged the sample involved in identifying these attributes for inclusion in the CARIES-QC-U classification system were recruited from a secondary care setting. As such, it was likely that the referral of these participants had been prompted by an experience of pain or related symptoms. This alone does not reflect the experiences of many children, whom we know from both the literature and clinical experience, may not have any symptoms from their caries (Milsom et al., 2002a, Tickle et al., 2002). This leads to an important question regarding the generation of QALYs in caries research and the wider field of oral health in general.

Dental caries is a chronic disease unlike any other. An individual may have multiple teeth with caries at different stages, hence they could experience a range of differing impacts from their teeth at any one time. An individual with more advanced disease does not necessarily experience more severe impacts, and notably, a tooth with extensive decay may die (pulpal necrosis) silently, without pain or infection. Those that do experience pain resulting from dental caries often find it is transient; severe toothache rarely lasts for more than a few days (Renton, 2011). Due to the severity of the pain, treatment is often sought quickly. Whilst the treatment provided may not always be appropriate to address the pain, such as prescription of antibiotics, curative options are available, such as removal of the irreversibly inflamed pulpal tissue (as part of a root canal treatment) or extraction of the tooth (Cope et al., 2016).

In light of these factors, there may be a number of difficulties in generating QALYs in caries research. For example, if an intervention has an effect on only one tooth, such as a type of restoration, it would not be possible to determine whether the impacts self-reported by a child using CARIES-QC-U related to that particular tooth, or another tooth in the mouth. As with CARIES-QC and other non-PBMs, children and young people may self-complete CARIES-QC-U thinking about other oral health problems that they may be experiencing, such as inflamed gums or ulcers (Krisdapong et al., 2012). Similarly, if an individual does not experience any symptoms from their carious teeth, there is unlikely to be a notable gain in QALYs noted following provision of an intervention to address the caries.

It is for these reasons, that CARIES-QC-U may be more applicable to the evaluation of upstream oral health promotion programmes, rather than individual or tooth-level interventions. Its use to evaluate population or community-based interventions may allow for differences in QALYs to be identified. Similarly, it would appear logical to expect QALYs to reflect changes in oral health status following the provision of emergency dental care, whereby the patient can undergo treatment to relieve their pain.

A related issue pertains to the durations used to obtain adult values for the DCE_{TTO} survey. As discussed in chapter 7, the use of durations of six months, a year, a year and a half and two years was considered preferable to longer, more conventional durations for surveys of this type. Whilst there is no real requirement for these durations to be clinically realistic, one cannot ignore the fact that toothache would not last for this length of time, nor would death be considered a potential consequence. The implications of these durations on the values provided by a participant, particularly an individual with previous experience of severe toothache, may have been substantial. The incorporation of a duration attribute detailing the lifespan remaining in a certain health state before death, was necessary to enable QALYs to be determined, though it does cast some doubt over the suitability of standard preference elicitation tasks for use to elicit preferences to generate QALYs in this field.

QALYs based upon adolescent preferences

Whilst CARIES-QC-U has both an adult and an adolescent value set, the emphasis of this thesis has been on the adolescent scoring system. This is due to the acknowledged need to actively involve children and young people in research and healthcare decisions (United Nations, 2009, Marshman et al., 2015).

There are a number of advantages to the use of an adolescent scoring set. Adolescents tend to have less engagement with professional treatment programmes, with poorer treatment compliance, which is thought to be perpetuated by exclusion from decisions relating to their own health, and healthcare in general (Dolgin et al., 1986, Goldston et al., 1995, Shaw, 2001, Cavet and Sloper, 2004). Incorporating the preferences of young people into healthcare research and subsequent health service development can result in programmes and interventions designed to be more relevant to their needs, which may result in improved service utilisation (Tylee et al., 2007).

Despite the numerous advantages to the generation of QALYs through use of the adolescent scoring system, it is also important to acknowledge that this may not be appropriate in all circumstances. For example, some studies may have a requirement for the use of preferences from the taxpaying and voting population, in which case the adult value set generated by modelling the adult DCE_{TTO} data could be applied.

As detailed in chapter 6, an adolescent population was selected to provide values for CARIES-QC-U, based upon the results of the qualitative school-based study that determined this age group were able to understand and complete best-worst scaling tasks independently, using a computer-based format. Yet, whilst the adolescent population may have been able to understand the tasks, the preferences that they provided may not reflect those of younger age groups, or even those of others within their age group. This is because adolescents is a unique period of development, characterised by behaviours that tend to differ from other aspects of childhood, such as increased risk-taking, with substantial variation between individuals (Dolgin et al., 1986). The adolescent value set can be applied to the responses provided by much younger children, since CARIES-QC can be used with children aged as young as five years old, though it is important to question how applicable the values of adolescents are to this age group. Nonetheless, there may not be a more suitable alternative, since younger children are not cognitively able to complete current preference elicitation tasks independently (Stevens, 2015).

To summarise, CARIES-QC-U is likely to have greater face and content validity for children with caries than GPBMs, in light of the extensive involvement of children and young people at every stage. The use of the adolescent value set may also offer wider benefits to healthcare research and services. Nonetheless, there are wider issues surrounding the use of QALYs in caries research that warrant further investigation.

8.3 Implication of CARIES-QC-U for research

As discussed in section 7.5.6, there is a necessity to validate CARIES-QC-U. One potential approach to determine the validity of the CARIES-QC-U would be to follow a similar model used to assess the validity of the generic paediatric PBM, CHU9D (Ratcliffe et al., 2012a). This would include an assessment of completion rates, and an assessment of its ability to discriminate between groups with different characteristics, particularly in relation to their oral health and previous caries experience.

An ideal dataset upon which to base the validation of CARIES-QC-U would be that arising from the ongoing BRIGHT trial (Marshman et al., 2019); a randomised controlled trial aiming to investigate the use of a short messaging service (SMS) behaviour change intervention to improve the oral health of young people aged 11 to 13 years living in deprived areas. The use of both CARIES-QC (to which the CARIES-QC-U value set can be applied) and CHU9D in this trial provides an ideal opportunity to evaluate and compare the empirical estimates of oral health- and health-related utility derived from these measures respectively (Marshman et al., 2019). Whilst this will allow any observed differences to be explored, it is important to acknowledge that the comparison of oral health-related utility with health-related utility is not analogous, and the value sets obtained for these two measures have been obtained from different populations using different techniques (Stevens, 2012). Nonetheless, there is currently no alternative comparator that is able to provide oral health-related utility values, given that CARIES-QC-U is the first of its kind. As further paediatric PBMs specifically designed to generate oral health-related utilities emerge, comparison with CARIES-QC-U in a similar way would be warranted.

Whilst the BRIGHT trial is still underway, the CARIES-QC-U value set could be applied to a recently completed study, investigating the OHRQoL of children and young people following provision of dental treatment under general anaesthetic, which also used the CARIES-QC measure (Knapp, 2019).

Following validation and application of the CARIES-QC-U algorithm to historical data collected using CARIES-QC, there is a role of the new PBM to be used in future trials. As the COVID-19 pandemic has accelerated the move towards minimally invasive dentistry and preventive care, there is an increasing array of new interventions and preventive programmes that will require systematic evaluation before adoption into practice. It is an expectation of funding applications for Health Technology Assessments (HTA) to have an economic evaluation planned alongside the clinical study. There have been notable difficulties in identifying a suitable tool to conduct cost-utility analyses for HTA in child oral health research (Tickle M, 2016, Chestnutt et al., 2017, Maguire et al., 2020). One common feature of many recent HTA studies in this field is that the most pivotal findings were actually economic in nature, rather than clinical (Tickle M, 2016, Chestnutt et al., 2017, Maguire et al., 2020). The incorporation of a utility measure specifically designed for the population and condition to be investigated could offer a number of advantages to future HTA studies in this area, such as ease of comparability of outcomes due to the generation of QALYS, and facilitation of involvement of children.

8.3.1 Operationalisation of CARIES-QC-U

The time taken for caries to develop or progress is highly variable, and dependent on a number of factors, including individual caries risk status. Similarly, the time required for the benefits of an intervention to be identified will be heavily dependent on these factors, as well as the nature of the intervention. Typically, trials investigating the caries preventive effect of an intervention will continue for a minimum of three years, with outcome assessments at baseline and three years (Tickle et al., 2011), with some studies adding additional assessment at 12 and 24 months (Chestnutt et al., 2012, Marshman et al., 2019). Others have adopted a pragmatic approach through incorporating outcome assessment at participants' scheduled or emergency dental visits between baseline and final assessment (Innes et al., 2013a). These studies tend to include a range of outcome measures, often including clinical assessment alongside parent- or patient-reported outcomes related to OHRQoL, which have been undertaken in clinical and non-clinical settings (Innes et al., 2013a, Marshman et al., 2019). In line with these studies, the authors would advocate the use of CARIES-QC-U as an outcome measure to be used at similar time points, and in a range of research settings, to allow the estimation of QALYs.

8.4 Implications of CARIES-QC-U for policy and commissioning

As discussed in chapter 2, the budget for any healthcare system must have a limit, and hence it must be determined which interventions or healthcare programmes to fund and allocate resources to, in order to maximise the benefit to patients (Mitton and Donaldson, 2009). Often these decisions can be difficult, as there are insufficient resources available to provide all potential services (Detels et al., 2015).

In the UK, decisions regarding the funding of healthcare interventions are typically made by the National Institute for Health and Care Excellence (NICE), whilst commissioning of services and allocation of resources is undertaken by NHS England and local clinical commissioning groups. NICE provide a reference case, outlining the methods that they consider to be most appropriate for analysis when developing technology appraisal guidance (NICE, 2013). This advises the conduct of cost utility analyses for adults using EQ-5D, a GPBM that has a value set elicited using TTO with a representative sample from the general adult population, though permits the use of a CSPBM where a GPBM is evidenced as unsuitable, and CSPBMs can be used in sensitivity analyses. For children, the NICE guidance is less prescriptive as there is no recommended single preference-based measure.

Although NICE play a leading role in decision-making in medical care, they have little involvement in oral healthcare. Further to this, commissioning in dentistry is somewhat different, whereby the dental budget is devolved to each region, yet no formal agreement exists between NHS England and these regions as to who holds the power and responsibility for strategic decisions (Vernazza et al., 2019). Since 2012, oral health promotion programmes have been commissioned by local authorities, with these decisions being based upon guidance from Public Health England. The document Commissioning Better Oral Health for Children and Young People advises local authorities to commission specific oral health programmes based upon the evidence base and the needs of the population, addressing the underlying causes of oral health inequalities, whilst putting children, young people and their families at the heart of commissioning (Public Health England, 2014a).

This guidance provides little formal advice on the financial considerations related to commissioning these programmes, though there is known to be little available evidence to inform these decisions (Public Health England, 2014a, Lord et al., 2015a). This has been highlighted by NICE, who have recommended that future research focuses on improving evidence on the cost-effectiveness of oral health promotion interventions (NICE, 2015).

Whilst QALYs are not directly required in the same vein that NICE would require for other areas of healthcare, nor is there a funding threshold set by NHS England related to QALY gains, the generation of QALYs in child oral health research could be of particular use in addressing this paucity. The use of CARIES-QC-U may provide more accurate data for the evidence base, which in turn can be used to inform commissioners regarding which specific approaches to caries prevention and management are funded. Furthermore, the use of an appropriate utility measure specifically designed for the relevant population and condition, such as CARIES-QC-U may also help to improve the quality of economic evaluations in this field. Interestingly, the use of CARIES-QC-U would be consistent with a number of the requirements of the aforementioned NICE reference case, namely through generating

QALYs, and the elicitation of values using a representative sample derived from the UK general population rather than patients.

Whilst CARIES-QC-U would require validation and potential adaptation before its use in other populations, baseline data has already been collected using CARIES-QC in a number of countries, which would facilitate the first steps in this process (Schuller et al., 2018, Foster Page et al., 2019, He and Wang, 2020). The generation of QALYS would satisfy the requests of numerous decision-makers around the world, including those in New Zealand and Australia.

8.5 Impact Statement

The studies outlined in this thesis describe the child-centred development of the first preference-based measure for children with caries; CARIES-QC-U. This measure can be used in economic evaluations of interventions to improve children's oral health, facilitating the generation of QALYs. In turn, this will contribute to the evidence base, allowing the most cost-effective programmes to be identified, supporting patients to receive the most effective intervention, whilst enabling the most efficient use of resources.

This chapter has summarised the key findings arising from this body of research, considering the key features of CARIES-QC-U and its ability to generate QALYs. Whilst further research is required to determine the validity and sensitivity of CARIES-QC-U and the suitability of generating QALYs in child oral health research, this chapter has demonstrated how use of this utility measure could address the need for high quality economic evaluations in this field, to inform commissioners of interventions to improve children's oral health. The next and final chapter of this thesis will consider whether the aims and objectives for this study have been addressed, and synthesise the key messages to take away from this study.

9 Conclusion

This thesis has presented a series of four interlinked studies, which together aimed to develop a PBM of OHRQoL specific to dental caries, based on the preferences of children and young people.

The following specific objectives were set in order to meet this aim:

1. To identify the quality and scope of published economic evaluations in child oral health research and the measures of benefit currently used
2. To develop and validate a classification system for child dental caries, based upon the CARIES-QC caries-specific measure of OHRQoL, that is amenable to health state valuation
3. To determine what age range of adolescents can complete ordinal health state valuation tasks, whether they prefer DCE or BWS and at what age they can use a computerised format independently
4. To undertake a population-based valuation survey with children and young people to determine their preferences
5. To model their preferences to produce a valuation algorithm that provides preference weights for each health state defined by the classification system

These objectives have been fulfilled, culminating in the development of CARIES-QC-U, a child-centred, PBM of OHRQoL specific to caries. Furthermore, additional knowledge has been generated surrounding the current quality and scope of economic evaluations in this field, the abilities of children and young people to complete ordinal valuation tasks and the practical challenges when using DCE_{TTO} and BWS tasks for oral health preference elicitation.

9.1 Key messages

- Few high quality economic evaluations have been conducted in the field of child oral health research, with a notable lack of cost-utility analyses, and little involvement of children and young people as active participants.

- The classification system for CARIES-QC-U was identified using a novel approach, combining Rasch analysis, classical psychometric testing, the views of the team that developed CARIES-QC, and the views of children, young people and parents. This approach can be applied to the identification of a classification system from measures of HRQoL, or to the development of short-form versions of measures.

- The use of qualitative techniques to validate a classification system is justifiable and feasible.

- Adolescents aged 11-16 years are able to understand and complete between 8 and 10 BWS tasks independently using a computerised format. The sample in the present study expressed a preference for, and greater understanding of BWS tasks, compared to DCE tasks.

- The responses to questions from CARIES-QC provided by adolescents regarding their OHRQoL 'today' were influenced by their past experiences, such as previous toothache, wobbly teeth and orthodontic treatment.

- Preference elicitation using BWS tasks with adolescents is both practical and feasible. These preferences can be modelled using the conditional logit, and mapped onto the QALY scale using adult values obtained using DCE_{TTO} tasks. The quadratic mapping model demonstrated the best fit between the observed and predicted values in the present study.

- Adult DCE_{TTO} values appeared to be strongly influenced by the length of the duration attribute. Although shorter durations were used in the present study, these still grossly overestimate the length of time an individual would typically experience impacts from dental caries; a condition characterised by symptoms of an acute and transient nature.

- The involvement of children and young people at every stage of development affords CARIES-QC-U the advantage of being pertinent to the population it is applied to.
- Children and young people have been involved at every stage of the development of CARIES-QC-U. This has enabled their views to be incorporated into the design of the measure, ensuring its relevance and suitability for the population in question.
- CARIES-QC-U enables utility values to be generated for all datasets that have collected CARIES-QC data, thus enabling the use of existing evidence to inform cost-utility analyses.

9.2 Recommendations for future research using CARIES-QC-U

- The CARIES-QC-U instrument should be validated through application of the recommended value set to recent data collected using CARIES-QC.
- The CARIES-QC-U measure should be incorporated into future studies to facilitate cost-utility analyses of interventions to improve children's oral health, addressing the paucity of high quality economic evaluations in this field.
- Future studies should compare outcomes generated using CARIES-QC-U to those identified using generic paediatric PBMs, such as the CHU9D and EQ-5D-Y. As other child oral health PBMs are developed, these will also require comparison to CARIES-QC-U.
- The CARIES-QC-U classification system should be validated and potentially adapted for use in other populations.

In conclusion, the development of a child-centred PBM of OHRQoL specific to dental caries will provide an appropriate instrument to facilitate the cost-utility analysis of interventions to improve children's oral health. Ultimately, evidence obtained using CARIES-QC-U can contribute to guideline development and be used to inform commissioning in this field.

10 References

ABANTO, J., PAIVA, S. M., RAGGIO, D. P., CELIBERTI, P., ALDRIGUI, J. M. & BÖNECKER, M. 2012. The impact of dental caries and trauma in children on family quality of life. *Community Dentistry and Oral Epidemiology,* 40**,** 323-331.

ABANTO, J., PAIVA, S. M., SHEIHAM, A., TSAKOS, G., MENDES, F. M., CORDESCHI, T., VIDIGAL, E. A. & BÖNECKER, M. 2016. Changes in preschool children's OHRQoL after treatment of dental caries: responsiveness of the B-ECOHIS. *International Journal of Paediatric Dentistry,* 26**,** 259-265.

ABED, R., BERNABE, E. & SABBAH, W. 2019. Family Impacts of Severe Dental Caries among Children in the United Kingdom. *International journal of environmental research and public health,* 17.

ACS, G., SHULMAN, R., NG, M. W. & CHUSSID, S. 1999. The effect of dental rehabilitation on the body weight of children with early childhood caries. *Pediatr Dent,* 21**,** 109-13.

ADAMS, J., TYRRELL, R., ADAMSON, A. & WHITE, M. 2012. Effect of Restrictions on Television Food Advertising to Children on Exposure to Advertisements for 'Less Healthy' Foods: Repeat Cross-Sectional Study. *PLoS One,* 7**,** e31578.

AHOVUO-SALORANTA, A., FORSS, H., WALSH, T., NORDBLAD, A., MAKELA, M. & WORTHINGTON, H. V. 2017. Pit and fissure sealants for preventing dental decay in permanent teeth. *Cochrane Database Syst Rev,* 7**,** Cd001830.

ALBADER, H. 2019. Changes in Oral Health Related Quality of Life in Children Following Dental Extractions Under a General Anaesthetic Using Two Child Self-Report Measures. University of Leeds.

ALKARIMI, H., WATT, R., PIKHART, H., SHEIHAM, A. & TSAKOS, G. 2014. Dental Caries and Growth in School-Age Children. *Pediatrics,* 133**,** e616.

ALKHADRA, T. 2004. Cost -effectiveness of a pit and fissure sealants program in a school-based setting in Saudi Arabia. *In:* GILLESPIE, K. (ed.). ProQuest Dissertations Publishing.

ALLAREDDY, V., NALLIAH, R. P., HAQUE, M., JOHNSON, H., RAMPA, S. B. & LEE, M. K. 2014. Hospital-based emergency department visits with dental conditions among children in the United States: nationwide epidemiological data. *Pediatric dentistry,* 36**,** 393-9.

ALSHARIF, A. T., KRUGER, E. & TENNANT, M. 2015. Dental hospitalization trends in Western Australian children under the age of 15 years: a decade of population-based study. *Int J Paediatr Dent,* 25**,** 35-42.

AMERICAN DENTAL ASSOCIATION 2001. ADA Policy on Evidence-Based Dentistry *Definition of Evidence-Based Dentistry*

AMIN, M. S., HARRISON, R. L. & WEINSTEIN, P. 2006. A qualitative look at parents' experience of their child's dental general anaesthesia. *International Journal of Paediatric Dentistry,* 16**,** 309-319.

AMMANN, P., KOLB, A., LUSSI, A. & SEEMANN, R. 2013. Influence of rubber dam on objective and subjective parameters of stress during dental treatment of children and adolescents – a randomized controlled clinical pilot study. *International Journal of Paediatric Dentistry,* 23**,** 110-115.

ANGST, F., VERRA, M. L., LEHMANN, S. & AESCHLIMANN, A. 2008. Responsiveness of five condition-specific and generic outcome assessment instruments for chronic pain. *BMC Medical Research Methodology,* 8**,** 26.

APPLETON, J. V. 1995. Analysing qualitative interview data: addressing issues of validity and reliability. *Journal of advanced nursing,* 22**,** 993-997.

ARHEIAM, A., BROWN, S. L., HIGHAM, S. M., ALBADRI, S. & HARRIS, R. V. 2016. The information filter: how dentists use diet diary information to give patients clear and simple advice. *Community Dentistry and Oral Epidemiology,* 44**,** 592-601.

ARROW, P. & FORREST, H. 2020. Atraumatic restorative treatments improve child oral health-related quality of life: A noninferiority randomized controlled trial. *Community dentistry and oral epidemiology*.

ASHLEY, P. F., ATTRILL, D. C., ELLWOOD, R. P., WORTHINGTON, H. V. & DAVIES, R. M. 1999. Toothbrushing habits and caries experience. *Caries Res,* 33**,** 401-2.

ASHLEY, P. F., WILLIAMS, C., MOLES, D. R. & PARRY, J. 2015. Sedation versus general anaesthesia for provision of dental treatment to patients younger than 18 years. *Cochrane Database of Systematic Reviews*.

ATKINS, C. Y., THOMAS, T. K., LENAKER, D., DAY, G. M., HENNESSY, T. W. & MELTZER, M. I. 2016. Cost-effectiveness of preventing dental caries and full mouth dental reconstructions among Alaska Native children in the Yukon-Kuskokwim delta region of Alaska. *Journal of Public Health Dentistry,* 76**,** 228-240.

BAINES, P. 2011. Assent for children's participation in research is incoherent and wrong. *Archives of Disease in Childhood,* 96**,** 960.

BANIHANI, A., DEERY, C., TOUMBA, J., MUNYOMBWE, T. & DUGGAL, M. 2018. The impact of dental caries and its treatment by conventional or biological approaches on the oral health-related quality of life of children and carers. *International Journal of Paediatric Dentistry,* 28**,** 266-276.

BANSAK, K., HAINMUELLER, J., HOPKINS, D. & YAMAMOTO, T. 2018. The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments. *Political Analysis,* 26**,** 112-119.

BANSBACK, N., BRAZIER, J., TSUCHIYA, A. & ANIS, A. 2012. Using a discrete choice experiment to estimate health state utility values. *J Health Econ,* 31**,** 306-18.

BARBER, S., BEKKER, H., MARTI, J., PAVITT, S., KHAMBAY, B. & MEADS, D. 2019. Development of a Discrete-Choice Experiment (DCE) to Elicit Adolescent and Parent Preferences for Hypodontia Treatment. *The Patient - Patient-Centered Outcomes Research,* 12**,** 137-148.

BDJ 2018. Sugar tax now in place. *Br. Dent. J.,* 224**,** 565-566.

BELL, S., MORGAN, A., MARSHMAN, Z. & RODD, H. 2010. Child and parental acceptance of preformed metal crowns. *European Archives of Paediatric Dentistry,* 11**,** 218-224.

BENJAMIN, R. M. 2010. Oral Health: The Silent Epidemic. *Public Health Reports,* 125**,** 158-159.

BERGSTROM, E. K., LINGSTROM, P., HAKEBERG, M., GAHNBERG, L. & SKOLD, U. M. 2016. Caries and costs: an evaluation of a school-based fluoride varnish programme for adolescents in a Swedish region. *Community Dental Health,* 33**,** 138-44.

BERINSKY, A. J., MARGOLIS, M. F. & SANCES, M. W. 2014. Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys. *American Journal of Political Science,* 58**,** 739-753.

BERKOWITZ, R. J. 2006. Mutans Streptococci: Acquisition and transmission. *Pediatric Dentistry* 28**,** 106-109.

BERTRAND, E., MALLIS, M., BUI, N. M. & REINHARZ, D. 2011. Cost-effectiveness simulation of a universal publicly funded sealants application program. *Journal of Public Health Dentistry,* 71**,** 38-45.

BHURIDEJ, P., KUTHY, R. A., FLACH, S. D., HELLER, K. E., DAWSON, D. V., KANELLIS, M. J. & DAMIANO, P. C. 2007. Four-year cost-utility analyses of sealed and nonsealed first permanent molars in Iowa Medicaid-enrolled children. *Journal of Public Health Dentistry,* 67**,** 191-8.

BIRCH, S., BRIDGMAN, C., BROCKLEHURST, P., ELLWOOD, R., GOMEZ, J., HELGESON, M., ISMAIL, A., MACEY, R., MARIOTTI, A., TWETMAN, S., PRESHAW, P. M., PRETTY, I. A. & WHELTON, H. 2015. Prevention in practice – a summary. *BMC Oral Health,* 15**,** S12-S12.

BLINKHORN, A. S. 1982. The caries experience and dietary habits of Edinburgh nursery school children. *Br Dent J,* 152**,** 227-30.

BLOM-HOFFMAN, J., LEFF, S., FRANKO, D., WEINSTEIN, E., BEAKLEY, K. & POWER, T. 2009. Consent Procedures and Participation Rates in School-Based Intervention and Prevention Research: Using a Multi-Component, Partnership-Based Approach to Recruit Participants. *A Multidisciplinary Research and Practice Journal,* 1**,** 3-15.

BOULTON, J., HASHEM, K. M., JENNER, K. H., LLOYD-WILLIAMS, F., BROMLEY, H. & CAPEWELL, S. 2016. How much sugar is hidden in drinks marketed to children? A survey of fruit juices, juice drinks and smoothies. *BMJ Open,* 6.

BRAGA, M. M., MENDES, F. M., DE BENEDETTO, M. S. & IMPARATO, J. C. 2009. Effect of silver diammine fluoride on incipient caries lesions in erupting permanent first molars: a pilot study. *J Dent Child (Chic),* 76**,** 28-33.

BRANCLAZZI, G. 2008. THE EUROPEAN ASSOCIATION FOR CHILDREN IN HOSPITAL (EACH) CHARTER. *Pediatrics,* 121**,** S97.

BRANNEN, J. 2005. Mixing Methods: The Entry of Qualitative and Quantitative Approaches into the Research Process. *International Journal of Social Research Methodology,* 8**,** 173-184.

BRAY, N., NOYES, J., HARRIS, N. & EDWARDS, R. T. 2017. Measuring the health-related quality of life of children with impaired mobility: examining correlation and agreement between children and parent proxies. *BMC research notes,* 10**,** 377-377.

BRAZIER, DEVERILL, GREEN, HARPER & BOOTH 1999. A review of the use of health status measures in economic evaluation. *Health Technology Assessment,* 3.

BRAZIER, J. 2017. *Measuring and valuing health benefits for economic evaluation,* Oxford, United Kingdom, Oxford, United Kingdom : Oxford University Press, 2017.

BRAZIER, J., AKEHURST, R., BRENNAN, A., DOLAN, P., CLAXTON, K., MCCABE, C., SCULPHER, M. & TSUCHYIA, A. 2005. Should patients have a greater role in valuing health states? *Applied Health Economics and Health Policy,* 4**,** 201-208.

BRAZIER, J., DIXON, S. & RATCLIFFE, J. 2009. The Role of Patient Preferences in Cost-Effectiveness Analysis. *PharmacoEconomics,* 27**,** 705-712.

BRAZIER, J. & LONGWORTH, L. 2011. NICE DSU Technical Support Document 8: An introduction to the measurement and valuation of health for NICE submissions. NICE Decision Support Unit.

BRAZIER, J., ROBERTS, J. & DEVERILL, M. 2002. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics,* 21**,** 271-292.

BRAZIER, J., ROWEN, D., YANG, Y. & TSUCHIYA, A. 2012a. Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. *Health Economics in Prevention and Care,* 13**,** 575-587.

BRAZIER, J. E., ROWEN, D., MAVRANEZOULI, I., TSUCHIYA, A., YOUNG, T., YANG, Y., BARKHAM, M. & IBBOTSON, R. 2012b. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health Technology Assessment,* 16**,** 1-109.

BRIGGS, A. H. & O'BRIEN, B. J. 2001. The death of cost-minimization analysis? *Health economics,* 10**,** 179-184.

BSPD 2015. Written evidence for the Select Committee Enquiry into the oral health of children in England - COH0014. London: British Society of Paediatric Dentistry.

BURGESS, J. O. & VAGHELA, P. M. 2018. Silver Diamine Fluoride: A Successful Anticarious Solution with Limits. *Advances in Dental Research,* 29**,** 131-134.

BURNARD, P. 1991. A method of analysing interview transcripts in qualitative research. *Nurse education today,* 11**,** 461-466.

BUX, S., PORRITT, J. & MARSHMAN, Z. 2019. Evaluation of Self-Help Cognitive Behavioural Therapy for Children's Dental Anxiety in General Dental Practice. *Dent J (Basel),* 7**,** 36.

CANADIAN PAEDIATRIC SOCIETY 2008. Ethical issues in health research in children. *Paediatrics & child health,* 13**,** 707-720.

CANAWAY, A., AL-JANABI, H., KINGHORN, P., BAILEY, C. & COAST, J. 2017. Development of a measure (ICECAP-Close Person Measure) through qualitative methods to capture the benefits of end-of-life care to those close to the dying for use in economic evaluation. *Palliative Medicine,* 31**,** 53-62.

CARLSSON, F. & MARTINSSON, P. 2003. Design techniques for stated preference methods in health economics. *Health Econ,* 12**,** 281-294.

CARVALHO, T. S., ABANTO, J., MENDES, F. M., RAGGIO, D. P. & BONECKER, M. 2012. Association between parental guilt and oral health problems in preschool children. *Braz Oral Res,* 26**,** 557-63.

CARVALHO, T. S., ABANTO, J., PINHEIRO, E. C. M., LUSSI, A. & BÖNECKER, M. 2018. Early childhood caries and psychological perceptions on child's oral health increase the feeling of guilt in parents: an epidemiological survey. *International Journal of Paediatric Dentistry,* 28**,** 23-32.

CASTILLO, J. L., RIVERA, S., APARICIO, T., LAZO, R., AW, T. C., MANCL, L. L. & MILGROM, P. 2011. The Short-term Effects of Diammine Silver Fluoride on Tooth Sensitivity: a Randomized Controlled Trial. *J Dent Res,* 90**,** 203-208.

CAVET, J. & SLOPER, P. 2004. The participation of children and young people in decisions about UK service development. *Child: Care, Health and Development,* 30**,** 613-621.

CHADWICK, B. & EVANS, D. 2007. Restoration of class II cavities in primary molar teeth with conventional and resin modified glass ionomer cements: a systematic review of the literature. *European Archives of Paediatric Dentistry,* 8**,** 14-21.

CHAMBERS, C. T., REID, C. J., CRAIG, K., MCGRATH, P. & FINLEY, G. 1998. Agreement between child and parent reports of pain. *Clin. J. Pain,* 14**,** 336-342.

CHATFIELD, C. & COLLINS, A. 1981. *Introduction to multivariate analysis*, CRC Press.

CHEN, G. & RATCLIFFE, J. 2015. A review of the development and application of generic multi-attribute utility instruments for paediatric populations. *Pharmacoeconomics,* 33**,** 1013-1028.

CHEN, G., XU, F., HUYNH, E., WANG, Z., STEVENS, K. & RATCLIFFE, J. 2019. Scoring the Child Health Utility 9D instrument: estimation of a Chinese child and adolescent-specific tariff. *Quality of Life Research,* 28**,** 163-176.

CHESTERS, R. K., HUNTINGTON, E., BURCHELL, C. K. & STEPHEN, K. W. 1992. Effect of oral care habits on caries in adolescents. *Caries Res,* 26**,** 299-304.

CHESTNUTT, I. G., CHADWICK, B. L., HUTCHINGS, S., PLAYLE, R., PICKLES, T., LISLES, C., KIRKBY, N., MORGAN, M. Z., HUNTER, L., HODELL, C., WITHERS, B., MURPHY, S., MORGAN-TRIMMER, S., FITZSIMMONS, D., PHILLIPS, C., NUTTALL, J. & HOOD, K. 2012. Protocol for "Seal or Varnish?" (SoV) trial: a randomised controlled trial to measure the relative cost and effectiveness of pit and fissure sealants and fluoride varnish in preventing dental decay. *BMC oral health,* 12**,** 51.

CHESTNUTT, I. G., HUTCHINGS, S., PLAYLE, R., TRIMMER, S. M., FITZSIMMONS, D., AAWAR, N., ANGEL, L., DERRICK, S., DREW, C., HODDELL, C., HOOD, K., HUMPHREYS, I., KIRBY, N., LAU, T. M. M., LISLES, C., MORGAN, M. Z., MURPHY, S., NUTTALL, J., ONISHCHENKO, K., PHILLIPS, C., PICKLES, T., SCOBLE, C., TOWNSON, J., WITHERS, B. & CHADWICK, B. L. 2017. Seal or varnish? A randomised controlled trial to determine the relative cost and effectiveness of pit and fissure sealant and fluoride varnish in preventing dental decay. *Health Technology Assessment,* 21.

CHI, D. L., VAN DER GOES, D. N. & NEY, J. P. 2014. Cost-effectiveness of pit-and-fissure sealants on primary molars in Medicaid-enrolled children. *American Journal of Public Health,* 104**,** 555-61.

CHIBINSKI, A. C., WAMBIER, L. M., FELTRIN, J., LOGUERCIO, A. D., WAMBIER, D. S. & REIS, A. 2017. Silver Diamine Fluoride Has Efficacy in Controlling Caries Progression in Primary Teeth: A Systematic Review and Meta-Analysis. *Caries Res,* 51**,** 527-541.

CHILDSMILE. 2018. *Childsmile: Professionals* [Online]. Scotland: NHS Scotland. Available: http://www.child-smile.org.uk/professionals/index.aspx [Accessed 28th August 2018].

CHIOU, C.-F., WEAVER, M. R., BELL, M. A., LEE, T. A. & KRIEGER, J. W. 2005. Development of the multi-attribute Pediatric Asthma Health Outcome Measure (PAHOM). *International Journal for Quality in Health Care,* 17**,** 23-30.

CHRISTENSEN, K., MAKRANSKY, G. & HORTON, M. 2016. Critical Values for Yens Q3: Identification of Local Dependence in the Rasch Model Using Residual Correlations. *Applied Psychological Measurement,* 41.

CLARKE, M., LOCKER, D., BERALL, G., PENCHARZ, P., KENNY, D. J. & JUDD, P. 2006. Malnourishment in a population of young children with severe early childhood caries. *Pediatric Dentistry,* 28**,** 254-259.

CLAXTON, K. 1999. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of health economics,* 18**,** 341-364.

COAST, J., FLYNN, T. N., NATARAJAN, L., SPROSTON, K., LEWIS, J., LOUVIERE, J. J. & PETERS, T. J. 2008a. Valuing the ICECAP capability index for older people. *Social Science & Medicine,* 67**,** 874-882.

COAST, J., MARLEY ANTHONY, A., LOUVIERE JORDAN, J., FLYNN TERRY, N. & PETERS TIM, J. 2008b. Rescaling quality of life values from discrete choice experiments for use as QALYs: a cautionary tale. *Population Health Metrics,* 6**,** 6.

COGHLAN, A. 2018. Ban the buzz? *New Scientist,* 237**,** 22-23.

COHEN, J. 1992. A power primer. *Psychological Bulletin,* 112**,** 155-159.

CONWAY, P., GOODRICH, K., MACHLIN, S., SASSE, B. & COHEN, J. 2011. Patient-centered care categorization of U.S. health care expenditures. *Health Services Research,* 46**,** 479-90.

COOK, R. D. & NACHTRHEIM, C. J. 1980. A Comparison of Algorithms for Constructing Exact D-Optimal Designs. *Technometrics,* 22**,** 315-324.

COOKSON, R., DRUMMOND, M. & WEATHERLY, H. 2009. Explicit incorporation of equity considerations into economic evaluation of public health interventions. *HEPL,* 4**,** 231-245.

COPE, A. L., FRANCIS, N. A., WOOD, F. & CHESTNUTT, I. G. 2016. Antibiotic prescribing in UK general dental practice: a cross-sectional study. *Community Dent Oral Epidemiol,* 44**,** 145-153.

COYNE, K., REVICKI, D., HUNT, T., COREY, R., STEWART, W., BENTKOVER, J., KURTH, H. & ABRAMS, P. 2002. Psychometric Validation of an Overactive Bladder Symptom and Health-Related Quality of Life Questionnaire: The OAB-q. *Qual Life Res,* 11**,** 563-574.

CREMEENS, J., EISER, C. & BLADES, M. 2006. Factors influencing agreement between child self-report and parent proxy-reports on the Pediatric Quality of Life Inventory 4.0 (PedsQL) generic core scales. *Health and quality of life outcomes,* 4**,** 58-58.

CRUMP, R. T., BEVERUNG, L. M., LAU, R., SIERACKI, R. & NICHOLSON, M. 2017. Reliability, Validity, and Feasibility of Direct Elicitation of Children's Preferences for Health States: A Systematic Review. *Medical Decision Making,* 37**,** 314-326.

CURZON, M. E. & CURZON, J. A. 1979. Dental caries prevalence in the Baffin Island Eskimo. *Pediatr Dent,* 1**,** 169-74.

D2S. 2018. *Designed to Smile - Information for professionals* [Online]. Available: http://www.designedtosmile.org/info-pro/ [Accessed 28th August 2018].

DAKIN, H. & WORDSWORTH, S. 2013. Cost-minimisation analysis versus cost-effectiveness analysis, revisited. *Health Economics,* 22**,** 22-34.

DALZIEL, K., CATCHPOOL, M., GARCÍA-LORENZO, B., GOROSTIZA, I., NORMAN, R. & RIVERO-ARIAS, O. 2020. Feasibility, Validity and Differences in Adolescent and Adult EQ-5D-Y Health State Valuation in Australia and Spain: An Application of Best–Worst Scaling. *PharmacoEconomics,* 38.

DAMSCHRODER, L. J., BARON, J., HERSHEY, J. C., ASCH, D. A., JEPSON, C. & UBEL, P. A. 2004. The Validity of Person Tradeoff Measurements: Randomized Trial of Computer Elicitation Versus Face-to-Face Interview. *Medical Decision Making,* 24**,** 170-180.

DAVIES, G., WORTHINGTON, H., ELLWOOD, R., BLINKHORN, A., TAYLOR, G., DAVIES, R. & CONSIDINE, J. 2003. An assessment of the cost effectiveness of a postal toothpaste programme to prevent caries among five-year-old children in the North West of England. *Community dental health* [Online], 20. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/813/CN-00459813/frame.html.

DE BEKKER-GROB, E., DONKERS, B., JONKER, M. & STOLK, E. 2015. Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *The Patient - Patient-Centered Outcomes Research,* 8**,** 373-384.

DE BEKKER-GROB, E. W., HOL, L., DONKERS, B., VAN DAM, L., HABBEMA, J. D. F., VAN LEERDAM, M. E., KUIPERS, E. J., ESSINK-BOT, M.-L. & STEYERBERG, E. W. 2010. Labeled versus Unlabeled Discrete Choice Experiments in Health Economics: An Application to Colorectal Cancer Screening. *Value in Health,* 13**,** 315-323.

DE BEKKER-GROB, E. W., RYAN, M. & GERARD, K. 2012. Discrete choice experiments in health economics: a review of the literature. *Health Economics,* 21**,** 145-172.

DEAN, A., BARK, J., SHERRIFF, A., MACPHERSON, L. & CAIRNS, A. 2011. Use of the 'Hall technique' for management of carious primary molars among Scottish general dental practitioners. *European Archives of Paediatric Dentistry,* 12**,** 159-162.

DEERY, C. 2013. Caries detection and diagnosis, sealants and management of the possibly carious fissure. *BDJ,* 214**,** 551.

DEERY, C. 2017. Clinical practice guidelines proposed the use of pit and fissure sealants to prevent and arrest noncavitated carious lesions. 48-50. ISSN 1532-3382.

DETELS, R., GULLIFORD, M., ABDOOL KARIM, Q. & TAN, C. C. 2015. *Oxford Textbook of Global Public Health*, Oxford University Press.

DEVLIN, N., HANSEN, P., KIND, P. & WILLIAMS, A. 2003. Logical inconsistencies in survey respondents' health state valuations - a methodological challenge for estimating social tariffs. *Health Econ.,* 12**,** 529-544.

DIAS, A. G. A., MAGNO, M. B., DELBEM, A. C. B., CUNHA, R. F., MAIA, L. C. & PESSAN, J. P. 2018. Clinical performance of glass ionomer cement and composite resin in Class II restorations in primary teeth: A systematic review and meta-analysis. *Journal of Dentistry,* 73**,** 1-13.

DO, L. G. 2012. Distribution of Caries in Children: Variations between and within Populations. *Journal of Dental Research,* 91**,** 536-543.

DOLAN, P., GUDEX, C., KIND, P. & WILLIAMS, A. 1996. The time trade-off method: Results from a general population study. *Health Econ,* 5**,** 141-154.

DOLAN, P. & ROBERTS, J. 2002. To what extent can we explain time trade-off values from other information about respondents? *Soc Sci Med,* 54**,** 919-929.

DOLDERS, M. G. T., ZEEGERS, M. P. A., GROOT, W. & AMENT, A. 2006. A meta-analysis demonstrates no significant differences between patient and population preferences. *Journal of Clinical Epidemiology,* 59**,** 653-664.

DOLGIN, M. J., KATZ, E. R., DOCTORS, S. R. & SIEGEL, S. E. 1986. Caregivers' perceptions of medical compliance in adolescents with cancer. *Journal of Adolescent Health Care,* 7**,** 22-27.

DRUMMOND, M., DRUMMOND, M. F. & DRUMMOND, M. F. 1997. *Methods for the economic evaluation of health care programmes,* Oxford, Oxford : Oxford University Press, 1997.

DRUMMOND, M., SCULPHER, M. J., TORRANCE, G. W., O'BRIEN, B. J. & STODDART, G. L. 2015. *Methods for the economic evaluation of health care programmes*, Oxford, United Kingdom : Oxford University Press, 2015.

DURLAK, J. A. 2009. How to Select, Calculate, and Interpret Effect Sizes. *Journal of Pediatric Psychology,* 34**,** 917-928.

ELICE, C. E. & FIELDS, H. W. 1990. Failure to thrive: review of the literature, case reports, and implications for dental treatment. *Pediatr Dent,* 12**,** 185-9.

ENGEL, G. L. 1977. The Need for a New Medical Model: A Challenge for Biomedicine. *Science,* 196**,** 129-36.

ESBENSEN, F.-A., MELDE, C., TAYLOR, T. J. & PETERSON, D. 2008. Active Parental Consent in School-Based Research: How Much Is Enough and How Do We Get It? *Evaluation Review,* 32**,** 335-362.

EUROQOL, G. 1990. EuroQol—a new facility for the measurement of health-related quality-of-life. *Health policy.,* 16**,** 199-208.

EYLES, H., MHURCHU, C., NGHIEM, N. & BLAKELY, T. 2012. Food Pricing Strategies, Population Diets, and Non-Communicable Disease: A Systematic Review of Simulation Studies. *PLos Med.,* 9.

FAYLE, S., WELBURY, R. & ROBERTS, J. 2001. British Society of Paediatric Dentistry: a policy document on management of caries in the primary dentition. *International Journal of Paediatric Dentistry,* 11**,** 153-157.

FAYLE, S. A. 1999. Stainless steel preformed crowns for primary molars. *International Journal of Paediatric Dentistry,* 9**,** 311-314.

FDI WORLD DENTAL FEDERATION 2015. The challenge of oral disease - a call for global action. *The Oral Health Atlas.* 2nd Edition ed. Geneva.

FEATHERSTONE, J. 2008. Dental caries: a dynamic disease process. *Australian Dental Journal,* 53**,** 286-291.

FEATHERSTONE, J. D. B. 1999. Prevention and reversal of dental caries: Role of low level fluoride. *Community Dentistry and Oral Epidemiology,* 27**,** 31-40.

FEENY, D., FURLONG, W., TORRANCE, G. W., GOLDSMITH, C. H., ZHU, Z., DEPAUW, S., DENTON, M. & BOYLE, M. 2002. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care,* 40**,** 113-28.

FEENY, H. D. & TORRANCE, W. G. 1989. Incorporating Utility-Based Quality-of-Life Assessment Measures in Clinical Trials: Two Examples. *Medical Care,* 27**,** S190-S204.

FERRANS, C. E. 1990. Quality of life: Conceptual issues. *Seminars in Oncology Nursing,* 6**,** 248-254.

FINKELSTEIN, J. W. 1998. Methods, models, and measures of health-related quality of life for children and adolescents. *Measuring Health-related Quality of Life in Children and Adolescents: Implications for Research and Practice. Mahwah, NJ: Lawrence Erlbaum Associates***,** 39-52.

FISHER, F. J. 1968. A field survey of dental caries, periodontal disease and enamel defects in Tristan da Cunha. *Br Dent J,* 125**,** 447-53.

FLOOD, C., SIMPSON, A., BURLS, A., PRICE, A. & CARTWRIGHT, M. 2018. What utility scores do mental health service users, healthcare professionals and members of the general public attribute to different health states? A co-produced mixed methods online survey. *PLoS One,* 13**,** e0205223.

FLYNN, T. N., LOUVIERE, J. J., PETERS, T. J. & COAST, J. 2007. Best–worst scaling: What it can do for health care research and how to do it. *Journal of Health Economics,* 26**,** 171-189.

FOSTER PAGE, L., GILCHRIST, F., BRODER, H. L., CLARK, E. & THOMSON, W. M. 2019. A Comparison of Three Child OHRQoL Measures.

FOSTER PAGE, L. A., BECKETT, D. M., CAMERON, C. M. & THOMSON, W. M. 2015. Can the Child Health Utility 9D measure be useful in oral health research? *International Journal of Paediatric Dentistry,* 25**,** 349-357.

FRAZAO, P. 2012. [Cost-effectiveness of conventional and modified supervised toothbrushing in preventing caries in permanent molars among 5-year-old children]. *Cadernos de Saude Publica,* 28**,** 281-90.

FYFFE, H. E. & KAY, E. J. 1992. Assessment of dental health state utilities. *Community dentistry and oral epidemiology,* 20**,** 269-73.

GAFFNEY, A., MCGRATH, P. J. & DICK, B. 2003. Measuring pain in children: developmental and instrument issues. *In:* SCHECHTER N.L., BERDE C.B. & M., Y. (eds.) *Pain in Infants, Children, and Adolescents.* 2 ed.: Lippincott Williams and Wilkins.

GAGE, H., KAYE, J., OWEN, C., TREND, P. & WADE, D. 2006. Evaluating rehabilitation using cost-consequences analysis: an example in Parkinson's disease. *Clin Rehabil,* 20**,** 232-8.

GALLOWAY, H. & NEWMAN, E. 2017. Is there a difference between child self-ratings and parent proxy-ratings of the quality of life of children with a diagnosis of attention-deficit hyperactivity disorder (ADHD)? A systematic review of the literature. *ADHD Attention Deficit and Hyperactivity Disorders,* 9**,** 11-29.

GAO, S. S., ZHAO, I. S., HIRAISHI, N., DUANGTHIP, D., MEI, M. L., LO, E. C. M. & CHU, C. H. 2016. Clinical Trials of Silver Diamine Fluoride in Arresting Caries among Children: A Systematic Review. *JDR Clinical & Translational Research,* 1**,** 201-210.

GERMAIN, N., ABALLÉA, S. & TOUMI, M. 2019. Measuring health-related quality of life in young children: how far have we come? *Journal of Market Access & Health Policy,* 7**,** 1618661.

GILCHRIST, F. 2015. Development of a child-centred, caries-specific measure of oral health-related quality of life. University of Sheffield, 2015.

GILCHRIST, F., MARSHMAN, Z., DEERY, C. & RODD, H. D. 2015. The impact of dental caries on children and young people: what they have to say? *International Journal of Paediatric Dentistry,* 25**,** 327-338.

GILCHRIST, F., RODD, H., DEERY, C. & MARSHMAN, Z. 2014. Assessment of the quality of measures of child oral health-related quality of life. *Acta Veterinaria Scandinavica,* 14**,** 40.

GILCHRIST, F., RODD, H. D., DEERY, C. & MARSHMAN, Z. 2018. Development and evaluation of CARIES-QC: a caries-specific measure of quality of life for children. *BMC Oral Health,* 18**,** 202.

GLENDOR, U. 2008. Epidemiology of traumatic dental injuries – a 12 year review of the literature. *Dental Traumatology,* 24**,** 603-611.

GOING, R. E., LOESCHE, W. J., GRAINGER, D. A. & SYED, S. A. 1978. The viability of microorganisms in carious lesions five years after covering with a fissure sealant. *The Journal of the American Dental Association,* 97**,** 455-462.

GOLDMAN, A., CHEN, X., FAN, M. & FRENCKEN, J. 2014. Methods and preliminary findings of a cost-effectiveness study of glass-ionomer-based and composite resin sealant materials after 2 yr. *European journal of oral sciences* [Online], 122. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/787/CN-01116787/frame.html.

GOLDMAN, A. S., CHEN, X., FAN, M. & FRENCKEN, J. E. 2016. Cost-effectiveness, in a randomized trial, of glass-ionomer-based and resin sealant materials after 4 yr. *European Journal of Oral Sciences,* 124**,** 472-479.

GOLDSTON, D. B., KOVACS, M., OBROSKY, D. S. & IYENGAR, S. 1995. A Longitudinal Study of Life Events and Metabolic Control Among Youths With Insulin-Dependent Diabetes Mellitus. *Health Psychology,* 14**,** 409-414.

GOODWIN, E. & GREEN, C. 2016. A Systematic Review of the Literature on the Development of Condition-Specific Preference-Based Measures of Health. *Applied Health Economics and Health Policy,* 14**,** 161-183.

GOODWIN, E., GREEN, C. & SPENCER, A. 2015a. Estimating a Preference-Based Index for an Eight-Dimensional Health State Classification System for Multiple Sclerosis. *Value in Health,* 18**,** 1025-1036.

GOODWIN, M., SANDERS, C., DAVIES, G., WALSH, T. & PRETTY, I. 2015b. Issues arising following a referral and subsequent wait for extraction under general anaesthetic: impact on children. *BMC Oral Health,* 15.

GOV.UK. 2019. *Find and compare schools in England* [Online]. GOV.UK. Available: https://www.compare-school-performance.service.gov.uk/ [Accessed 3rd May 2020].

GRANVILLE-GARCIA, A. F., GOMES, M. C., DANTAS, L. R., DANTAS, L. R., DA SILVA, B. R., PERAZZO MDE, F. & SIQUEIRA, M. B. 2016. Parental influence on children's answers to an oral-health-related quality of life questionnaire. *Braz Oral Res,* 30.

GREWAL, I., LEWIS, J., FLYNN, T., BROWN, J., BOND, J. & COAST, J. 2006. Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science & Medicine,* 62**,** 1891-1901.

GRIFFIN, S. O., GRIFFIN, P. M., GOOCH, B. F. & BARKER, L. K. 2002. Comparing the costs of three sealant delivery strategies. *Journal of Dental Research,* 81**,** 641-5.

GUBA, E. G. & LINCOLN, Y. S. 1981. *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*, Jossey-Bass.

GUSTAFSSON, B. E., QUENSEL, C. E., LANKE, L. S., LUNDQVIST, C., GRAHNEN, H., BONOW, B. E. & KRASSE, B. 1954. The Vipeholm dental caries study; the effect of different levels of carbohydrate intake on caries activity in 436 individuals observed for five years. *Acta Odontol Scand,* 11**,** 232-64.

GUYATT, G., FEENY, D. & PATRICK, D. L. 1993. Measuring health-related quality of life. *Ann. Intern. Med.*

GUYATT, G. H., BOMBARDIER, C. & TUGWELL, P. X. 1986. Measuring disease-specific quality of life in clinical trials. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne,* 134**,** 889-895.

HANDELMAN, S. L., BUONOCORE, M. G. & HESECK, D. J. 1972. A preliminary report on the effect of fissure sealant on bacteria in dental caries. *The Journal of Prosthetic Dentistry,* 27**,** 390-392.

HARDEN, J., SCOTT, S., BACKETT-MILBURN, K. & JACKSON, S. 2000. Can't Talk, Won't Talk?: Methodological Issues in Researching Children. *Sociological Research Online,* 5**,** 1-12.

HARPER, R. 1997. A comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax.,* 52**,** 879-887.

HAYDEN, C., BOWLER, J. O., CHAMBERS, S., FREEMAN, R., HUMPHRIS, G., RICHARDS, D. & CECIL, J. E. 2013. Obesity and dental caries in children: a systematic review and meta-analysis. *Community Dentistry and Oral Epidemiology,* 41**,** 289-308.

HE, S. & WANG, J. 2020. Validation of the Chinese version of the Caries Impacts and Experiences Questionnaire for Children (CARIES-QC. *International Journal of Paediatric Dentistry,* 30**,** 50-56.

HERNANDEZ ALAVA, M., PUDNEY, S. & WAILOO, A. 2020. The EQ-5D-5L Value Set for England: Findings of a Quality Assurance Program. *Value Health,* 23**,** 642-648.

HETTIARACHCHI, R. M., KULARATNA, S., DOWNES, M. J., BYRNES, J., KROON, J., LALLOO, R., JOHNSON, N. W. & SCUFFHAM, P. A. 2017. The cost-effectiveness of oral health interventions: A systematic review of cost- utility analyses. *Community Dentistry and Oral Epidemiology*.

HICHENS, L., ROWLAND, H., WILLIAMS, A., HOLLINGHURST, S., EWINGS, P., CLARK, S., IRELAND, A. & SANDY, J. 2007. Cost-effectiveness and patient satisfaction: Hawley and vacuum-formed retainers. *European Journal of Orthodontics,* 29**,** 372-8.

HIETASALO, P., SEPPA, L., LAHTI, S., NIINIMAA, A., KALLIO, J., ARONEN, P., SINTONEN, H. & HAUSEN, H. 2009. Cost-effectiveness of an experimental caries-control regimen in a 3.4-yr randomized clinical trial among 11-12-yr-old Finnish schoolchildren. *European Journal of Oral Sciences,* 117**,** 728-33.

HIRSCH, G., EDELSTEIN, B., FROSH, M. & ANSELMO, T. 2012. A simulation model for designing effective interventions in early childhood caries (Provisional abstract). *Preventing Chronic Disease* [Online], 9:e66. Available: http://onlinelibrary.wiley.com/o/cochrane/cleed/articles/NHSEED-22012024215/frame.html.

HOGG, C. 1996. *Health services for children and young people : a guide for commissioners and providers,* London, London : Action for Sick Children, 1996.

HOLLAND, T. J., CONSIDINE, J. & CREEDON, P. 2001. The effectiveness and cost of two fluoride program for children. *European Journal of Paediatric Dentistry,* 2**,** 61-66.

HOLT, R. D., NUNN, J. H., ROCK WP & J., P. 1996. British Society of Paediatric Dentistry: A Policy Document on Fluoride Dietary Supplements and Fluoride Toothpastes for Children. *International Journal of Paediatric Dentistry,* 6**,** 139-142.

HOOLEY, M., SKOUTERIS, H., BOGANIN, C., SATUR, J. & KILPATRICK, N. 2012. Body mass index and dental caries in children and adolescents: A systematic review of literature published 2004 to 2011. *Systematic Reviews,* 1**,** <xocs:firstpage xmlns:xocs=""/>.

HORSMAN, J., FURLONG, W., FEENY, D. & TORRANCE, G. 2003. The Health Utilities Index (HUI®): concepts, measurement properties and applications. *Health and Quality of Life Outcomes,* 1**,** 54.

HOSEY, M. T. 2002. UK National Clinical Guidelines in Paediatric Dentistry. Managing anxious children: the use of conscious sedation in paediatric dentistry. *Int J Paediatr Dent,* 12**,** 359-72.

HSCIC 2015. Children’s Dental Health Survey 2013 Report 1: Attitudes, Behaviours and Children’s Dental Health England, Wales and Northern Ireland, 2013 Health and Social Care Information Centre.

HSCIC 2016. Hospital Episode Statistics England 2015-2016. Health and Social Care Information Centre.

HUINC. 2018. *Questionnaire Development, Translations and Support* [Online]. Health Utilities Inc. [Accessed September 2018].

HURLEY, J. C. & UNDERWOOD, M. K. 2002. Children’s Understanding of Their Research Rights before and after Debriefing: Informed Assent, Confidentiality, and Stopping Participation. *Child Development,* 73**,** 132-143.

HUSEREAU, D., DRUMMOND, M., PETROU, S., CARSWELL, C., MOHER, D., GREENBERG, D., AUGUSTOVSKI, F., BRIGGS, A., MAUSKOPF, J. & LODER, E. 2013. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)-Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value in Health,* 16**,** 231-250.

HUYNH, E., COAST, J., ROSE, J., KINGHORN, P. & FLYNN, T. 2017. Values for the ICECAP-Supportive Care Measure (ICECAP-SCM) for use in economic evaluation at end of life. *Social Science & Medicine,* 189**,** 114.

ICOHIRP 2015. Social inequalities in oral health: from evidence to action *In:* WATT, R. G., LISTL, S., PERES, M. & HEILMANN, A. (eds.). International Centre for Oral Health Inequalities Research & Policy.

IHEOZOR-EJIOFOR, Z., WORTHINGTON, H. V., WALSH, T., O'MALLEY, L., CLARKSON, J. E., MACEY, R., ALAM, R., TUGWELL, P., WELCH, V. & GLENNY, A. M. 2015.

Water fluoridation for the prevention of dental caries. *Cochrane Database Syst Rev***,** Cd010856.

INDEX OF MULTIPLE DEPRIVATION. 2015. *Index of Multiple Deprivation* [Online]. Ministry of Housing, Community and Local Government. Available: http://opendatacommunities.org/resource?uri=http%3A%2F%2Fopendatacommunities.org%2Fdata%2Fsocietal-wellbeing%2Fdeprivation%2Fimd-score-2010 [Accessed September 2018].

INNES, N. P., CLARKSON, J. E., SPEED, C., DOUGLAS, G. V., MAGUIRE, A. & FI, C. T. C. 2013a. The FiCTION dental trial protocol - filling children's teeth: indicated or not? *BMC Oral Health,* 13**,** 25.

INNES, N. P. T., CLARKSON, J. E., SPEED, C., DOUGLAS, G. V. A. & MAGUIRE, A. 2013b. The FiCTION dental trial protocol - filling children's teeth: Indicated or not? *BMC Oral Health,* 13.

INNES, N. P. T., EVANS, D. J. P., BONIFACIO, C. C., GENESER, M., HESSE, D., HEIMER, M., KANELLIS, M., MACHIULSKIENE, V., NARBUTAITÉ, J., OLEGÁRIO, I. C., OWAIS, A., ARAUJO, M. P., RAGGIO, D. P., SPLIETH, C., AMERONGEN, E. V., WEBER-GASPARONI, K. & SANTAMARIA, R. M. 2017. The Hall Technique 10 years on: Questions and answers. *BDJ,* 222**,** 478.

INNES, N. P. T., EVANS, D. J. P. & STIRRUPS, D. R. 2011. Sealing Caries in Primary Molars: Randomized Control Trial, 5-year Results. *Journal of Dental Research,* 90**,** 1405-1410.

INVOLVE. 2020. *Involvement cost calculator* [Online]. NIHR. Available: https://www.invo.org.uk/resource-centre/payment-and-recognition-for-public-involvement/involvement-cost-calculator/ [Accessed 6th April 2020].

JACKSON, S., VANN, W., KOTCH, J., PAHEL, B. & LEE, J. 2011. Impact of Poor Oral Health on Children's School Attendance and Performance. *American Journal of Public Health,* 101**,** 1900-6.

JOFFER, J., JERDÉN, L., ÖHMAN, A. & FLACKING, R. 2016. Exploring self-rated health among adolescents: a think-aloud study. *BMC Public Health,* 16**,** 1.

JOHANNESSON, M. & JÖNSSON, B. 1991. Economic evaluation in health care: Is there a role for cost-benefit analysis? *Health policy,* 17**,** 1-23.

JOHNSON, F. R., LANCSAR, E., MARSHALL, D., KILAMBI, V., MÜHLBACHER, A., REGIER, D. A., BRESNAHAN, B. W., KANNINEN, B. & BRIDGES, J. F. P. 2013. Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value in Health,* 16**,** 3-13.

JOKELA, J. & PIENIHAKKINEN, K. 2003. Economic evaluation of a risk-based caries prevention program in preschool children. *Acta Odontologica Scandinavica,* 61**,** 110-114.

KAAKKO, T., SKARET, E., GETZ, T., HUJOEL, P., GREMBOWSKI, D., MOORE, C. & MILGROM, P. 2002. An ABCD program to increase access to dental care for children enrolled in Medicaid in a rural county. *Journal of public health dentistry* [Online], 62. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/843/CN-00459843/frame.html.

KAEWKAMNERDPONG, I. & KRISDAPONG, S. 2018. Oral diseases associated with condition-specific oral health-related quality of life and school performance of Thai primary school children: A hierarchical approach. *Community Dentistry and Oral Epidemiology,* 46**,** 270-279.

KAHNEMAN, D. 2008. *QALYs versus Experience: A Perspective from Experimental Economics*, Office of Health Economics.

KARIMI, M. & BRAZIER, J. 2016. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? *Pharmacoeconomics,* 34**,** 645-9.

KASSEBAUM, N. J., BERNABE, E., DAHIYA, M., BHANDARI, B., MURRAY, C. J. & MARCENES, W. 2015. Global burden of untreated caries: a systematic review and metaregression. *J Dent Res,* 94**,** 650-8.

KEARNEY, K. A., HOPKINS, R. H., MAUSS, A. L. & WEISHEIT, R. A. 1983. Sample Bias Resulting From a Requirement for Written Parental Consent. *The Public Opinion Quarterly,* 47**,** 96-102.

KEENEY, R. L. & RAIFFA, H. 1993. *Decisions with multiple objectives: preferences and value trade-offs*, Cambridge university press.

KILPATRICK, N. M. 1993. Durability of restorations in primary molars. *Journal of Dentistry,* 21**,** 67-73.

KIND, P., KLOSE, K., GUSI, N., OLIVARES, P. R. & GREINER, W. 2015. Can adult weights be used to value child health states? Testing the influence of perspective in valuing EQ-5D-Y. *Quality of Life Research,* 24**,** 2519-2539.

KINDELAN, S. A., DAY, P., NICHOL, R., WILLMOTT, N. & FAYLE, S. A. 2008. UK National Clinical Guidelines in Paediatric Dentistry: stainless steel preformed crowns for primary molars. *International Journal of Paediatric Dentistry,* 18**,** 20-28.

KING, M. T., COSTA, D. S., AARONSON, N. K., BRAZIER, J. E., CELLA, D. F., FAYERS, P. M., GRIMISON, P., JANDA, M., KEMMLER, G., NORMAN, R., PICKARD, A. S., ROWEN, D., VELIKOVA, G., YOUNG, T. A. & VINEY, R. 2016. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. *Qual Life Res,* 25**,** 625-36.

KIRK, S. 2007. Methodological and ethical issues in conducting qualitative research with children and young people: A literature review. *International Journal of Nursing Studies,* 44**,** 1250-1260.

KLEIN, H., PALMER, C. & KNUTSON, J. 1938. Studies on dental caries I: dental status and needs of elementary school children. *Public Health Reports,* 53**,** 751-765.

KLEIN, H. & PALMER, C. E. 1940. Studies On Dental Caries:X. a Procedure for the Recording and Statistical Processing of Dental Examination Findings. *Journal of Dental Research,* 19**,** 243-256.

KNAPP, R. 2019. The impact of dental caries and its treatment under general anaesthetic on the everyday lives of children and their families. University of Sheffield (Dentistry)

KOH, R., PUKALLUS, M., KULARATNA, S., GORDON, L. G., BARNETT, A. G., WALSH, L. J. & SEOW, W. K. 2015. Relative cost-effectiveness of home visits and telephone contacts in preventing early childhood caries. *Community Dentistry & Oral Epidemiology,* 43**,** 560-8.

KOWASH, M. B., TOUMBA, K. J. & CURZON, M. E. 2006. Cost-effectiveness of a long-term dental health education program for the prevention of early childhood caries. *European Archives of Paediatric Dentistry: Official Journal of the European Academy of Paediatric Dentistry,* 7**,** 130-5.

KRISDAPONG, S., PRASERTSOM, P., RATTANARANGSIMA, K. & SHEIHAM, A. 2012. Relationships between oral diseases and impacts on Thai schoolchildren's quality of life: Evidence from a Thai national oral health survey of 12- and 15-year-olds. *Community Dentistry and Oral Epidemiology,* 40**,** 550-559.

KRISDAPONG, S., SHEIHAM, A. & TSAKOS, G. 2009. Oral health-related quality of life of 12- and 15-year-old Thai children: findings from a national survey. *Community Dentistry & Oral Epidemiology,* 37**,** 509-17.

KROSNICK, J. A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology,* 5**,** 213-236.

KROSNICK, J. A. 1999. SURVEY RESEARCH. *Annu. Rev. Psychol.,* 50**,** 537-567.

KRUCIEN, N., SICSIC, J. & RYAN, M. 2019. For better or worse? Investigating the validity of best–worst discrete choice experiments in health. *Health Economics,* 28**,** 572-586.

KRUCIEN, N., WATSON, V. & RYAN, M. 2017. Is Best–Worst Scaling Suitable for Health State Valuation? A Comparison with Discrete Choice Experiments. *Health Economics,* 26**,** e1-e16.

LANCSAR, E. & LOUVIERE, J. 2006. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics,* 15**,** 7.

LANDIS, J. R. & KOCH, G. G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics,* 33**,** 159-174.

LESKINEN, K., SALO, S., SUNI, J. & LARMAS, M. 2008. Practice-based study of the cost-effectiveness of fissure sealants in Finland. *Journal of Dentistry,* 36**,** 1074-9.

LEVEY, C., INNES, N., SCHWENDICKE, F., LAMONT, T. & GÖSTEMEYER, G. 2017. Outcomes in randomised controlled trials in prevention and management of carious lesions: a systematic review. 18**,** 515.

LEVINE, R., PITTS, N. & NUGENT, Z. 2002. The fate of 1,587 unrestored carious deciduous teeth: a retrospective general dental practice based study from northern England. *British Dental Journal,* 193**,** 99-103.

LLODRA, J. C., RODRIGUEZ, A., FERRER, B., MENARDIA, V., RAMOS, T. & MORATO, M. 2016. Efficacy of Silver Diamine Fluoride for Caries Reduction in Primary Teeth and First Permanent Molars of Schoolchildren: 36-month Clinical Trial. *J Dent Res,* 84**,** 721-724.

LLOYD, A., NAFEES, B., NAREWSKA, J., DEWILDE, S. & WATKINS, J. 2006. Health state utilities for metastatic breast cancer. *British journal of cancer,* 95**,** 683-690.

LOCKER, D. & ALLEN, F. 2007. What do measures of 'oral health-related quality of life' measure? *Community Dent Oral Epidemiol,* 35**,** 401-11.

LOCKER, D., JOKOVIC, A., STEPHENS, M., KENNY, D., TOMPSON, B. & GUYATT, G. 2002. Family impact of child oral and oro-facial conditions. *Community Dent Oral Epidemiol,* 30**,** 438-48.

LORD, J., LONGWORTH, L., SINGH, J., ONYIMADU, O., FRICKE, J. & BAYLISS, S. 2015a. *Economic Analysis of Oral Health Promotion Approaches for Dental Teams: NICE Public Health Guideline on Oral Health Promotion Approaches for Dental Teams (GID-PHG60),* London, National Institute for Health and Care Excellence.

LORD, J., LONGWORTH, L., SINGH, J., ONYIMADU, O., FRICKE, J., BAYLISS, S. & MEADS, C. 2015b. Oral Health Guidance – Economic analysis of oral health promotion approaches for dental teams Birmingham & Brunel Consortium: External Assessment Centre.

LOUVIERE, J. J., FLYNN, T. N. & CARSON, R. T. 2010. Discrete Choice Experiments Are Not Conjoint Analysis. *Journal of Choice Modelling,* 3**,** 57-72.

LUOTO, A., LAHTI, S., NEVANPERA, T., TOLVANEN, M. & LOCKER, D. 2009. Oral-health-related quality of life among children with and without dental fear. *International Journal of Paediatric Dentistry,* 19**,** 115-20.

MACKENZIE, R. & DIXON, A. K. 1995. Measuring the effects of imaging: an evaluative framework. *Clin Radiol,* 50**,** 513-8.

MACPHERSON, L. M., BALL, G. E., BREWSTER, L., DUANE, B., HODGES, C., WRIGHT, W., GNICH, W., RODGERS, J., MCCALL, D. R., TURNER, S. & CONWAY, D. I. 2010. Childsmile: the national child oral health improvement programme in Scotland. Part 1: establishment and development. *BDJ,* 209**,** 73.

MACPHERSON, L. M. D., ANOPA, Y., CONWAY, D. I. & MCMAHON, A. D. 2013. National Supervised Toothbrushing Program and Dental Decay in Scotland. *Journal of Dental Research,* 92**,** 109-113.

MAGUIRE, A., CLARKSON, J. E., DOUGLAS, G. V. A., RYAN, V., HOMER, T., MARSHMAN, Z., MCCOLL, E., WILSON, N., VALE, L., ROBERTSON, M., ABOUHAJAR, A., HOLMES, R. D., FREEMAN, R., CHADWICK, B., DEERY, C., WONG, F. & INNES, N. P. T. 2020. Best-practice prevention alone or with conventional or biological caries management for 3- to 7-year-olds: the FiCTION three-arm RCT. 24**,** 1.

MAJEWSKI, R. F., SNYDER, C. W. & BERNAT, J. E. 1988. Dental emergencies presenting to a children's hospital. *ASDC J Dent Child,* 55**,** 339-42.

MANFREDA, K. L. & VEHOVAR, V. Survey design features influencing response rates in web surveys. The international conference on improving surveys proceedings, 2002. Citeseer, 25-28.

MARCHIONI, D. M. L., DIAS DE OLIVEIRA LATORRE, M. D. R., ELUF-NETO, J., WÜNSCH-FILHO, V. & FISBERG, R. M. 2005. Identification of dietary patterns using factor analysis in an epidemiological study in São Paulo. *São Paulo Medical Journal,* 123**,** 124-127.

MARINHO, V. C., CHONG, L. Y., WORTHINGTON, H. V. & WALSH, T. 2016. Fluoride mouthrinses for preventing dental caries in children and adolescents. *Cochrane Database Syst Rev,* 7**,** Cd002284.

MARINHO, V. C., HIGGINS, J. P., SHEIHAM, A. & LOGAN, S. 2004. Combinations of topical fluoride (toothpastes, mouthrinses, gels, varnishes) versus single topical fluoride for preventing dental caries in children and adolescents. *Cochrane Database Syst Rev***,** Cd002781.

MARINHO, V. C., WORTHINGTON, H. V., WALSH, T. & CLARKSON, J. E. 2013. Fluoride varnishes for preventing dental caries in children and adolescents. *Cochrane Database Syst Rev***,** Cd002279.

MARINHO, V. C. C., HIGGINS, J., LOGAN, S. & SHEIHAM, A. 2003. Fluoride toothpastes for preventing dental caries in children and adolescents. *Cochrane Database of Systematic Reviews*.

MARINO, R., FAJARDO, J., ARANA, A., GARCIA, C. & PACHAS, F. 2011. Modeling an economic evaluation of a salt fluoridation program in Peru (Provisional abstract). *Journal of Public Health Dentistry* [Online], 71. Available: http://onlinelibrary.wiley.com/o/cochrane/cleed/articles/NHSEED-22011001755/frame.html.

MARINO, R., FAJARDO, J. & MORGAN, M. 2012. Cost-effectiveness models for dental caries prevention programmes among Chilean schoolchildren. *Community Dental Health,* 29**,** 302-8.

MARINO, R., MORGAN, M., WEITZ, A. & VILLA, A. 2007. The cost-effectiveness of adding fluorides to milk-products distributed by the National Food Supplement Programme (PNAC) in rural areas of Chile. *Community Dental Health,* 24**,** 75-81.

MARIÑO, R. J., KHAN, A. R. & MORGAN, M. 2013. Systematic review of publications on economic evaluations of caries prevention programs. *Caries Research,* 47**,** 265-272.

MARSHALL, D., BRIDGES, J., HAUBER, B., CAMERON, R., DONNALLEY, L., FYIE, K. & JOHNSON, R. 2010. Conjoint Analysis Applications in Health - How are Studies being Designed and Reported? An Update on Current Practice in the Published Literature between 2005 and 2008. *Patient,* 3**,** 249-256.

MARSHMAN, Z., AINSWORTH, H., CHESTNUTT, I. G., DAY, P., DEY, D., EL YOUSFI, S., FAIRHURST, C., GILCHRIST, F., HEWITT, C., JONES, C., KELLAR, I., PAVITT, S., ROBERTSON, M., SHAH, S., STEVENS, K., TORGERSON, D. & INNES, N. 2019. Brushing RemInder 4 Good oral HealTh (BRIGHT) trial: Does an SMS behaviour change programme with a classroom-based session improve the oral health of young people living in deprived areas? A study protocol of a randomised controlled trial. *Trials,* 20.

MARSHMAN, Z., GIBSON, B. J., OWENS, J., RODD, H. D., MAZEY, H., BAKER, S. R., BENSON, P. E. & ROBINSON, P. G. 2007. Seen but not heard: a systematic review of the place of the child in 21st-century dental research. *Int J Paediatr Dent,* 17**,** 320-7.

MARSHMAN, Z., GUPTA, E., BAKER, S. R., ROBINSON, P. G., OWENS, J., RODD, H. D., BENSON, P. E. & GIBSON, B. 2015. Seen and heard: towards child participation in dental research. *Int J Paediatr Dent,* 25**,** 375-82.

MARSHMAN, Z. & HALL, M. J. 2008. Oral health research with children. *International Journal of Paediatric Dentistry,* 18**,** 235-242.

MARSHMAN, Z., MORGAN, A., PORRITT, J., GUPTA, E., BAKER, S., CRESWELL, C., NEWTON, T., STEVENS, K., WILLIAMS, C., PRASAD, S., KIRBY, J. & RODD, H. 2016. Protocol for a feasibility study of a self-help cognitive behavioural therapy resource for the reduction of dental anxiety in young people.

MAVRANEZOULI, I., BRAZIER, J. E., ROWEN, D. & BARKHAM, M. 2013. Estimating a Preference-Based Index from the Clinical Outcomes in Routine Evaluation–Outcome Measure (CORE-OM): Valuation of CORE-6D. *Med Decis Making,* 33**,** 381-395.

MCCABE, C., STEVENS, K., ROBERTS, J. & BRAZIER, J. 2005. Health state values for the HUI 2 descriptive system: results from a UK survey. *Health Economics,* 14**,** 231-244.

MCDONAGH, M. S., WHITING, P. F., WILSON, P. M., SUTTON, A. J., CHESTNUTT, I., COOPER, J., MISSO, K., BRADLEY, M., TREASURE, E. & KLEIJNEN, J. 2000. Systematic review of water fluoridation. *Bmj,* 321**,** 855-9.

MCDOWELL, I. 1996. Measuring health a guide to rating scales and questionnaires. *Measuring health*.

MCFADDEN, D. 1973. Conditional logit analysis of qualitative choice behavior.

MCHUGH, M. L. 2012. Interrater reliability: the kappa statistic. *Biochemia medica,* 22**,** 276-282.

MCKNIGHT, K. 2018. The first generation of true digital natives. *Research World,* 2018**,** 14-17.

MCTAGGART-COWAN, H. M., BRAZIER, J. E. & TSUCHIYA, A. 2010. Clustering Rasch Results: A Novel Method for Developing Rheumatoid Arthritis States for Use in Valuation Studies. *Value in Health,* 13**,** 787-795.

MEHREZ, A. & GAFNI, A. 1993. Healthy-years Equivalents versus Quality-adjusted Life Years:In Pursuit of Progress. *Medical Decision Making,* 13**,** 287-292.

MENEZES ABREU, D., LEAL, S., MULDER, J. & FRENCKEN, J. 2011. Patterns of dental anxiety in children after sequential dental visits. *European Archives of Paediatric Dentistry,* 12**,** 298-302.

MERTZ-FAIRHURST, E. J., CURTIS, J. W., ERGLE, J. W., RUEGGEBERG, F. A. & ADAIR, S. M. 1998. ULTRACONSERVATIVE AND CARIOSTATIC SEALED RESTORATIONS: RESULTS AT YEAR 10. *The Journal of the American Dental Association,* 129**,** 55-66.

MICHALEK, S. M., KATZ, J. & CHILDERS, N. K. 2001. A vaccine against dental caries: an overview. *BioDrugs,* 15**,** 501-8.

MILLER, D. 1996. Comparisons of pain ratings from postoperative children, their mothers, and their nurses. *Pediatr Nurs,* 22**,** 145-9.

MILLER, J., VAUGHAN-WILLIAMS, E., FURLONG, R. & HARRISON, L. 1982. Dental caries and children's weights. *Journal of Epidemiology and Community Health,* 36**,** 49.

MILSOM, K., TICKLE, M. & BLINKHORN, A. 2002a. Dental pain and dental treatment of young children attending the general dental service. *British Dental Journal,* 192**,** 280-4.

MILSOM, K. M., TICKLE, M. & BLINKHORN, A. 2002b. The prescription and relative outcomes of different materials used in general dental practice in the north west region of England to restore the primary dentition. *Journal of Dentistry,* 30**,** 77-82.

MILSOM, K. M., TICKLE, M. & KING, D. 2003. Does the dental profession know how to care for the primary dentition? *British Dental Journal,* 195**,** 301.

MINISTRY OF HOUSING, COMMUNITIES & LOCAL GOVERNMENT,. 2019. *English indices of deprivation* [Online]. Available: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [Accessed 25th May 2020].

MITTON, C. & DONALDSON, C. 2009. *Priority setting toolkit: guide to the use of economics in healthcare decision making*, John Wiley & Sons.

MOKKINK, L., TERWEE, C., PATRICK, D., ALONSO, J., STRATFORD, P., KNOL, D., BOUTER, L. & VET, H. 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation - Official Journal of the International Society of Quality of Life Research,* 19**,** 539-549.

MONSE, B., HEINRICH-WELTZIEN, R., MULDER, J., HOLMGREN, C. & VAN PALENSTEIN HELDERMAN, W. H. 2012. Caries preventive efficacy of silver diammine fluoride (SDF) and ART sealants in a school-based daily fluoride toothbrushing program in the Philippines. *BMC Oral Health,* 12**,** 52-52.

MOODIE, M., RICHARDSON, J., RANKIN, B., IEZZI, A. & SINHA, K. 2010. Predicting Time Trade-Off Health State Valuations of Adolescents in Four Pacific Countries Using the Assessment of Quality-of-Life (AQoL-6D) Instrument. *Value in Health,* 13**,** 1014-1027.

MORGAN, D. L. 1998. Practical Strategies for Combining Qualitative and Quantitative Methods: Applications to Health Research. *Qualitative Health Research,* 8**,** 362-376.

MORGAN, M. V., CROWLEY, S. J. & WRIGHT, C. 1998. Economic evaluation of a pit and fissure dental sealant and fluoride mouthrinsing program in two nonfluoridated regions of Victoria, Australia. *Journal of Public Health Dentistry,* 58**,** 19-27.

MORRIS, S., DEVLIN, N., PARKIN, D. & SPENCER, A. 2014. *Economic Analysis in Healthcare*, Wiley.

MOTT, D., SHAH, K., RAMOS-GOÑI, J., DEVLIN, N. & RIVERO-ARIAS, O. 2019. Valuing EQ-5D-Y health states using a discrete choice experiment: do adult and adolescent preferences differ?

MOYNIHAN, P. 2016. Sugars and Dental Caries: Evidence for Setting a Recommended Threshold for Intake. *Advances in Nutrition,* 7**,** 149-156.

MOYNIHAN, P. J. & KELLY, S. A. 2014. Effect on caries of restricting sugars intake: systematic review to inform WHO guidelines. *J Dent Res,* 93**,** 8-18.

MUKURIA, C., ROWEN, D., HARNAN, S., RAWDIN, A., WONG, R., ARA, R. & BRAZIER, J. 2019. An Updated Systematic Review of Studies Mapping (or Cross-Walking) Measures of Health-Related Quality of Life to Generic Preference-Based Measures to Generate Utility Values. *Applied health economics and health policy,* 17**,** 295-313.

MULHERN, B., LABEIT, A., ROWEN, D., KNOWLES, E., MEADOWS, K., ELLIOTT, J. & BRAZIER, J. 2017. Developing preference-based measures for diabetes: DHP-3D and DHP-5D. *Diabet Med,* 34**,** 1264-1275.

MULHERN, B., LONGWORTH, L., BRAZIER, J., ROWEN, D., BANSBACK, N., DEVLIN, N. & TSUCHIYA, A. 2013a. Binary Choice Health State Valuation and Mode of Administration: Head-to-Head Comparison of Online and CAPI. *Value in Health,* 16**,** 104-113.

MULHERN, B., NORMAN, R., SHAH, K., BANSBACK, N., LONGWORTH, L. & VINEY, R. 2018. How Should Discrete Choice Experiments with Duration Choice Sets Be Presented for the Valuation of Health States? *Medical Decision Making,* 38**,** 306-318.

MULHERN, B., ROWEN, D., BRAZIER, J., SMITH, S., ROMEO, R., TAIT, R., WATCHURST, C., CHUA, K. C., LOFTUS, V., YOUNG, T., LAMPING, D., KNAPP, M., HOWARD, R. & BANERJEE, S. 2013b. Development of DEMQOL-U and DEMQOL-PROXY-

U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. *Health Technol Assess,* 17**,** v-xv, 1-140.

MULHERN, B., SMITH, S. C., ROWEN, D., BRAZIER, J. E., KNAPP, M., LAMPING, D. L., LOFTUS, V., YOUNG, T. A., HOWARD, R. J. & BANERJEE, S. 2012. Improving the measurement of QALYs in dementia: developing patient- and carer-reported health state classification systems using Rasch analysis. *Value Health,* 15**,** 323-33.

MULHERN, B. J., BANSBACK, N., NORMAN, R. & BRAZIER, J. 2020. Valuing the SF-6Dv2 classification system in the United Kingdom using a discrete-choice experiment with duration.

NADLER, J. T., WESTON, R. & VOYLES, E. C. 2015. Stuck in the Middle: The Use and Interpretation of Mid-Points in Items on Questionnaires. *The Journal of General Psychology,* 142**,** 71-89.

NATIONAL HEALTH AND HOSPITALS REFORM COMMISSION 2009. A healthier future for all Australians- Final Report of the National Health and Hospitals Reform Commission.

NAWCH 1990. *Setting standards for adolescents in hospital,* London, London : National Association for the Welfare of Children in Hospital, 1990.

NEIDELL, M., SHEARER, B. & LAMSTER, I. B. 2016. Cost-Effectiveness Analysis of Dental Sealants versus Fluoride Varnish in a School-Based Setting. *Caries Research,* 50**,** 78-82.

NETTEN, A., BURGE, P., MALLEY, J., POTOGLOU, D., TOWERS, A. M., BRAZIER, J., FLYNN, T., FORDER, J. & WALL, B. 2012. Outcomes of social care for adults: Developing a preference-weighted measure. *Health Technology Assessment,* 16**,** 1-165.

NEVILLE, B. A., FORSTER, S. C. & LAWLEY, T. D. 2018. Commensal Koch's postulates: establishing causation in human microbiota research. *Current Opinion in Microbiology,* 42**,** 47-52.

NEY, J. P., VAN DER GOES, D. N. & CHI, D. L. 2014. Economic modeling of sealing primary molars using a "value of information" approach. *Journal of Dental Research,* 93**,** 876-81.

NHS 2004. Improving the patient experience; Friendly healthcare environments for children and young people. London.

NHS DIGITAL 2019. Hospital Admitted Patient Care and Adult Critical Care Activity. *In:* DIGITAL, N. (ed.).

NHS DIGITAL. 2020. *Ethnic Category Code* [Online]. NHS. Available: https://www.datadictionary.nhs.uk/data_dictionary/attributes/e/end/ethnic_category_code_de.asp?shownav=1 [Accessed 6th April 2020].

NICE 2010. Cardiovascular disease prevention (PH25). Manchester: National Institute for Health and Care Excellence.

NICE 2012. NICE Process and Methods Guides. *Methods for the Development of NICE Public Health Guidance.* London: National Institute for Health and Care Excellence.

NICE 2013. Guide to the methods of technology appraisal. London: National Institute of Health and Care Excellence.

NICE 2015. Oral health promotion: general dental practice. National Institute for Health and Care Excellence London.

NICE 2017. Position statement on use of the EQ-5D-5L valuation set. London: National Institute for Health and Care Excellence.

NICE 2018. Technology appraisal guidance. Manchester: National Institute for Health and Care Excellence.

NISRA. 2017. *Northern Ireland Multiple Deprivation Measure* [Online]. Northern Ireland Statistics and Research Agency. Available: https://www.nisra.gov.uk/statistics/deprivation/northern-ireland-multiple-deprivation-measure-2017-nimdm2017 [Accessed 25th May 2020].

NORMAN, G. R. 2014. *Biostatistics [electronic resource] : the bare essentials*, 2014.

NORMAN, R., CRONIN, P. & VINEY, R. 2013. A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. *Appl Health Econ Health Policy,* 11**,** 287-98.

NORMAN, R., MERCIECA-BEBBER, R., ROWEN, D., BRAZIER, J., CELLA, D., A. SIMON, P., DEBORAH J, S., VINEY, R., REVICKI, D. & KING, M. 2019. UK utility weights for the EORTC QLU-C10D.

NORMAN, R., VINEY, R., AARONSON, N. K., BRAZIER, J. E., CELLA, D., COSTA, D. S. J., FAYERS, P. M., KEMMLER, G., PEACOCK, S., PICKARD, A. S., ROWEN, D., STREET, D. J., VELIKOVA, G., YOUNG, T. A. & KING, M. T. 2016. Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format.

NUTTALL, N. M., GILBERT, A. & MORRIS, J. 2008. Children's dental anxiety in the United Kingdom in 2003. *Journal of Dentistry,* 36**,** 857-860.

OFCOM 2010. HFSS advertising restrictions: final review. *In:* OFCOM (ed.).

OFFICE FOR NATIONAL STATISTICS 2009. General Household Survey - 2007 report. London: ONS.

OFSTED. 2020. *Find an inspection report* [Online]. Ofsted. Available: https://reports.ofsted.gov.uk/ [Accessed 3rd May 2020].

OLIVEIRA, B. H., RAJENDRA, A., VEITZ-KEENAN, A. & NIEDERMAN, R. 2019. The Effect of Silver Diamine Fluoride in Preventing Caries in the Primary Dentition: A Systematic Review and Meta-Analysis. *Caries Res,* 53**,** 24-32.

ONS. 2011. *2011 Census* [Online]. Office for National Statistics. Available: https://www.ons.gov.uk/census/2011census [Accessed 9th June 2020].

OSCARSON, N., KALLESTAL, C., FJELDDAHL, A. & LINDHOLM, L. 2003. Cost-effectiveness of different caries preventive measures in a high-risk population of Swedish adolescents. *Community Dentistry & Oral Epidemiology,* 31**,** 169-78.

OTTO, M. 2017. How can a child die of toothache in the US? *The Guardian*.

OUYANG, W. 2009. Cost -effectiveness analysis of dental sealant using econometric modeling.

PALFREYMAN, S. & BRAZIER, J. 2012. PCV80 The Development and Valuation of the Sheffield Preference Based Leg Ulcer Questionnaire (SPVU-5D). *Value in Health,* 15**,** A126-A126.

PALLANT, J. F. & TENNANT, A. 2007. An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS. *British Journal of Clinical Psychology,* 46**,** 1-18.

PARACHA, N., ABDULLA, A. & MACGILCHRIST, K. S. 2018. Systematic review of health state utility values in metastatic non-small cell lung cancer with a focus on previously treated patients. *Health and quality of life outcomes,* 16**,** 179-179.

PARKIN, D., MORRIS, S. & DEVLIN, N. 2015. *Economic appraisal in public healthcare: assessing efficiency and equity*, Oxford University Press.

PATRICK, D. I. 1973. Methods for measuring levels of well-being for a health status index. *Health services research : HSR.,* 8**,** 228-245.

PATRICK, D. L., BURKE, L. B., GWALTNEY, C. J., LEIDY, N. K., MARTIN, M. L., MOLSEN, E. & RING, L. 2011. Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2—Assessing Respondent Understanding. *Value in Health,* 14**,** 978-988.

PAU, A., BAXEVANOS, K. & CROUCHER, R. 2007. Family structure is associated with oral pain in 12-year-old Greek schoolchildren. *Int. J. Paediatr. Dent.,* 17**,** 345-351.

PAYNE, K., FARGHER, E. A., ROBERTS, S. A., TRICKER, K., ELLIOTT, R. A., RATCLIFFE, J. & NEWMAN, W. G. 2011. Valuing pharmacogenetic testing services: A comparison of patients' and health care professionals' preferences. *Value in Health,* 14**,** 121-134.

PEPPER, D., HODGEN, J., LAMESOO, K., KÕIV, P. & TOLBOOM, J. 2018. Think aloud: using cognitive interviewing to validate the PISA assessment of student self-efficacy in mathematics. *International Journal of Research & Method in Education,* 41**,** 3-16.

PERES, M. A., MACPHERSON, L. M. D., WEYANT, R. J., DALY, B., VENTURELLI, R., MATHUR, M. R., LISTL, S., CELESTE, R. K., GUARNIZO-HERREÑO, C. C., KEARNS, C., BENZIAN, H., ALLISON, P. & WATT, R. G. 2019. Oral diseases: a global public health challenge. *The Lancet,* 394**,** 249-260.

PERES, M. A., PERES, K. G., DE BARROS, A. J. & VICTORA, C. G. 2007. The relation between family socioeconomic trajectories from childhood to adolescence and dental caries and associated oral behaviours. *Journal of Epidemiology & Community Health,* 61**,** 141-145.

PETRÉN, S., BJERKLIN, K., MARKÉ, L. Å. & BONDEMARK, L. 2013. Early correction of posterior crossbite-a cost-minimization analysis. *European Journal of Orthodontics,* 35**,** 14-21.

PETROU, S. 2003. Methodological issues raised by preference-based approaches to measuring the health status of children. *Health Economics,* 12**,** 697-702.

PETTITT, D., RAZA, S., NAUGHTON, B., ROSCOE, A., RAMAKRISHNAN, A., ALI, A., DAVIES, B., DOPSON, S., HOLLANDER, G. & SMITH, J. 2016. The limitations of QALY: a literature review. *Journal of Stem Cell Research and Therapy,* 6.

PHATTARATARATIP, E., OLSON, B., BROFFITT, B., QIAN, F., BROGDEN, K. A., DRAKE, D. R., LEVY, S. M. & BANAS, J. A. 2011. Streptococcus mutans strains recovered from caries-active or caries-free individuals differ in sensitivity to host antimicrobial peptides. *Molecular oral microbiology,* 26**,** 187-199.

PIKO, B. F. & BAK, J. 2006. Children's perceptions of health and illness: images and lay concepts in preadolescence. *Health Education Research,* 21**,** 643-653.

PINE, C., ADAIR, P., BURNSIDE, G., ROBINSON, L., EDWARDS, R. T., ALBADRI, S., CURNOW, M., GHAHREMAN, M., HENDERSON, M., MALIES, C., WONG, F., MUIRHEAD, V., WESTON-PRICE, S. & WHITEHEAD, H. 2015. A new primary dental care service compared with standard care for child and family to reduce the re-occurrence of childhood dental caries (Dental RECUR): study protocol for a randomised controlled trial. *Trials,* 16**,** 505.

PINE, C. M., ADAIR, P. M., BURNSIDE, G., BRENNAN, L., SUTTON, L., EDWARDS, R. T., EZEOFOR, V., ALBADRI, S., CURNOW, M. M., DEERY, C., HOSEY, M. T., WILLIS-LAKE, J., LYNN, J., PARRY, J. & WONG, F. S. L. 2020. Dental RECUR Randomized Trial to Prevent Caries Recurrence in Children. *J Dent Res,* 99**,** 168-174.

PITTS, N., CHADWICK, B. & ANDERSON, T. 2015. Children's Dental Health Survey 2013. Report 2: Dental Disease and Damage in Children, England, Wales and Northern Ireland. Health and Social Care Information Centre.

PITTS, N. B., ZERO, D. T., MARSH, P. D., EKSTRAND, K., WEINTRAUB, J. A., RAMOS-GOMEZ, F., TAGAMI, J., TWETMAN, S., TSAKOS, G. & ISMAIL, A. 2017. Dental caries. *Nature Reviews Disease Primers,* 3**,** 17030.

POKORNY, S. B., JASON, L. A., SCHOENY, M. E., TOWNSEND, S. M. & CURIE, C. J. 2001. Do Participation Rates Change when Active Consent Procedures Replace Passive Consent. *Evaluation Review,* 25**,** 567-580.

POST 2016. Sugar and Health Policy. *In:* POSTNOTE (ed.). Houses of Parliament: Office of Science and Technology.

POTOGLOU, D., BURGE, P., FLYNN, T., NETTEN, A., MALLEY, J., FORDER, J. & BRAZIER, J. E. 2011. Best–worst scaling vs. discrete choice experiments: An empirical comparison using social care data. *Social Science & Medicine,* 72**,** 1717-1727.

POWELL, P. A., CARLTON, J., ROWEN, D. & BRAZIER, J. E. 2019. Producing a preference-based quality of life measure for people with Duchenne muscular dystrophy: a mixed-methods study protocol. *BMJ Open,* 9.

PROCHASKA, J. J., SPRING, B. & NIGG, C. R. 2008. Multiple health behavior change research: an introduction and overview. *Preventive medicine,* 46**,** 181-188.

PUBLIC HEALTH ENGLAND 2014a. Local authorities improving oral health: commissioning better oral health for children and young people. An evidence-informed toolkit for local authorities. . London: Public Health England.

PUBLIC HEALTH ENGLAND 2014b. Oral health survey of three-year-old children 2013 - a report on the prevalence and severity of dental decay London: PHE.

PUBLIC HEALTH ENGLAND 2015. Sugar Reduction: the evidence for action *In:* ENGLAND, P. H. (ed.). London.

PUBLIC HEALTH ENGLAND 2017. Delivering better oral health: an evidence-based toolkit for prevention *In:* HEALTH, D. O. (ed.) 3rd Edition ed. London: PHE.

PUBLIC HEALTH ENGLAND 2018. Oral health survey of 5-year-olds 2017. A report on the variations in prevalence and severity of dental decay. *National Dental Epidemiology Programme for England.* London.

PUBLIC HEALTH ENGLAND 2020. Oral health survey of 5-year-olds 2019. A report on the variations in prevalence and severity of dental decay. *National Dental Epidemiology Programme for England.* London.

PUKALLUS, M., PLONKA, K., KULARATNA, S., GORDON, L., BARNETT, A. G., WALSH, L. & SEOW, W. K. 2013. Cost-effectiveness of a telephone-delivered education programme to prevent early childhood caries in a disadvantaged area: a cohort study. *Bmj Open,* 3**,** 8.

QUIÑONEZ, R. B., DOWNS, S. M., SHUGARS, D., CHRISTENSEN, J. & VANN JR, W. F. 2005. Assessing cost-effectiveness of sealant placement in children. *Journal of Public Health Dentistry,* 65**,** 82-89.

QUINONEZ, R. B., STEARNS, S. C., TALEKAR, B. S., ROZIER, R. G. & DOWNS, S. M. 2006. Simulating cost-effectiveness of fluoride varnish during well-child visits for medicaid-enrolled children. *Archives of Pediatrics & Adolescent Medicine,* 160**,** 164-170.

RABIN, R. & DE CHARRO, F. 2001. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med,* 33**,** 337.

RAMAEKERS, B. L. T., JOORE, M. A. & GRUTTERS, J. P. C. 2013. How Should We Deal with Patient Heterogeneity in Economic Evaluation: A Systematic Review of National Pharmacoeconomic Guidelines. *Value in Health,* 16**,** 855-862.

RAMOS-GOMEZ, F. J. & SHEPARD, D. S. 1999. Cost-effectiveness model for prevention of early childhood caries. *Journal of the California Dental Association,* 27**,** 539-44.

RAMP, M., KHAN, F., MISAJON, R. A. & PALLANT, J. F. 2009. Rasch analysis of the Multiple Sclerosis Impact Scale MSIS-29. *Health Qual Life Outcomes,* 7**,** 58.

RASCH, G. 1960. *Probabilistic models for some intelligence and attainment tests,* Chicago IL, University of Chicago Press.

RASMUSSEN, J. K., FREDERIKSEN, J. A., HALLONSTEN, A. L. & POULSEN, S. 2005. Danish dentists' knowledge, attitudes and management of procedural dental pain in children: association with demographic characteristics, structural factors, perceived stress during the administration of local analgesia and their tolerance towards pain. *Int J Paediatr Dent,* 15**,** 159-68.

RATCLIFFE, J. 2000. THE USE OF CONJOINT ANALYSIS TO ELICIT WILLINGNESS-TO-PAY VALUES Proceed With Caution? *Inter. J. Tech. Assess. Health Care,* 16.

RATCLIFFE, J., CHEN, G., STEVENS, K., BRADLEY, S., COUZNER, L., BRAZIER, J., SAWYER, M., ROBERTS, R., HUYNH, E. & FLYNN, T. 2015. Valuing Child Health Utility 9D Health States with Young Adults: Insights from a Time Trade Off Study.

RATCLIFFE, J., COUZNER, L., FLYNN, T., SAWYER, M., STEVENS, K., BRAZIER, J. & BURGESS, L. 2011. Valuing child health utility 9D health states with a young adolescent sample. *Applied Health Economics and Health Policy,* 9**,** 15-27.

RATCLIFFE, J., HUYNH, E., CHEN, G., STEVENS, K., SWAIT, J., BRAZIER, J., SAWYER, M., ROBERTS, R. & FLYNN, T. 2016a. Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. *Social Science & Medicine,* 157**,** 48-59.

RATCLIFFE, J., HUYNH, E., STEVENS, K., BRAZIER, J., SAWYER, M. & FLYNN, T. 2016b. Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best–Worst Scaling. *Health Economics,* 25**,** 486-496.

RATCLIFFE, J., STEVENS, K., FLYNN, T., BRAZIER, J. & SAWYER, M. 2012a. An assessment of the construct validity of the CHU9D in the Australian adolescent general population. *Quality of Life Research,* 21**,** 717-725.

RATCLIFFE, J., STEVENS, K., FLYNN, T., BRAZIER, J. & SAWYER, M. G. 2012b. Whose values in health? An empirical comparison of the application of adolescent and adult values for the CHU-9D and AQOL-6D in the Australian adolescent general population. *Value in Health,* 15**,** 730-6.

RATNAYAKE, N. & EKANAYAKE, L. 2005. Prevalence and impact of oral pain in 8-year-old children in Sri Lanka. *International Journal of Paediatric Dentistry,* 15**,** 105-112.

RAVENS-SIEBERER, U., WILLE, N., BADIA, X., BONSEL, G., BURSTRÖM, K., CAVRINI, G., DEVLIN, N., EGMAR, A. C., GUSI, N., HERDMAN, M., JELSMA, J., KIND, P., OLIVARES, P. R., SCALONE, L. & GREINER, W. 2010. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study.

RCPCH 2019. Safeguarding children and young people: roles and competencies for paediatricians London, UK: Royal College of Paediatrics and Child Health

RCS ENGLAND 2015. The state of children's oral health in England. London: Royal College of Surgeons of England, Faculty of Dental Surgery.

RENSLOW, J. & MAUPIN, J. 2018. Draw-and-write technique elicits children's perceptions of health in the USA and Guatemala. *Health Education Journal,* 77**,** 15-29.

RENTON, T. 2011. Dental (Odontogenic) Pain. *Reviews in Pain,* 5**,** 2-7.

RIBEIRO, G., GOMES, M. C., DE LIMA, K., MARTINS, C., PAIVA, S. M. & GRANVILLE-GARCIA, A. 2015. Work absenteeism by parents because of oral conditions in preschool children. *Int. Dent. J.,* 65**,** 331-337.

RICKETTS, D., LAMONT, T., INNES, N., KIDD, E. & CLARKSON, J. E. 2013. Operative caries management in adults and children. *Cochrane Database of Systematic Reviews,* Art. No.: CD003808.

RISA HOLE, A. 2015. DCREATE: Stata module to create efficient designs for discrete choice experiments. *Statistical Software Components.* Boston College Department of Economics.

RITCHIE, J., RITCHIE, J., LEWIS, J., MCNAUGHTON NICHOLLS, C., ORMSTON, R. & NICHOLLS, C. M. 2014. *Qualitative research practice : a guide for social science students and researchers,* Los Angeles, Los Angeles : SAGE, 2014.

RODD, H. D. & BLANKENSTEIN, R. 1995. Iron-deficiency anaemia in young children requiring dental extractions under general anaesthesia. *International Journal of Paediatric Dentistry,* 5**,** 9-14.

ROGERS, H. J., VERMAIRE, J. H., GILCHRIST, F. & SCHULLER, A. A. 2019. The Relationship between Caries-Specific Quality of Life and Generic Wellbeing in a Dutch Pediatric Population. *Dentistry Journal,* 7**,** 67.

ROLLAND, S., TREASURE, E., BURDEN, D., FULLER, E. & VERNAZZA, C. 2016. The orthodontic condition of children in England, Wales and Northern Ireland 2013. *British Dental Journal,* 221**,** 415-419.

RONEN, G. M., STREINER, D. L., ROSENBAUM, P. & CANADIAN PEDIATRIC EPILEPSY, N. 2003. Health-related Quality of Life in Children with Epilepsy: Development and Validation of Self-report and Parent Proxy Measures. *Epilepsia,* 44**,** 598-612.

ROSENBAUM, P. 1996. Measuring health-related quality of life in pediatric populations. *Quality of life and pharmacoeconomics in clinical trials*, 785-791.

ROUXEL, P. & CHANDOLA, T. 2018. Socioeconomic and ethnic inequalities in oral health among children and adolescents living in England, Wales and Northern Ireland. *Community Dentistry and Oral Epidemiology,* 46**,** 426-434.

ROWEN, D., BRAZIER, J., ARA, R. & AZZABI ZOURAQ, I. 2017. The Role of Condition-Specific Preference-Based Measures in Health Technology Assessment. *PharmacoEconomics,* 35**,** 33-41.

ROWEN, D., BRAZIER, J., KEETHARUTH, A., TSUCHIYA, A. & MUKURIA, C. 2016. Comparison of Modes of Administration and Alternative Formats for Eliciting Societal Preferences for Burden of Illness. *Applied Health Economics and Health Policy,* 14**,** 89-104.

ROWEN, D., BRAZIER, J. & ROBERTS, J. 2009. Mapping SF-36 onto the EQ-5D index: how reliable is the relationship? *Health and Quality of Life Outcomes,* 7**,** 27.

ROWEN, D., BRAZIER, J., TSUCHIYA, A., YOUNG, T. & IBBOTSON, R. 2012. It's All in the Name, or Is It? The Impact of Labeling on Health State Values. *Medical Decision Making,* 32**,** 31-40.

ROWEN, D., BRAZIER, J. & VAN HOUT, B. 2015. A Comparison of Methods for Converting DCE Values onto the Full Health-Dead QALY Scale. *Medical Decision Making,* 35**,** 328-340.

ROWEN, D., BRAZIER, J., YOUNG, T., GAUGRIS, S., CRAIG, B. M., KING, M. T. & VELIKOVA, G. 2011. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value Health,* 14**,** 721-31.

ROWEN, D., LABEIT, A., STEVENS, K., ELLIOTT, J., MULHERN, B., CARLTON, J., BASARIR, H. & BRAZIER, J. 2018a. Estimating a Preference-Based Single Index Measuring the Quality-of-Life Impact of Self-Management for Diabetes. *Medical Decision Making,* 38**,** 699-707.

ROWEN, D., MULHERN, B., STEVENS, K. & VERMAIRE, J. H. 2018b. Estimating a Dutch Value Set for the Paediatric Preference-Based CHU9D Using a Discrete Choice Experiment with Duration. *Value in Health*.

ROWEN, D., RIVERO-ARIAS, O., DEVLIN, N. & RATCLIFFE, J. 2020. Review of Valuation Methods of Preference-Based Measures of Health for Economic Evaluation in Child and Adolescent Populations: Where are We Now and Where are We Going? *PharmacoEconomics,* 38**,** 325-340.

ROZIER, R. G., ADAIR, S., GRAHAM, F., IAFOLLA, T., KINGMAN, A., KOHN, W., KROL, D., LEVY, S., POLLICK, H., WHITFORD, G., STROCK, S., FRANTSVE-HAWLEY, J., ARAVAMUDHAN, K. & MEYER, D. M. 2010. Evidence-based clinical recommendations on the prescription of dietary fluoride supplements for caries prevention: a report of the American Dental Association Council on Scientific Affairs. *J Am Dent Assoc,* 141**,** 1480-9.

RUGG-GUNN, A. J., HACKETT, A. F., APPLETON, D. R., JENKINS, G. N. & EASTOE, J. E. 1984. Relationship between dietary habits and caries increment assessed over two years in 405 English adolescent school children. *Arch Oral Biol,* 29**,** 983-92.

RUGG-GUNN, A. J., HACKETT, A. F., APPLETON, D. R. & MOYNIHAN, P. J. 1986. The dietary intake of added and natural sugars in 405 English adolescents. *Hum Nutr Appl Nutr,* 40**,** 115-24.

RYAN, D., ASHLEY, P. F., PAREKH, S., AL-CHIHABI, M. & PARRY, J. 2010. Sedation for dental treatment of children in the primary care sector (UK). *Br Dent J,* 208**,** E21-E21.

SAKUMA, S., YOSHIHARA, A., MIYAZAKI, H. & KOBAYASHI, S. 2010. Economic Evaluation of a School-based Combined Program with a Targeted Pit and Fissure Sealant and Fluoride Mouth Rinse in Japan. *The Open Dentistry Journal,* 4**,** 230-236.

SAMNALIEV, M., WIJERATNE, R., GRACE KWON, E., OHIOMOBA, H. & WAI NG, M. 2015. Cost-effectiveness of a disease management program for early childhood caries. *Journal of Public Health Dentistry,* 75**,** 24-33.

SANDELOWSKI, M. 1995. Sample size in qualitative research. *Research in Nursing & Health,* 18**,** 179-183.

SANDELOWSKI, M. & GIVEN, L. M. 2008. *The SAGE encyclopedia of qualitative research methods,* Thousand Oaks, California, USA, SAGE Publications.

SANDOR, Z. & WEDEL, M. 2001. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research,* 38**,** 430-444.

SANTAMARIA, R., INNES, N., MACHIULSKIENE, V., EVANS, D. & SPLIETH, C. 2014. Caries management strategies for primary molars: 1-yr randomized control trial results. *Journal of dental research,* 93**,** 1062-1069.

SANTOS PAIS, M. & BISSELL, S. 2006. Overview and implementation of the UN Convention on the Rights of the Child. *The Lancet,* 367**,** 689-690.

SAUNDERS, B., SIM, J., KINGSTONE, T., BAKER, S., WATERFIELD, J., BARTLAM, B., BURROUGHS, H. & JINKS, C. 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity,* 52**,** 1893-1907.

SCHAFER, J. L. 1999. Multiple imputation: a primer. *Statistical Methods in Medical Research,* 8**,** 3-15.

SCHILPZAND, E. J., SCIBERRAS, E., EFRON, D., ANDERSON, V. & NICHOLSON, J. M. 2015. Improving survey response rates from parents in school-based research using a multi-level approach. *PLoS ONE,* 10**,** e0126950.

SCHROTH, R. J., LEVI, J., KLIEWER, E., FRIEL, J. & MOFFATT MICHAEL, E. 2013. Association between iron status, iron deficiency anaemia, and severe early childhood caries: a case–control study. *BMC Pediatrics,* 13**,** 22.

SCHUCH, H. S., DOS SANTOS COSTA, F., TORRIANI, D. D., DEMARCO, F. F. & GOETTEMS, M. L. 2015. Oral health-related quality of life of schoolchildren: impact of clinical and psychosocial variables. *International Journal of Paediatric Dentistry,* 25**,** 358-365.

SCHULLER, A. A., VERMAIRE, J. H., VAN KEMPEN, C. P. F., VAN DOMMELEN, P. & VERRIPS, G. H. W. 2018. *Kies voor Tanden [a survey of oral health and oral health behaviour in children and adolescents],* Leiden, The Netherlands, TNO.

SCHWENDICKE, F., DÖRFER, C. E., SCHLATTMANN, P., PAGE, L. F., THOMSON, W. M. & PARIS, S. 2015a. Socioeconomic Inequality and Caries: A Systematic Review and Meta-Analysis. Los Angeles, CA.

SCHWENDICKE, F., LAMONT, T. & INNES, N. 2015b. Outcomes in Trials for Management of Caries Lesions (OuTMaC): protocol. *Trials,* 16**,** 397.

SCHWENDICKE, F., STOLPE, M., MEYER-LUECKEL, H., PARIS, S. & DORFER, C. E. 2013. Cost-effectiveness of one- and two-step incomplete and complete excavations. *Journal of Dental Research,* 92**,** 880-7.

SCHWENDICKE, F., THOMSON, W. M., BROADBENT, J. M. & STOLPE, M. 2016. Effects of taxing sugar-sweetened beverages on caries and treatment costs. *Journal of Dental Research,* 95**,** 1327-1332.

SCOTTISH EXECUTIVE 2005. Improving Oral Health and Modernising NHS Dental Services in Scotland. Edinburgh.

SCOTTISH GOVERNMENT. 2020. *Scottish Index of Multiple Deprivation* [Online]. Available: https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020/ [Accessed 25th May 2020].

SDCEP 2010. Prevention and management of dental caries in children: dental clinical guidance. *In:* PROGRAMME, S. D. C. E. (ed.). Dundee.

SDCEP 2018. Restricting the Use of Dental Amalgam in Specific Patient Groups: Implementation Advice for Article 10(2) of Regulation (EU) 2017/852 on Mercury. Scotland.

SEVERENS, J. L. & MILNE, R. J. 2004. Discounting health outcomes in economic evaluation: the ongoing debate. *Value Health,* 7**,** 397-401.

SHAHNAVAZ, S., HEDMAN-LAGERLÖF, E., HASSELBLAD, T., REUTERSKIÖLD, L., KALDO, V. & DAHLLÖF, G. 2018. Internet-Based Cognitive Behavioral Therapy for Children and Adolescents With Dental Anxiety: Open Trial. *Journal of Medical Internet Research,* 20**,** e12.

SHAHNAVAZ, S., HEDMAN, E., GRINDEFJORD, M., REUTERSKIÖLD, L. & DAHLLÖF, G. 2016. Cognitive Behavioral Therapy for Children with Dental Anxiety: A Randomized Controlled Trial. *JDR Clinical & Translational Research,* 1**,** 234-243.

SHAOUL, R., GAITINI, L., KHAROUBA, J., DARAWSHI, G., MAOR, I. & SOMRI, M. 2012. The association of childhood iron deficiency anaemia with severe dental caries. *Acta Paediatrica,* 101**,** e76-e79.

SHAW, J. W. 2011. Use of Patient Versus Population Preferences in Economic Evaluations of Health Care Interventions.

SHAW, R. J. 2001. Treatment Adherence in Adolescents: Development and Psychopathology. *Clinical Child Psychology and Psychiatry,* 6**,** 137-150.

SHEIHAM, A. 2006. Dental caries affects body weight, growth and quality of life in pre-school children. *British Dental Journal,* 201**,** 625-6.

SHEMILT, I., MUGFORD, M., BYFORD, S., DRUMMOND, M., EISENSTEIN, E., KNAPP, M., MALLENDER, J., MCDAID, D., VALE, L. & WALKER, D. 2008. Incorporating Economics Evidence. *In:* GREEN, J. P. H. A. S. (ed.) *Cochrane Handbook for Systematic Reviews of Interventions.*

SHEPHERD, C. C. J., LI, J. & ZUBRICK, S. R. 2012. Socioeconomic disparities in physical health among Aboriginal and Torres Strait Islander children in Western Australia. *Ethnicity & Health,* 17**,** 439-461.

SHEPHERD, M. A., NADANOVSKY, P. & SHEIHAM, A. 1999. The prevalence and impact of dental pain in 8-year-old school children in Harrow, England. *British Dental Journal,* 187**,** 38-41.

SINGH, J., LORD, J., LONGWORTH, L., ORR, S., MCGARRY, T., SHELDON, R. & BUXTON, M. 2012. Does Responsibility Affect the Public's Valuation of Health Care

Interventions? A Relative Valuation Approach to Health Care Safety. *Value in Health,* 15**,** 690-698.

SKOLD, U. M., PETERSSON, L. G., BIRKHED, D. & NORLUND, A. 2008. Cost-analysis of school-based fluoride varnish and fluoride rinsing programs. *Acta Odontologica Scandinavica,* 66**,** 286-92.

SLADE, G. D. & SANDERS, A. E. 2017. Two decades of persisting income-disparities in dental caries among U.S. children and adolescents. *Journal of Public Health Dentistry*.

SOARES, F., SOUTO, G., LOFRANO, M. & COLARES, V. 2015. Anxiety related to dental care in children and adolescents in a low-income Brazilian community. *European Archives of Paediatric Dentistry,* 16**,** 149-152.

SOUTHALL, D., BURR, S., SMITH, R. & BULL, D. 2000. The child-friendly healthcare initiative (CFHI): Healthcare provision in accordance with the UN convention on the rights of the child. *Pediatrics,* 106**,** 1054-64.

STATS WALES. 2019. *Welsh Index of Multiple Deprivation* [Online]. Available: https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation [Accessed 25th May 2020].

STEARNS, S. C., ROZIER, R. G., KRANZ, A. M., PAHEL, B. T. & QUINONEZ, R. B. 2012. Cost-effectiveness of preventive oral health care in medical offices for young Medicaid enrollees. *Archives of Pediatrics & Adolescent Medicine,* 166**,** 945-51.

STEINBERG, L. 2005. Cognitive and affective development in adolescence. *Trends in cognitive sciences,* 9**,** 69-74.

STEPHENSON, J., CHADWICK, B., PLAYLE, R. & TREASURE, E. 2010. A Competing Risk Survival Analysis Model to Assess the Efficacy of Filling Carious Primary Teeth. *Caries Research,* 44**,** 285-93.

STEVENS, K. 2009. *Developing a descriptive system for a new preference based measure of health related quality of life for children*, Quality of Life Research.

STEVENS, K. 2012. Valuation of the Child Health Utility 9D Index. *PharmacoEconomics,* 30**,** 729-747.

STEVENS, K. 2015. "Because that's what matters to me". A pilot study to test the feasibility and reliability of ordinal valuation methods for health state valuation with children. *Health Economics and Decision Science Discussion Paper Series (University of Sheffield),* 15.

STEVENS, K. & PALFREYMAN, S. 2011. *The use of qualitative methods in developing the descriptive systems of preference based measures of health related quality of life for use in economic evaluation*.

STEVENS, K. J. 2010. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. *Qual Health Res,* 20**,** 340-51.

STEVENS, K. J., BRAZIER, J. E., MCKENNA, S. P., DOWARD, L. C. & CORK, M. J. 2005. The development of a preference-based measure of health in children with atopic dermatitis. *British Journal of Dermatology,* 153**,** 372-377.

STURGIS, P., ROBERTS, C. & SMITH, P. 2014. Middle Alternatives Revisited: How the neither/nor Response Acts as a Way of Saying "I Don't Know"? *Sociological Methods & Research,* 43**,** 15-38.

SUNG, L. & REGIER, D. A. 2013. Decision making in pediatric oncology: Evaluation and incorporation of patient and parent preferences. Hoboken.

SUTTON, E. & COAST, J. 2014. Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods. *Palliative Medicine,* 28**,** 151-157.

TAGLIAFERRO, E., MARINHO, D., PEREIRA, C., PARDI, V., AMBROSANO, G., MENEGHIM, M. C. & PEREIRA, A. 2013. [Cost-effectiveness analysis of preventive methods for occlusal surface according to caries risk: results of a controlled clinical trial]. *Cadernos de saúde pública* [Online], 29 Suppl 1. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/405/CN-01067405/frame.html.

TENNANT, A. & CONAGHAN, P. G. 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research,* 57**,** 1358-1362.

TENNANT, A. & PALLANT, J. 2007. DIF matters: A practical approach to test if Differential Item Functioning makes a difference. *Rasch Measurement Transactions,* 20**,** 1082-1084.

TERVONEN, T., SCHMIDT-OTT, T., MARSH, K., BRIDGES, J. F. P., QUAIFE, M. & JANSSEN, E. 2018. Assessing Rationality in Discrete Choice Experiments in Health: An Investigation into the Use of Dominance Tests. *Value in Health,* 21**,** 1192-1197.

TERWEE, C. B., BOT, S. D., DE BOER, M. R., VAN DER WINDT, D. A., KNOL, D. L., DEKKER, J., BOUTER, L. M. & DE VET, H. C. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology,* 60**,** 34-42.

TESIO, L. 2003. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *J Rehabil Med,* 35**,** 105-15.

THRELFALL, A., HUNT, C., MILSOM, K., TICKLE, M. & BLINKHORN, A. 2007. Exploring factors that influence general dental practitioners when providing advice to help prevent caries in children. *British Dental Journal,* 202**,** E10-E10.

THRELFALL, A., PILKINGTON, L., MILSOM, K., BLINKHORN, A. & TICKLE, M. 2005. General dental practitioners' views on the use of stainless steel crowns to restore primary molars. *British Dental Journal,* 199**,** 453-5.

THURSTONE, L. L. 1927. A law of comparative judgment. *Psychological Review,* 34**,** 273-286.

TICKLE, M., MILSOM, K. & KENNEDY, A. 1999. Is it better to leave or restore carious deciduous molar teeth? A preliminary study. *Prim Dent Care,* 6**,** 127-31.

TICKLE, M., MILSOM, K., KING, D., KEARNEY-MITCHELL, P. & BLINKHORN, A. 2002. The fate of the carious primary teeth of children who regularly attend the general dental service. *British Dental Journal,* 192**,** 219.

TICKLE, M., MILSOM, K. M., DONALDSON, M., KILLOUGH, S., O'NEILL, C., CREALEY, G., SUTTON, M., NOBLE, S., GREER, M. & WORTHINGTON, H. V. 2011. Protocol for Northern Ireland Caries Prevention in Practice Trial (NIC-PIP) trial: a randomised controlled trial to measure the effects and costs of a dental caries prevention regime for young children attending primary care dental services. *BMC oral health,* 11**,** 27.

TICKLE M, O. N. C., DONALDSON M, BIRCH S, NOBLE S, KILLOUGH S, MURPHY L, GREER M, BRODISON J, VERGHIS R, WORTHINGTON HV 2016. A randomised controlled trial to measure the effects and costs of a dental caries prevention regime for young children attending primary care dental services: the Northern Ireland Caries Prevention In Practice (NIC-PIP) trial. *NIHR Health Technology Assessment programme,* 20.

TIMMS, L., SUMNER, O., DEERY, C. & ROGERS, H. J. 2020. Everyone else is using it, so why isn't the UK? Silver diamine fluoride for children and young people. *Community Dent Health,* 37**,** 143-149.

TOLLEY, K. 2009. What are health utilities. *London: Hayward Medical Communications*.

TONG, A., SAINSBURY, P. & CRAIG, J. 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care,* 19**,** 349-357.

TONMUKAYAKUL, U. & ARROW, P. 2017. Cost-effectiveness analysis of the atraumatic restorative treatment-based approach to managing early childhood caries. *Community Dentistry and Oral Epidemiology,* 45**,** 92-100.

TONMUKAYAKUL, U., CALACHE, H., CLARK, R., WASIAK, J. & FAGGION, C. M. 2015. Systematic Review and Quality Appraisal of Economic Evaluation Publications in Dentistry. *Journal of dental research,* 94**,** 1348-1354.

TORGERSON, D. J. & RAFTERY, J. 1999. Discounting. *BMJ,* 319**,** 914-915.

TORRANCE, G. W. 1987. Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases,* 40**,** 593-600.

TORRANCE, W. G., FEENY, H. D., FURLONG, J. W., BARR, D. R., ZHANG, D. Y. & WANG, D. Q. 1996. Multiattribute Utility Function for a Comprehensive Health Status Classification System: Health Utilities Index Mark 2. *Medical Care,* 34**,** 702-722.

TSAKOS, G., HILL, K., CHADWICK, B. & T., A. 2015. Children's dental health survey 2013, Report 1: Attitudes, behaviours and children's dental health, England, Wales and Northern Ireland.

TWETMAN, S. 2018. Prevention of dental caries as a non-communicable disease. *European Journal of Oral Sciences,* 126**,** 19-25.

TWISS, J., MCKENNA, S. P., GRAHAM, J., SWETZ, K., SLOAN, J. & GOMBERG-MAITLAND, M. 2016. Applying Rasch analysis to evaluate measurement equivalence of different administration formats of the Activity Limitation scale of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR). *Health and quality of life outcomes,* 14**,** 57-57.

TYLEE, A., HALLER, D. M., GRAHAM, T., CHURCHILL, R. & SANCI, L. A. 2007. Youth-friendly primary-care services: how are we doing and what more needs to be done? *The Lancet,* 369**,** 1565-1573.

UK ASSOCIATION FOR ACCESSIBLE FORMATS 2012. Creating clear print and large print documents: Guidance from UKAAF. *Ref G003.* UK: UKAAF.

UK HEALTH FORUM & CANCER RESEARCH UK 2016. Short and sweet: Why the Government should introduce a sugary drinks tax. UK Health Forum and Cancer Research UK.

UNGAR, W. 2011. Challenges in Health State Valuation in Paediatric Economic Evaluation. *PharmacoEconomics,* 29**,** 641-652.

UNGAR, W. J., BOYDELL, K., DELL, S., FELDMAN, B. M., MARSHALL, D., WILLAN, A. & WRIGHT, J. G. 2012. A Parent-Child Dyad Approach to the Assessment of Health Status and Health-Related Quality of Life in Children with Asthma. *PharmacoEconomics,* 30**,** 697-712.

UNITED NATIONS 2009. General Comment No. 12: The right of the child to be heard (No.UN=CRC=C=GC=12). Geneva, Switzerland: United Nations.

VANDEVELDE, S., VAN KEER, H., SCHELLINGS, G. & VAN HOUT-WOLTERS, B. 2015. Using think-aloud protocol analysis to gain in-depth insights into upper primary school children's self-regulated learning. *Learning and Individual Differences,* 43**,** 11-30.

VANNIYASINGAM, T., CUNNINGHAM, C. E., FOSTER, G. & THABANE, L. 2016. Simulation study to determine the impact of different design features on design efficiency in discrete choice experiments. *BMJ Open,* 6.

VERMAIRE, J., LOVEREN, C., BROUWER, W. & KROL, M. 2014. Value for money: economic evaluation of two different caries prevention programmes compared with standard care in a randomized controlled trial. *Caries research* [Online], 48. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/533/CN-01043533/frame.html.

VERNAZZA, C. R., TAYLOR, G., DONALDSON, C., GRAY, J., HOLMES, R., CARR, K. & EXLEY, C. 2019. How does priority setting for resource allocation happen in commissioning dental services in a nationally led, regionally delivered system: a qualitative study using semistructured interviews with NHS England dental commissioners. *BMJ open,* 9**,** e024995-e024995.

VERSTEEGH, M. M. & BROUWER, W. B. F. 2016. Patient and general public preferences for health states: A call to reconsider current guidelines. *Social Science & Medicine,* 165**,** 66-74.

VERSTEEGH, M. M., LEUNIS, A., UYL-DE GROOT, C. A. & STOLK, E. A. 2012. Condition-Specific Preference-Based Measures: Benefit or Burden? *Value in Health,* 15**,** 504-513.

VETTER, T. R., BRIDGEWATER, C. L. & MCGWIN, G. 2012. An observational study of patient versus parental perceptions of health-related quality of life in children and adolescents with a chronic pain condition: who should the clinician believe? *Health and Quality of Life Outcomes,* 10**,** 85.

WALKER, D. G. 2012. *Best practices for conducting economic evaluations in health care: a systematic review of quality assessment tools,* Rockville, MD, Rockville, MD Agency for Healthcare Research and Quality.

WALSH, T., WORTHINGTON, H. V., GLENNY, A. M., APPELBE, P., MARINHO, V. C. & SHI, X. 2010. Fluoride toothpastes of different concentrations for preventing dental caries in children and adolescents. *Cochrane Database Syst Rev***,** Cd007868.

WATSON, V., PORTEOUS, T., BOLT, T. & RYAN, M. 2019. Mode and Frame Matter: Assessing the Impact of Survey Mode and Sample Frame in Choice Experiments. *Medical Decision Making,* 39**,** 827-841.

WATT, R. G. 2012. Social determinants of oral health inequalities: implications for action. *Community Dentistry and Oral Epidemiology,* 40**,** 44-48.

WEERHEIJM, K. 2003. Molar Incisor Hypomineralisation (MIH). *European Journal of Paediatric Dentistry,* 4**,** 115-120.

WEINSTEIN, P., HARRISON, R. & BENTON, T. 2004. Motivating parents to prevent caries in their young children: One-year findings: One-year findings. *The Journal of the American Dental Association,* 135**,** 731-738.

WEINTRAUB, J. A., STEARNS, S. C., ROZIER, R. G. & HUANG, C. 2001. Treatment outcomes and costs of dental sealants among children enrolled in Medicaid. *American Journal of Public Health,* 91**,** 1877-1881.

WELSH GOVERNMENT 2013. Together for Health: A National Oral Health Plan for Wales 2013-18. Wales: Welsh Government.

WEST, N., OAKES, L., HINDS, P. S., SANDERS, L., HOLDEN, R., WILLIAMS, S., FAIRCLOUGH, D. & BOZEMAN, P. 1994. Measuring pain in pediatric oncology ICU patients. *J Pediatr Oncol Nurs,* 11**,** 64-8; discussion 69-70.

WHITE, D. A., TSAKOS, G., PITTS, N. B., FULLER, E., DOUGLAS, G. V. A., MURRAY, J. J. & STEELE, J. G. 2012. Adult Dental Health Survey 2009: common oral health conditions and their impact on the population. *BDJ,* 213**,** 567.

WHITTY, J., WALKER, R., GOLENKO, X. & RATCLIFFE, J. 2014. A Think Aloud Study Comparing the Validity and Acceptability of Discrete Choice and Best Worst Scaling Methods. *PLoS One,* 9.

WHITTY, J. A. & OLIVEIRA GONÇALVES, A. S. 2018. A Systematic Review Comparing the Acceptability, Validity and Concordance of Discrete Choice Experiments and Best–Worst Scaling for Eliciting Preferences in Healthcare. *The Patient - Patient-Centered Outcomes Research,* 11**,** 301-317.

WIEDEL, A., NORLUND, A., PETRÉN, S. & BONDEMARK, L. 2016. A cost minimization analysis of early correction of anterior crossbite-a randomized controlled trial. *European journal of orthodontics* [Online], 38. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/280/CN-01215280/frame.html.

WILKINSON, D. 2012. Dissent about assent in paediatric research. *Journal of Medical Ethics,* 38**,** 2.

WILLE, N., BADIA, X., BONSEL, G., BURSTRÖM, K., CAVRINI, G., DEVLIN, N., EGMAR, A.-C., GREINER, W., GUSI, N., HERDMAN, M., JELSMA, J., KIND, P., SCALONE, L. & RAVENS-SIEBERER, U. 2010. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Quality of Life Research,* 19**,** 875-886.

WILLIS, G. B. 2005. *Cognitive interviewing [electronic resource] : a tool for improving questionnaire design,* Thousand Oaks, California, USA, SAGE Publications.

WILSON, I. B. & CLEARY, P. D. 1995. Linking Clinical Variables With Health-Related Quality of Life: A Conceptual Model of Patient Outcomes. *JAMA,* 273**,** 59-65.

WOLSTENHOLME, J. L., BARGO, D., WANG, K., HARNDEN, A., RÄISÄNEN, U. & ABEL, L. 2018. Preference-based measures to obtain health state utility values for use in economic evaluations with child-based populations: a review and UK-based focus group assessment of patient and parent choices. *Quality of Life Research,* 27**,** 1769-1780.

WORLD HEALTH OGANISATION 2011. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity (WHO/NMH/NHD/MNM/11.1).

Vitamin and Mineral Nutrition Information System. . Geneva: World Health Organization.

WORLD HEALTH ORGANISATION 2002. International Classification of Functioning, Disability and Health Geneva: WHO.

WORLD HEALTH ORGANISATION. 2005. *Constitution of the World Health Organization: Principles* [Online]. Geneva. Available: http://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf?ua=1 [Accessed September 2018].

WORLD HEALTH ORGANISATION 2013. Oral Health Surveys: Basic Methods. Switzerland: WHO.

WU, Y., WANG, J. & MAO, Z. 2002. [Cost-minimization analysis of two methods during the prevention of dental fear during caries filling treatments]. *Chung-Hua Liu Hsing Ping Hsueh Tsa Chih Chinese Journal of Epidemiology,* 23**,** 387-90.

YAWARY, R., ANTHONAPPA, R. P., EKAMBARAM, M., MCGRATH, C. & KING, N. M. 2016. Changes in the oral health-related quality of life in children following comprehensive oral rehabilitation under general anaesthesia. *International Journal of Paediatric Dentistry,* 26**,** 322-329.

YEE, R., MCDONALD, N. & WALKER, D. 2004. A cost-benefit analysis of an advocacy project to fluoridate toothpastes in Nepal. *Community Dental Health,* 21**,** 265-270.

YORK HEALTH ECONOMICS CONSORTIUM 2016. A rapid review of evidence on the cost-effectiveness of interventions to improve the oral health of children aged 0-5 years. London.

YOUNG, T., YANG, Y., BRAZIER, J., TSUCHIYA, A. & COYNE, K. 2009. The first stage of developing preference-based measures: constructing a health-state classification using Rasch analysis. *Qual. Life Res.,* 18**,** 253-265.

YOUNG, T. A., YANG, Y., BRAZIER, J. E. & TSUCHIYA, A. 2011. The Use of Rasch Analysis in Reducing a Large Condition-Specific Instrument for Preference Valuation: The Case of Moving from AQLQ to AQL-5D. *Medical Decision Making,* 31**,** 195-210.

YURGELUN-TODD, D. 2007. Emotional and cognitive changes during adolescence. *Current opinion in neurobiology,* 17**,** 251-257.

ZABOS, G. P., GLIED, S. A., TOBIN, J. N., AMATO, E., TURGEON, L., MOOTABAR, R. N. & NOLON, A. K. 2002. Cost-effectiveness analysis of a school-based dental sealant program for low-socioeconomic-status children: a practice-based report. *Journal of Health Care for the Poor & Underserved,* 13**,** 38-48.

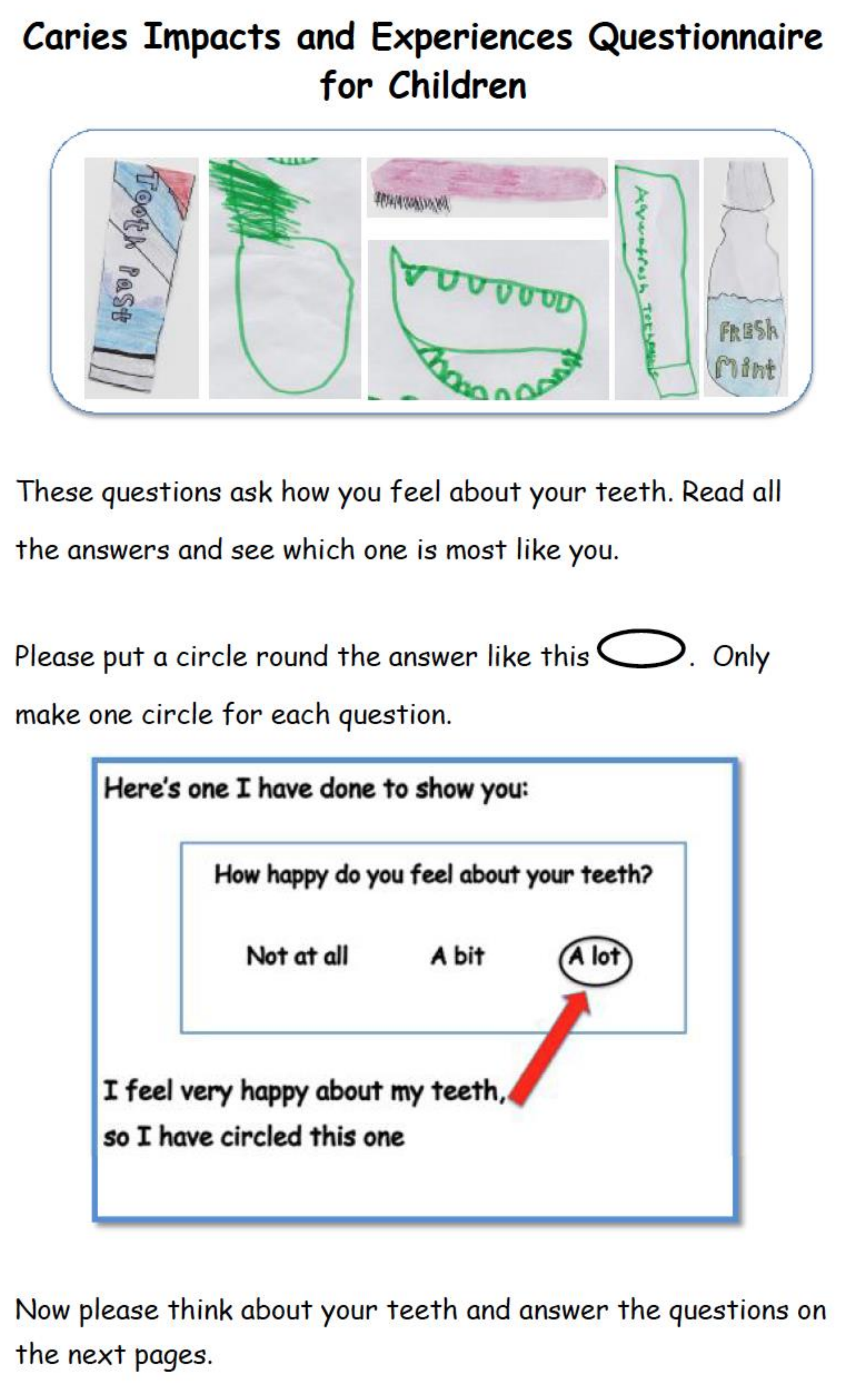
ŽALOUDÍKOVÁ, I. 2010. Children's conceptions of health, illness, death and the anatomy of the human body. *School and Health,* 21**,** 123-140.

ZHOU, H., ROBERTS, P. & HORGAN, L. 2008. Association between self-report pain ratings of child and parent, child and nurse and parent and nurse dyads: meta-analysis. Oxford, UK.

ZWERINA, K., HUBER, J. & KUHFELD, W. F. 1996. A general method for constructing efficient choice designs. *Durham, NC: Fuqua School of Business, Duke University*.

Appendices

A. CARIES-QC



Caries Impacts and Experiences Questionnaire for Children

These questions ask how you feel about your teeth. Read all the answers and see which one is most like you.

Please put a circle round the answer like this ◯. Only make one circle for each question.

Here's one I have done to show you:

How happy do you feel about your teeth?

Not at all **A bit** [✓ **A lot**]

I feel very happy about my teeth, so I have circled this one

Now please think about your teeth and answer the questions on the next pages.

Please **circle one answer** for each question.

1. How much do your teeth hurt you?

Not at all　A bit　A lot

2. Do your teeth make it hard to eat some foods?

Not at all　A bit　A lot

3. Do you have to eat on one side of your mouth because of your teeth?

Not at all　A bit　A lot

4. Do you get food stuck in your teeth?

Not at all　A bit　A lot

5. How much do you get kept awake by your teeth?

Not at all　A bit　A lot

6. How much do your teeth annoy you?

Not at all　A bit　A lot

Please go to the next page.

Please **circle one answer** for each question.

7. How much do your teeth hurt when you brush them?

Not at all A bit A lot

8. Do you have to eat more carefully because of your teeth?

Not at all A bit A lot

9. Do you have to eat more slowly because of your teeth?

Not at all A bit A lot

10. Do you feel cross because of your teeth?

Not at all A bit A lot

11. How much have you cried because of your teeth?

Not at all A bit A lot

Please go to the next page.

Please **circle one answer** for each question.

12. Do your teeth make it hard to do your schoolwork?

Not at all A bit A lot

13. How much of a problem are your teeth for you?

Not at all A bit A lot

Thank you for answering the questions!

B. Classification system for CHU9D

Dimension	Severity	Wording
Worry	0	I don’t feel worried today
	1	I feel a little bit worried today
	2	I feel a bit worried today
	3	I feel quite worried today
	4	I feel very worried today
Sad	0	I don’t feel sad today
	1	I feel a little bit sad today
	2	I feel a bit sad today
	3	I feel quite sad today
	4	I feel very sad today
Pain	0	I don’t have any pain today
	1	I have a little bit of pain today
	2	I have a bit of pain today
	3	I have quite a lot of pain today
	4	I have a lot of pain today
Tired	0	I don’t feel tired today
	1	I feel a little bit tired today
	2	I feel a bit tired today
	3	I feel quite tired today
	4	I feel very tired today
Annoyed	0	I don’t feel annoyed today
	1	I feel a little bit annoyed today

	2	I feel a bit annoyed today
	3	I feel quite annoyed today
	4	I feel very annoyed today
School work/homework	0	I have no problems with my school work/homework today
	1	I have a few problems with my school work/homework today
	2	I have some problems with my school work/homework today
	3	I have many problems with my school work/homework today
	4	I can’t do my schoolwork/homework today
Sleep	0	Last night I had no problems sleeping
	1	Last night I had a few problems sleeping
	2	Last night I had some problems sleeping
	3	Last night I had many problems sleeping
	4	Last night I couldn’t sleep at all
Daily routine	0	I have no problems with my daily routine today
	1	I have a few problems with my daily routine today
	2	I have some problems with my daily routine today
	3	I have many problems with my daily routine today
	4	I can’t do my daily routine today
Able to join in activities	0	I can join in with any activities today
	1	I can join in with most activities today
	2	I can join in with some activities today
	3	I can join in with a few activities today
	4	I can join in with no activities today

C. Search strategy

Medline

1. "Dental Care for Children"/ec [economics]

2. Health Education, Dental/ec [economics]

3. Oral health/ec [economics]

4. (Dent* or Teeth or "oral health" or "oral surgery" or orthodont* or "fluoride" or "hypomineralisation" or "hypomineralization" or "caries").ti.

5. 1 or 2 or 4

6. Child/

7. Adolescent/ or Infant/

8. (child* or p?ediatric* or infan* or adolescen* or teen*).ti.

9. 6 or 7 or 8

10. 5 and 9

11. Economics/

12. exp "costs and cost analysis"/

13. Economics, Dental/

14. exp economics, hospital/

15. (economic$ or cost or costs or costly or costing or price or prices or pricing).ti.

16. (expenditure$ not energy).ti.

17. "value for money".ti.

18. budget$.ti.

19. or/11-18

20. ((energy or oxygen) adj cost).ti,ab.

21. (metabolic adj cost).ti,ab.

22. ((energy or oxygen) adj expenditure).ti,ab.

23. or/20-22

24. 19 not 23

25. letter.pt.

26. editorial.pt.

27. historical article.pt.

28. or/25-27

29. 24 not 28

30. exp animals/ not humans/

31. 29 not 30

32. bmj.jn.

33. "cochrane database of systematic reviews".jn.

34. health technology assessment winchester england.jn.

35. or/32-34

36. 31 not 35

37. 10 and 36

38. (model or models or modelling or markov or decision or analytic or lifetime or horizon).mp. [mp=tx, bt, ti, ab, ct, hw, ar, ax, bx, cx, ft, fd, ad, jh, jx, ca, pc, ot, nm, fx, kf, ox, px, rx, ui, sy, tc, id, tm, mh]

39. 37 and 38

40. (socioeconomic or socio-economic or ecology or ecological).ti.

41. 39 not 40

Econlit

1. (dent* or Teeth or gum* or tooth or oral care or oral hygiene or fluoride or orthodont* or oral surgery or caries).ti,ab.

2. (child* or p?ediatric* or infan* or adolescen* or teen* or school*).ti,ab.

3. economics.ti,ab.

4. (costs and cost analysis).mp.

5. (economics and medical).mp. [mp=heading words, abstract, title, country as subject]

6. (economics and hospital).mp. [mp=heading words, abstract, title, country as subject]

7. (economics and dental).mp. [mp=heading words, abstract, title, country as subject]

8. (economics and pharmaceutical).mp. [mp=heading words, abstract, title, country as subject]

9. (economics and nursing).mp. [mp=heading words, abstract, title, country as subject]

10. (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.

11. (expenditure$ not energy).ti,ab.

12. value for money.ti,ab.

13. budget$.ti,ab.

14. or/3-13

15. ((energy or oxygen) adj cost).ti,ab.

16. (metabolic adj cost).ti,ab.

17. ((energy or oxygen) adj expenditure).ti,ab.

18. or/15-17

19. 14 not 18

20. 1 and 2 and 19

D. Drummond 10-point checklist

From Drummond M et al. Methods for the economic evaluation of health care programmes. 2nd ed. Oxford. Oxford University Press. 1997

1. Was a well-defined question posed in answerable form?

1.1. Did the study examine both costs and effects of the service(s) or programme(s)?
1.2. Did the study involve a comparison of alternatives?
1.3. Was a viewpoint for the analysis stated and was the study placed in any particular decision-making context?

2. Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?

2.1. Were there any important alternatives omitted?
2.2. Was (should) a do-nothing alternative be considered?

3. Was the effectiveness of the programme or services established?

3.1. Was this done through a randomised, controlled clinical trial? If so, did the trial protocol reflect what would happen in regular practice?
3.2. Was effectiveness established through an overview of clinical studies?
3.3. Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?

4. Were all the important and relevant costs and outcomes for each alternative identified?

4.1. Was the range wide enough for the research question at hand?
4.2. Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of patients and third-party payers. Other viewpoints may also be relevant depending upon the particular analysis.)
4.3. Were the capital costs, as well as operating costs, included?

5a. Were outcomes measured accurately in appropriate physical units (e.g. gained life years)?

5b. Were costs measured accurately in appropriate physical units (e.g. hours of nursing time, number of physician visits)?

5.1. Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?
5.2. Were there any special circumstances (e.g., joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?

6a. Were the outcomes valued credibly?

6b. Were the costs valued credibly?

6.1. Were the sources of all values clearly identified? (Possible sources include market values, patient or client preferences and views, policy-makers' views and health professionals' judgements)
6.2. Were market values employed for changes involving resources gained or depleted?
6.3. Where market values were absent (e.g. volunteer labour), or market values did not reflect actual values (such as clinic space donated at a reduced rate), were adjustments made to approximate market values?
6.4. Was the valuation of outcomes appropriate for the question posed (i.e. has the appropriate type or types of analysis – cost-effectiveness, cost-benefit, cost-utility – been selected)?

7a. Were outcomes adjusted for differential timing?

7b. Were costs adjusted for differential timing?

7.1. Were costs and outcomes that occur in the future 'discounted' to their present values?
7.2. Was there any justification given for the discount rate used?

8. Was an incremental analysis of the outcomes and costs of alternatives performed?

8.1. Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits, or utilities generated?

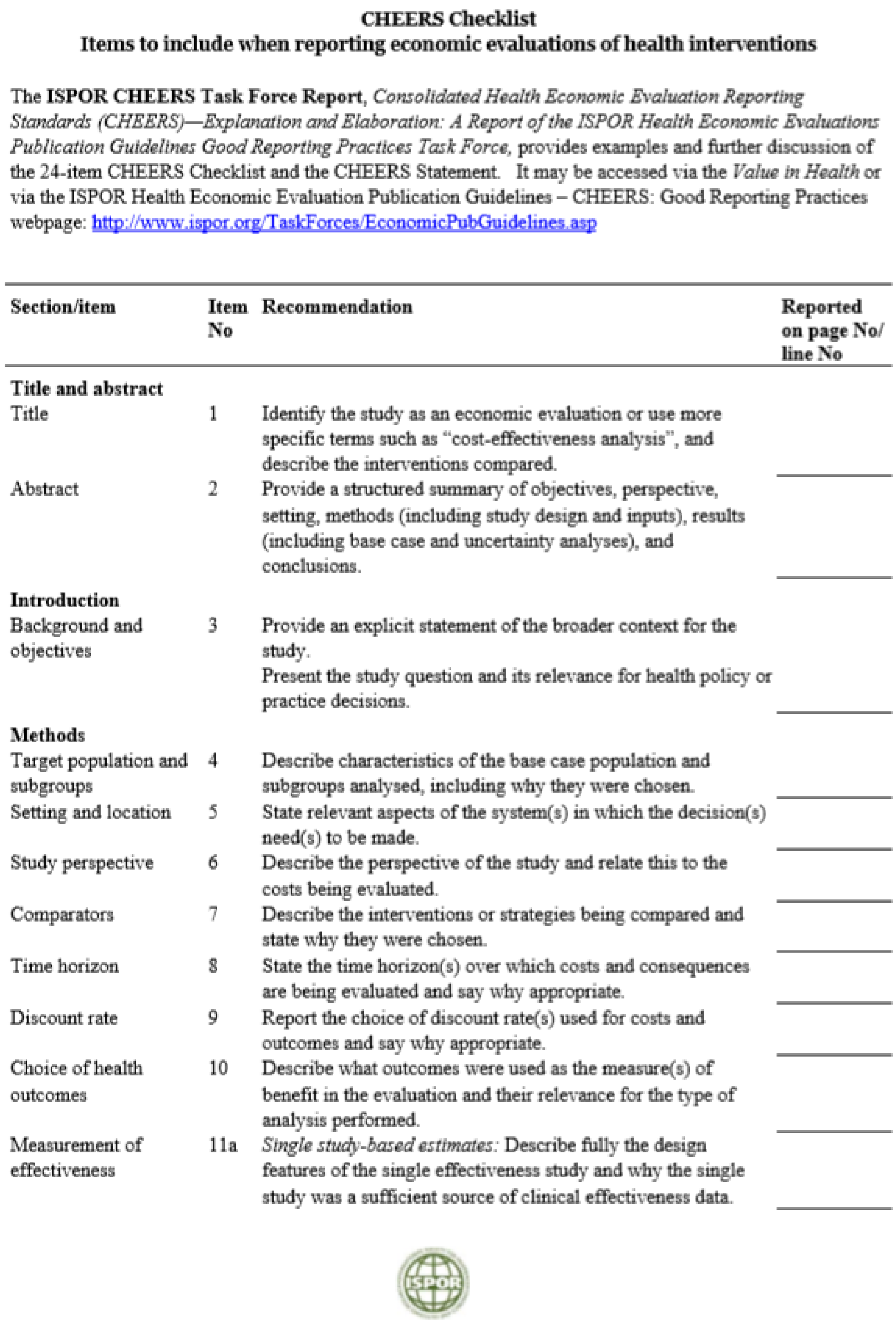
9. Was allowance made for uncertainty in the estimates of costs and outcomes?

9.1. If data on costs and outcomes were stochastic (randomly determined sequence of observations), were appropriate statistical analyses performed?
9.2. If a sensitivity analysis was employed, was justification provided for the range of values (or for key study parameters)?
9.3. Were the study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to outcomes)?

10. Did the presentation and discussion of study results include all issues of concern to users?

10.1. Were the conclusions of the analysis based on some overall index or ratio of costs to outcomes (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?
10.2. Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?
10.3. Did the study discuss the generalisability of the results to other settings and patient/client groups?
10.4. Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and outcomes, or relevant ethical issues)?
10.5. Did the study discuss issues of implementation, such as the feasibility of adopting the ‘preferred’ programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?

E. CHEERS checklist

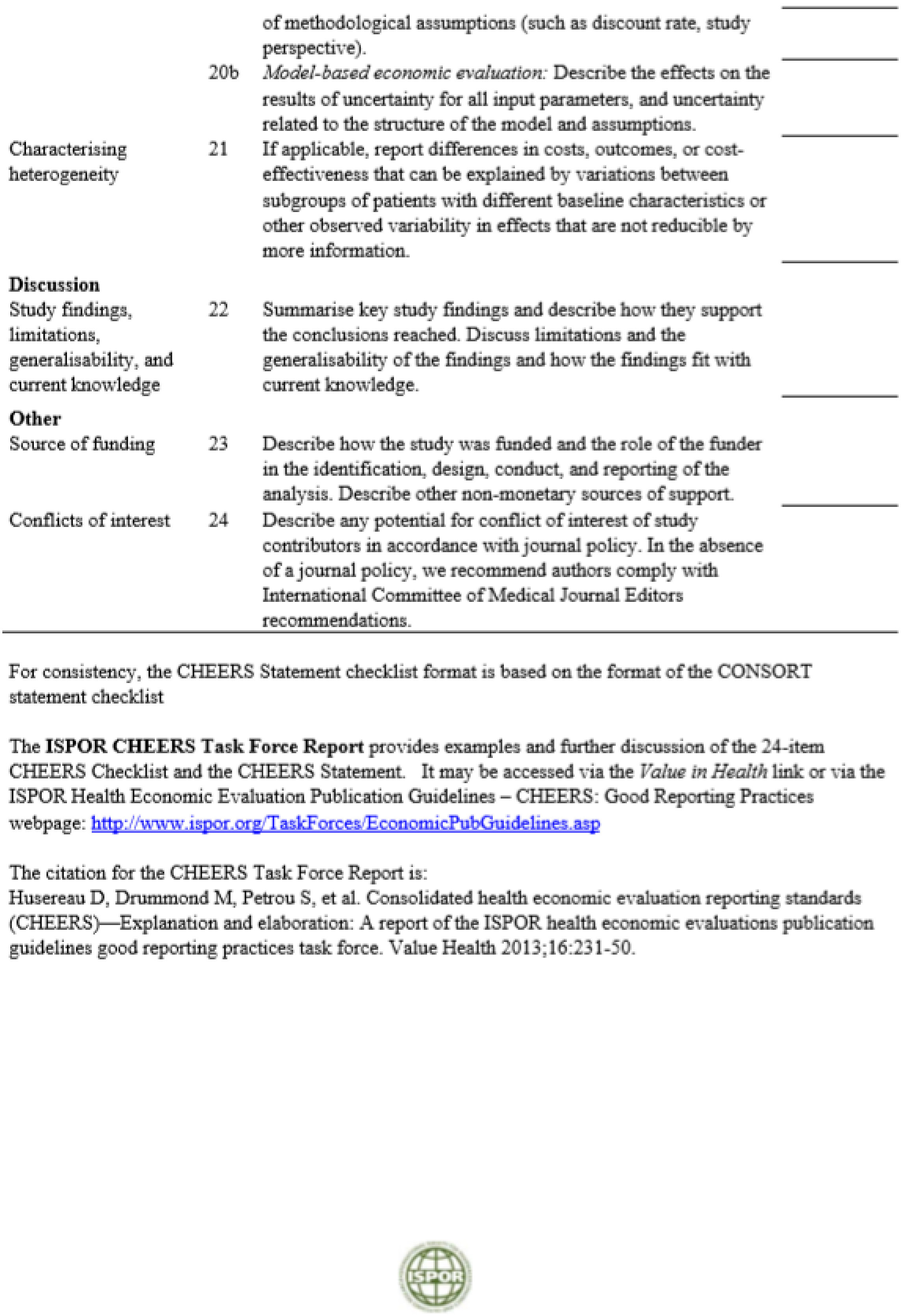


CHEERS Checklist
Items to include when reporting economic evaluations of health interventions

The **ISPOR CHEERS Task Force Report**, *Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluations Publication Guidelines Good Reporting Practices Task Force*, provides examples and further discussion of the 24-item CHEERS Checklist and the CHEERS Statement. It may be accessed via the *Value in Health* or via the ISPOR Health Economic Evaluation Publication Guidelines – CHEERS: Good Reporting Practices webpage: http://www.ispor.org/TaskForces/EconomicPubGuidelines.asp

Section/item	**Item No**	**Recommendation**	**Reported on page No/ line No**
Title and abstract			
Title	1	Identify the study as an economic evaluation or use more specific terms such as "cost-effectiveness analysis", and describe the interventions compared.	
Abstract	2	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.	
Introduction			
Background and objectives	3	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for health policy or practice decisions.	
Methods			
Target population and subgroups	4	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.	
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	
Choice of health outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	
Measurement of effectiveness	11a	*Single study-based estimates:* Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.	

	11b	*Synthesis-based estimates:* Describe fully the methods used for identification of included studies and synthesis of clinical effectiveness data.	
Measurement and valuation of preference based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	
Estimating resources and costs	13a	*Single study-based economic evaluation:* Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	
	13b	*Model-based economic evaluation:* Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	
Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.	
Choice of model	15	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.	
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytical model.	
Analytical methods	17	Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.	
Results			
Study parameters	18	Report the values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.	
Incremental costs and outcomes	19	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.	
Characterising uncertainty	20a	*Single study-based economic evaluation:* Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact	



		of methodological assumptions (such as discount rate, study perspective).	
	20b	*Model-based economic evaluation:* Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.	
Characterising heterogeneity	21	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.	
Discussion			
Study findings, limitations, generalisability, and current knowledge	22	Summarise key study findings and describe how they support the conclusions reached. Discuss limitations and the generalisability of the findings and how the findings fit with current knowledge.	
Other			
Source of funding	23	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support.	
Conflicts of interest	24	Describe any potential for conflict of interest of study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors recommendations.	

For consistency, the CHEERS Statement checklist format is based on the format of the CONSORT statement checklist

The **ISPOR CHEERS Task Force Report** provides examples and further discussion of the 24-item CHEERS Checklist and the CHEERS Statement. It may be accessed via the *Value in Health* link or via the ISPOR Health Economic Evaluation Publication Guidelines – CHEERS: Good Reporting Practices webpage: http://www.ispor.org/TaskForces/EconomicPubGuidelines.asp

The citation for the CHEERS Task Force Report is:
Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS)—Explanation and elaboration: A report of the ISPOR health economic evaluations publication guidelines good reporting practices task force. Value Health 2013;16:231-50.

F. Summary of studies included in systematic review

Table 10-1: Summary of characteristics and quality of studies included in the systematic review

First author	Title	Year of publication	Country	Type of EE	Outcome reporting	% applicable Drummond criteria met	% applicable CHEERS criteria met
Alkhadra, T	Cost -effectiveness of a pit and fissure sealants program in a school-based setting in Saudi Arabia	2004	Saudi Arabia	CEA	Clinician	38	65
Atkins, C.	Cost-effectiveness of preventing dental caries and full mouth dental reconstructions among Alaska Native children in the Yukon–Kuskokwim delta region of Alaska	2016	USA	CEA	Clinician	46	96
Bergström, E	Caries and costs: An evaluation of a school-based fluoride varnish programme for adolescents in a Swedish region	2016	Sweden	CMA	Clinician	23	70
Bertrand, É	Cost-effectiveness simulation of a universal publicly funded sealants application program	2011	Canada	CEA	Clinician	54	91
Bhuridej, P	Four-year cost-utility analyses of sealed and nonsealed first permanent molars in Iowa Medicaid-enrolled children	2007	USA	CUA	Clinician	92	86
Chi, D	Cost-Effectiveness of Pit-and-Fissure Sealants on Primary Molars in Medicaid-Enrolled Children	2014	USA	CEA	Clinician	69	96
Davies, G	An assessment of the cost effectiveness of a postal toothpaste programme to prevent caries among five-year-old children in the North West of England	2003	UK	CEA	Clinician	54	90
Frazão, P	Cost-effectiveness of conventional and modified supervised toothbrushing in preventing caries in permanent molars among 5-year-old children	2012	Brazil	CEA	Clinician	54	48
Goldman, A	Methods and preliminary findings of a cost-effectiveness study of glass-ionomer-based and composite resin sealant materials after 2 yr	2014	China	CEA	Clinician	85	90

First author	Title	Year of publication	Country	Type of EE	Outcome reporting	% applicable Drummond criteria met	% applicable CHEERS criteria met
Goldman, A	Cost-effectiveness, in a randomized trial, of glass-ionomer-based and resin sealant materials after 4 yr	2016	China	CEA	Clinician	100	95
Griffin, S	Comparing the costs of three sealant delivery strategies	2002	USA	CEA	Clinician	77	87
Hichens, L.	Cost-effectiveness and patient satisfaction: Hawley and vacuum-formed retainers	2007	UK	CEA	Child and clinician	36	84
Hietasalo, P.	Cost-effectiveness of an experimental caries-control regimen in a 3.4-yr randomized clinical trial among 11-12-yr-old Finnish schoolchildren	2009	Finland	CEA	Clinician	69	50
Hirsch, G	A simulation model for designing effective interventions in early childhood caries	2012	USA	CEA	N/A	0	39
Holland, T	The effectiveness and cost of two fluoride program for children	2001	Ireland	CEA	Clinician	62	35
Jokela, J.	Economic evaluation of a risk-based caries prevention program in preschool children	2003	Finland	CEA	Clinician	8	60
Kaakko, T	An ABCD program to increase access to dental care for children enrolled in Medicaid in a rural county	2002	USA	CEA	Clinician	15	52
Koh, R.	Relative cost-effectiveness of home visits and telephone contacts in preventing early childhood caries	2015	Australia	CEA + CUA	N/A	77	96
Kowash, M.	Cost-effectiveness of a long-term dental health education program for the prevention of early childhood caries	2006	UK	CEA + CBA	Clinician	15	85
Leskinen, K.	Practice-based study of the cost-effectiveness of fissure sealants in Finland	2008	Finland	CEA	Clinician	8	76
Marino, R	Modeling an economic evaluation of a salt fluoridation program in Peru	2011	Peru	CEA	Clinician	77	87

First author	Title	Year of publication	Country	Type of EE	Outcome reporting	% applicable Drummond criteria met	% applicable CHEERS criteria met
Mariño, R	Cost-effectiveness models for dental caries prevention programmes among Chilean school children	2012	Chile	CEA	Clinician	54	67
Mariño, R.	The cost-effectiveness of adding fluorides to milk-products distributed by the National Food Supplement Programme (PNAC) in rural areas of Chile	2007	Chile	CEA	Clinician	54	76
Morgan, M	Economic evaluation of a pit and fissure dental sealant and fluoride mouthrinsing program in two nonfluoridated regions of Victoria, Australia	1998	Australia	CEA	Clinician	38	81
Neidell, M.	Cost-Effectiveness Analysis of Dental Sealants versus Fluoride Varnish in a School-Based Setting	2016	USA	CEA	Clinician	38	77
Ney, J. P.	Economic modeling of sealing primary molars using a "value of information" approach	2014	USA	CEA	Clinician	38	87
Oscarson, N	Cost-effectiveness of different caries preventive measures in a high-risk population of Swedish adolescents	2003	Sweden	CEA	Clinician	85	90
Ouyang, W	Cost -effectiveness analysis of dental sealant using econometric modeling	2009	USA	CEA	Clinician	69	87
Petrén, S	Early correction of posterior crossbite-a cost-minimization analysis	2013	Sweden	CMA	Clinician	38	62
Pukallus, M.	Cost-effectiveness of a telephone-delivered education programme to prevent early childhood caries in a disadvantaged area: a cohort study	2013	Australia	CEA	Clinician	54	100
Quiñonez, R.	Assessing cost-effectiveness of sealant placement in children	2005	USA	CEA	Clinician	54	82
Quinonez, R	Simulating cost-effectiveness of fluoride varnish during well-child visits for Medicaid-enrolled children	2006	USA	CEA	Clinician	62	95

First author	Title	Year of publication	Country	Type of EE	Outcome reporting	% applicable Drummond criteria met	% applicable CHEERS criteria met
Ramos-Gomez, F.	Cost-effectiveness model for prevention of early childhood caries	1999	USA	CEA	Clinician	8	55
Sakuma, S	Economic Evaluation of a School-based Combined Program with a Targeted Pit and Fissure Sealant and Fluoride Mouth Rinse in Japan	2010	Japan	CEA	Clinician	62	62
Samnaliev, M	Cost-effectiveness of a disease management program for early childhood caries	2015	USA	CEA	Clinician	46	100
Sköld, U	Cost-analysis of school-based fluoride varnish and fluoride rinsing programs	2008	Sweden	CEA	Clinician	62	91
Stearns, S	Cost-effectiveness of preventive oral health care in medical offices for young medicaid enrollees	2012	USA	CEA	Clinician	46	100
Tagliaferro, E	Cost-effectiveness analysis of preventive methods for occlusal surface according to caries risk: results of a controlled clinical trial	2013	Brazil	CEA	Clinician	77	43
Tickle M,	A randomised controlled trial to measure the effects and costs of a dental caries prevention regime for young children attending primary care dental services	2016	UK	CEA	Clinician	38	95
Tonmukayakul, U	Cost-effectiveness analysis of the atraumatic restorative treatment-based approach to managing early childhood caries	2017	Australia	CEA	Clinician	46	95
Vermaire, J	Value for money: economic evaluation of two different caries prevention programmes compared with standard care in a randomized controlled trial	2014	Netherlands	CEA	Clinician	100	90
Weintraub, J	Treatment outcomes and costs of dental sealants among children enrolled in Medicaid	2001	USA	CEA	Clinician	0	78
Wiedel, A.	A cost minimization analysis of early correction of anterior crossbite - A randomized controlled trial	2016	Sweden	CMA	Clinician	23	71

First author	Title	Year of publication	Country	Type of EE	Outcome reporting	% applicable Drummond criteria met	% applicable CHEERS criteria met
Wu, Y	Cost-minimization analysis of two methods during the prevention of dental fear during caries filling treatments	2002	China	CMA	Clinician	8	33
Yee, R	A cost-benefit analysis of an advocacy project to fluoridate toothpastes in Nepal	2004	Nepal	CBA	Clinician	23	71
Zabos, G	Cost-effectiveness analysis of a school-based dental sealant program for low-socioeconomic-status children: A practice-based report	2002	USA	CEA	Clinician	31	57

Notes: CEA: Cost-effectiveness analysis, CUA: Cost-utility analysis, CBA: Cost-benefit analysis, CMA: Cost-minimisation analysis

Table 10-2: Summary of methods and economic characteristic of studies included in systematic review

First author	Nature of the intervention	Oral health condition studied	Setting	Study duration	Number of participants	Age of participants	Measure of effect	Costs	Effects	Cost-effectiveness
Alkhadra, T	Pit and fissure sealants plus annual check-ups plus oral hygiene instruction through videotape presentations vs oral hygiene instruction through videotape presentations only	Caries	Community (non-dental)	4 years	200	6 years	Number of teeth saved	234,309.63 (intervention) and 198,055.57 (control) Saudi Riyals	220 teeth saved in intervention arm	165 Riyals per tooth saved
Atkins, C.	Water fluoridation, dental sealants, fluoride varnish applications, home tooth brushing with fluoride toothpaste, and conducting initial dental	Caries	Combination	10 years	1536	0-60 months	Number of caries prevented	$17,597(fluoridation),$12,058 (sealants), $61,090 (varnish) $60,500 (toothbrushing)	137 (fluoridation), 39 (sealants), 133 (fluoride varnish), 168 (toothbrushing) caries	Toothbrushing prevented the greatest number of caries

	exams on children less than 18 months of age with parents receiving parental counselling								prevented	
Bergström, E	Population-based programme with fluoride varnish applications	Caries	Community (non-dental)	2005-2008 & 2010-2013	16751	12-15years	DFT, DFSa, DeSa	400SEK per participant in the intervention group over 4 years	Caries increment and prevalence significantly lower in intervention group	Costs and gains due to prevented fillings broke even
Bertrand, É	Universally public-funded pit & fissure sealants in three situations: mixed (school and private), private clinics, schools	Caries	Combination	10 years	Virtual population: 78,372	8 years	Number of children without decay on first permanent molars	Private (14,257,324) School (11,723,584)	Private (64,672) School (65,626)	Incremental C/E ratio Private vs mixed (868) School vs mixed (172)
Bhuridej, P	Sealed vs non-sealed permanent molars	Caries	General Dental Practice	1996-2000	2132	6 to 10	QATY	Non sealed 7.81 and sealed 13.3	Non sealed 0.93, sealed 0.94	Relative incremental cost to gain one additional QATY ranged $171.1 to $510.3
Chi, D	Always vs never seal in Medicaid enrolled children	Caries	General Dental Practice	2008-2011	N/A	Younger than 6 years	Teeth restored or extracted	Standard care cost $214510, Always seal $232141, Never seal $186010	Relative to standard care always seal 2389 to 340, never seal increased 2389 to 2853	Always seal $8.12 per restoration avoided, Standard care compared to never seal $65.62
Davies, G	Postal toothpaste programme vs standard care	Caries	Community (non-dental)	4 years	5344	12 to 60 months	dmft	Intervention £27.93 per child	DMFT intervention 2.15, control 2.57	Estimated cost per tooth saved from caries £80.83
Frazão, P	Bucco-lingual supervised toothbrushing vs standard supervised toothbrushing	Caries	Community (non-dental)	Not stated	280	N/S	Incidence density (ID)	$1.79 per capita modified brushing	ID 50% lower among boys in modified brushing	Marginal C/E ratio $6.30 per avoided caries
Goldman, A	Glass-ionomer based sealants vs composite resin sealants after 2	Caries	Community (non-dental)	2 years	380	7-9 years	Dentine lesions	$3.73 composite to $7.5 GIC	New dentine lesions composite 11,	GIC to prevent one additional carious lesion

	years								GIC 10	compared to composite $1106
Goldman, A	Glass-ionomer based sealants vs composite resin sealants after 4 years	Caries	Community (non-dental)	4 years	365	7-9 years	% effectiveness dentine carious lesions	HVGIC (4.57) LED HVGIC (4.65) Glass carbomer 7.84 Composite 3.05	HGVIC (97.3%) LED HVGIC 98% Glass carbomer 94.5% Composite 96.4%	The additional cost per LED thermocured HVGIC restoration over composite resin is $1.70 for producing 1000 sealants and $0.96 for producing 6000 sealants
Griffin, S	Seal all (SA) vs seal for children at high risk (Target) vs seal none (SN) (permanent 1st molars)	Caries	Not stated	9 years	Not stated	72-83 months	Permanent 1st molar occlusal surface caries increment	Seal none ($0), Seal all ($27), Target ($12.06)	SN vs SA (0.278) Target vs SA (0.101) SN vs Target (0.177)	SN (23.42), SA (73.96), Target (Cost saving relative to SN)
Hichens, L.	Hawley retainers vs vacuum-formed retainers (VFR)	Malocclusion	Specialist Practice	6 months	355	14-15 years	Little's irregularity Index (LII) & patient satisfaction questionnaire	Cost to NHS Hawley 152.42 euros, VFR 121.08 euros, cost to patient Hawley 152 euros, VFR 122.02 euros	LII statistically significant greater in hawley group compared with VFR group over 6 months	VFR were dominant over hawley retainer
Hietasalo, P.	Individual caries-control regimen vs standard care	Caries	Community (dental)	2001-2005	497	11 to 12	DMFS	Control 354.78 euros, experimental 298.28 euros	Experimental 3.33, control 4.58	Intervention cost effective
Hirsch, G	6 categories of ECC intervention: applying fluorides, limiting cariogenic bacterial transmission from mothers to children, using xylitol directly with children, clinical treatment, motivational interviewing, and combinations of these	Caries	Not stated	10 year	N/A	younger than 6 years	DFT	16,881 fewer affected teeth	$6 million for water fluoridation intervention	Community water fluoridation produces 6.6% relative reduction in cavity prevalence in children with $8 million net savings over ten years

Holland, T	School-based mouth-rinse programme compared with community water fluoridation	Caries	Community (non-dental)	6 years	380	12 years	DMFT	Mouth rinse 3.26 euros, Fluoridated 0.21 euros	Mouth rinse 1.32, no fluoride 1.82, fluoridated 1.25	Population size below 3168 MR intervention cost effective, population above 3168 fluoridated intervention cost effective
Jokela, J.	Risk-based caries prevention program vs conventional prevention	Caries	Community (non-dental)	3 years	525	2-5 years	Time spent on treatment	Control group 69 euros, risk based group 54 euros	Risk based 107mins 48 secs, routine 107mins 72 secs	Intervention cost effective if assistant delivered
Kaakko, T	Medicaid-enrolled children assigned to either ABCD programme or to regular benefits	Caries	General Dental Practice	1997-1999	437	1-4 years	Rate of utilization & dmft	$67.32 ABCD, $52.44 routine	Utilisation rate ABCD 33%, routine 21.5%	ABCD children had fewer teeth with initial caries, the average incremental cost per child per initial lesion prevented was $31.44
Koh, R.	Home visits conducted by OH therapists vs telephone-based vs no intervention	Caries	Community (non-dental)	5.5 years	N/S	6 months-6 years	QALYs	Usual care ($348903), telephone ($204193), home visits ($181870)	Usual care (258 lesions / QALYs 540) telephone (158 caries / QALYs 546) home visits (caries 145 / QALYs 547)	Both interventions were dominant
Kowash, M.	Long-term dental health education home visits vs slow releasing fluoride device, community water fluoridation, school-based fissure sealant program	Caries	Community (non-dental)	3 years	179	8 -44 months	dmft/s	Leeds DHE $20626, Leeds F glass devices $20,413, Community water F $12,126, school fissure sealant 40,826	Caries reduction Leeds DHE 96% (10,752 carious surface saved, Leeds F glass devices 76% 8512, community water F 10%	Leeds DHE 1.92, Leeds F glass devices 2.4, community water F 8.66, school fissure sealant 23.74

									1400, school fissure sealant 90% 1728	
Leskinen, K.	Sealing only high-caries risk patients vs all patients routinely sealed	Caries	Community (dental)	6 years	2,922	5-12 years	surface-specific filling increments of permanent first molars ans incisors	All sealed 235 euros, high risk sealed 185 euros	Restored surfaces all sealed 98/100, restored surfaces of high risk sealed 85/100	Sealing high risk patients only is cost effective
Marino, R	Salt fluoridation vs non-intervention	Caries	Community (non-dental)	6 years	N/A	6-12 years	DMFT/	Salt fluoridated 982328 sol, non fluoridated 1,741,717 sol	Salt fluoridated DMFT 3.07, non intervention DMFT 5.46	Incremental benefits (DMFT avoided per child) 59,500 sol. Incremental cost per DMFT avoided 11.95
Mariño, R	3 community based programmes: water fluoridation, salt-fluoridation and dental sealants and 4 school-based programmes: milk-fluoridation, fluoridated mouthrinses, APF_Gel and supervised toothbrushing with fluoride toothpaste	Caries	Community (non-dental)	6 years	N/A	6-12 years	DMFT	water F $112177, salt F $17311, sealants $2348850, APF gel $143219, FMR 38983, milk F $16746, toothbrush $112686	water F 40%, salt F 44%, selaants 50%, APF gel 21%, FMW 26%, milk F 53%, toothpaste 24%	Incremental cost per DMFT avoided water F 14.89, salt F 16.21, sealants 11.56, milk F 14.78, FMW 8.63, toothbrushing 8.55, APF 21.3
Mariño, R.	Fluoridated powdered milk and milk cereal vs non-fluoridated milk products	Caries	Community (non-dental)	1995-1999	2000	6 months - 6years	dmft	Intervention $17191589.78, control $28351391.58	Intervention DMFT 2.08, control DMFT 3.49	Incremental cost per DMFT avoided 2695.61
Morgan, M	Pit and fissure sealant, weekly fluoride mouthrinsing and annual oral hygiene education session vs only oral hygiene education	Caries	Community (non-dental)	1989-1991	444	12-13 years	DMFS	$24, 750 intervention, 46,750 control	Intervention DMFS 1.12; control DMFS 2.35	Incremental cost $11.80 per DMFS avoided
Neidell, M.	Sealant vs varnish vs no treatment (RCT)	Caries	Community (non-dental)	4 years	N/S	6-8 years	% Caries reduction	Sealants $104.25, Varnish $44.96	Sealants 76.3%, Varnish 43.9%	ACER sealants $137, ACER for varnish $102.

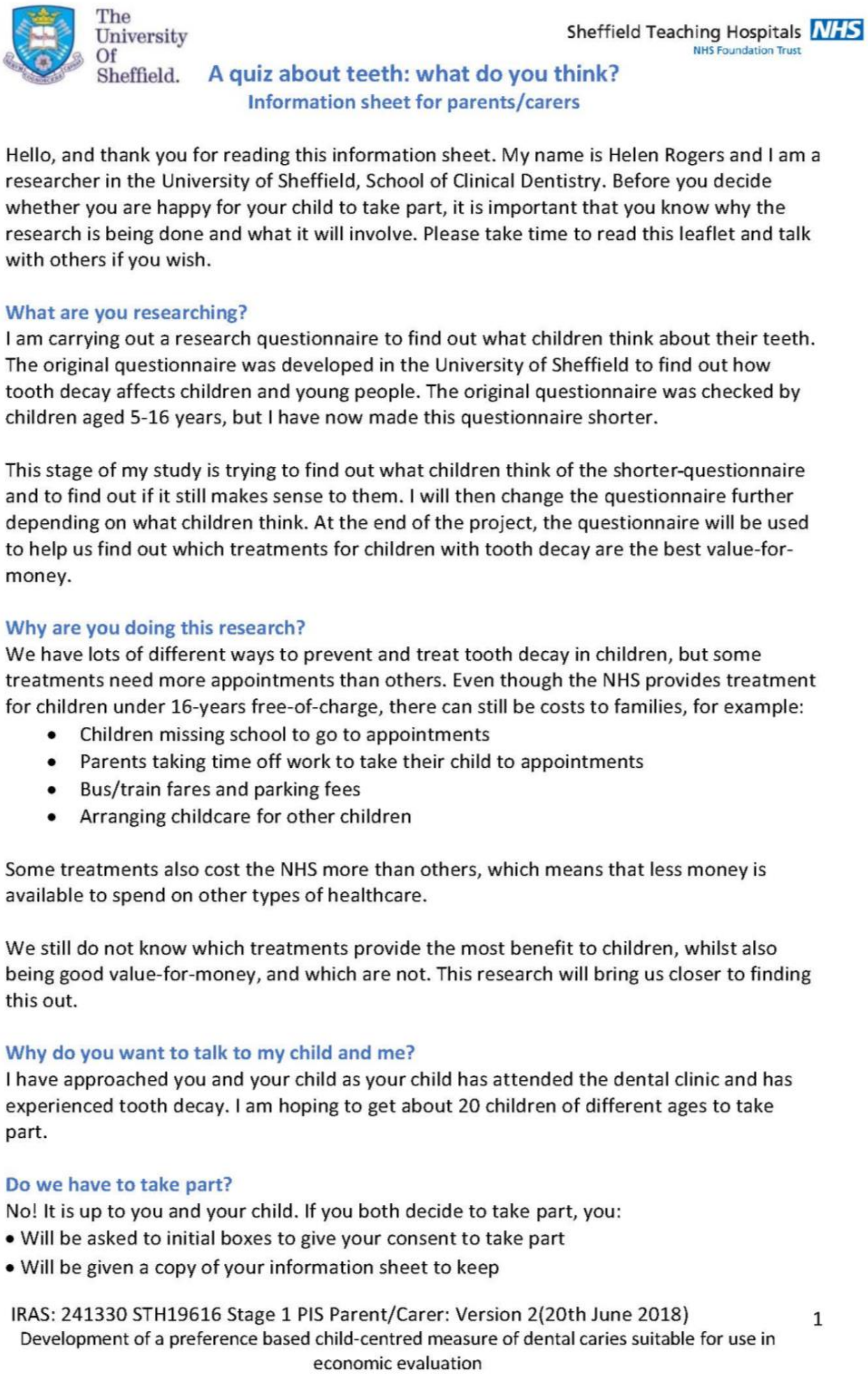
										ACER ratio 1.33 varnish more cost effective
Ney, J. P.	Always seal (AS) vs standard care (SC)	Caries	General Dental Practice	10 years	1250	3	Restorations or extractions averted	AS $182.47, SC $168.95	Numbers of restorations and extractions averted AS 0.28, SC 2.08	ICER $7.49, intervention more cost effective
Oscarson, N	Assigned to one of 4 preventative programmes, representing a step-wise increase in fluoride content, contact with dental personnel and cost	Caries	Community (dental)	1995-1999	3373	12	DMFS	Letter SEK199, fluoride lozenges SEK412 , duraphat 6 months SE1297, duraphat 3 months SEK1014	Letter 3.5, fluoride lozenges 3.0 , duraphat 6 months 3.0, duraphat 3 months 2.9	Incremental cost effectiveness of SEK2043 per averted DMFS
Ouyang, W	Pit and fissure sealants vs routine care	Caries	Community (non-dental)	4 years	200	6-10 years	Presence of caries	Sealant $56.84, unsealed $13.13	FPM healthy sealed 94.3%, FPM healthy unsealed 91.8%	ICER sealant cost $42.16 more than non sealant treatment to get one more caries free month FPM
Petrén, S	Quad Helix (QH) vs expansion plates (EP)	Malocclusion	Combination	3 years	40	N/S	Success rate of crossbite correction and degree of maxillary expansion in mm	QH SEK1031, EP SEK 1395	QH 2.8mm, EP 2.6mm	The QH had lower direct and indirect costs and fewer failures needing retreatment
Pukallus, M.	Telephone prevention programme vs usual care	Caries	Community (non-dental)	5.5 years	N/S	6 months to 6 years	Number of carious teeth	Telephone £19926 usual care £89910	Usual care 54 carious teeth, tel 11 carious teeth	Telephone intervention dominant
Quiñonez, R.	Seal all (SA) vs risk-based seal (RBS) vs seal none (SN)	Caries	Not stated	10 years	N/S	N/S	Cavity-free months	SA $54.60 RBS $53.8 SN $68.10	SA 97.4% RBS 86.4% SN 76.3%	C/E SA 0.56, RBS 0.62, SN 0.89, incr C/E NS dominated, SA 0.08
Quinonez, R	Application of fluoride varnish at 9, 18, 24 and 36 months vs no	Caries	General Dental Practice	42 months	5171	9- 42 months	cavity-free months	FVA $181.66, FVN $170.73	FVA cavity free mouth 31.49, FVN 29.97	ICER 7.18

	intervention									
Ramos-Gomez, F.	Minimal prevention (risk assessment and preventative treatment) vs intermediate (minimal + counseling) vs comprehensive (intermediate + outreach and incentives)	Caries	Combination	5 years	N/S	0-6years	dmfs	Minimal $314, intermediate $497, comprehensive $570	Minimal 40% 4.32 intermediate 70% 7.32 comprehensive 80% 8.36	Cost per carious tooth averted minimal $72.69 intermediate $65.74 comprehensive $66.28
Sakuma, S	Fluoride mouth rinse programme and targetted sealants vs control	Caries	Community (non-dental)	1992-1999/ 1995-1999	221	8-11years	DFT	Intevention Yen 2841, control -	Intervention 131 Yen, control 5348 Yen	1:184 cost benefit ratio
Samnaliev, M	Disease management intervention vs historial control	Caries	Hospital	2008-2010	395	0-60 months	Hospital based visits for restorative treatment or extractions	Reduction in societal costs ($669) per patient over 12 months	Reduction in hospital visits by 0.45 over 12 months for intervention	98.6% probability that intervention was less costly and more effective at 12 months
Sköld, U	Fluoride varnish treatment (FVT) vs fluoride mouth rinsing (FMR) vs control	Caries	Community (non-dental)	8 years	N/S	N/S	Prevented fillings	FVT SEK 35.80, FMR SEK 63	FVT (16.8 + 8.3) FMR (14.9 + 7.3)	Expressed as ratio of expected benefits to costs FVT 1.8:1 FMR 0.9:1
Stearns, S	Four or more medical office-based preventative oral health program (IBM) visits vs 0 IMB visits	Caries	Combination	2000-2006	209285	6-72 months	Visits with dental-related treatment	Cumulative reduction in Medicaid payments for dental services by $231	Intervention results in a statistically significant reduction in the likelihood of a child having a hospital episode related to caries	The intervention is 95% likely to be cost-effective if Medicaid is willing to pay at least $2,331 to avert a hospital episode.
Tagliaferro, E	OHI and fluoride varnish vs OHI and modified GIC sealant on first permanent molars	Caries	Community (non-dental)	Not stated	268	6 to 8	DMFS/ number of occlusal surfaces	Intervention R$18.17 control R$3.44	Control 7.2, intervention 1.2	Sealant C/E ratio $119.80, varnish $108.36, sealing most cost

							saved			effective
Tickle M,	Intervention: fluoride varnish containing 22,600 parts per million (p.p.m.) fluoride, a toothbrush and a 50-ml tube of toothpaste containing 1450 p.p.m. fluoride; plus standardised, evidence-based prevention advice provided at 6-monthly intervals over 3 years. Control group: prevention advice alone	Caries	General Dental Practice	2010-2013	1096	2-3 years	Conversion from caries-free to caries-active state and dmfs	Intervention £155.74, control £48.21	Intervention (34% 7.2, control 39% 9.6)	The intervention was unlikely to be cost-effective
Tonmukayakul, U	Atraumatic restorative treatment-based (ART-based) approach vs standard care (SC)	Caries	Combination	1 year	254	N/S	Number of referrals to specialists/ dental treatments for caries carried out (number of fillings/ extractions)	ART total $142595.59, control $179448.59	ART 6, control 62	ICER ART $654.05, control $677.70
Vermaire, J	Increased professional fluoride application (IPFA) vs non-operative caries treatment and prevention (NOCTP) vs standard care	Caries	General Dental Practice	2006-2008	179	6 years (+/- 3 months)	DMFS (prevented DMFS)	control 310, NOCTP 329, IPFA 494	control 0.57, NOCTP 0.41, IPFA 0.47	ICER - NOCTP v control 24, IPFA v control 265, NOCTP cost effective
Weintraub, J	Sealant(s) vs no sealants	Caries	General Dental Practice	1985-1992	15438	5-12 years	Caries-related services involving the occlusal surface (CRSOs)	Medicaid reimbursement for a sealant: $11.60; for an amalgam £18.57	Sealants effective in reducing the number of caries-related services involving occlusal	Sealant placement was broadly associated with expenditure savings to Medicaid for high risk

									surfaces	children
Wiedel, A.	Fixed appliance vs removable appliance	Malocclusion	Combination	Not stated	62	N/S, though mean age 9.8	Success rate of anterior crossbite correction and size of overjet(mm)	Fixed appliance: 630 euros; Removable appliance: 945 euros	Equivalence assumed between the two approaches	Fixed appliance therapy is more cost-effective
Wu, Y	Dental familiarisation video vs Local anaesthetic vs nil	Dental Fear	Combination	1996 to 1998	150	7 to 12	Venhams anxiety scale	Grp A 83.99 yuan, grp B 87.09 Control 87.76 yuan	Equivalence assumed	Grp A better clinical effect in dental fear prevention with lower cost
Yee, R	Advocacy to increase the availability and consumption of fluoride toothpaste vs no availability of fluoride toothpaste	Caries	Combination	2002-2007	N/S	6-18 years	DMFS	Intervention $594446	For every $1 potential saving $87 to $356	CBA shows intervention efficient
Zabos, G	School-based dental sealand program vs no intervention	Caries	Community (non-dental)	1983-1992	30 at follow-up	6-14 years	DMFS	Sealant $1720, control $2100	Caries incidence across 5 years control: 6.8 DMFS; intervention 2.2 DMFS	Administering sealants to children of low socio-economic status saves money relative to ordinary care

G. Parental information sheet and consent form for validation of the classification system



The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A quiz about teeth: what do you think?

Information sheet for parents/carers

Hello, and thank you for reading this information sheet. My name is Helen Rogers and I am a researcher in the University of Sheffield, School of Clinical Dentistry. Before you decide whether you are happy for your child to take part, it is important that you know why the research is being done and what it will involve. Please take time to read this leaflet and talk with others if you wish.

What are you researching?

I am carrying out a research questionnaire to find out what children think about their teeth. The original questionnaire was developed in the University of Sheffield to find out how tooth decay affects children and young people. The original questionnaire was checked by children aged 5-16 years, but I have now made this questionnaire shorter.

This stage of my study is trying to find out what children think of the shorter-questionnaire and to find out if it still makes sense to them. I will then change the questionnaire further depending on what children think. At the end of the project, the questionnaire will be used to help us find out which treatments for children with tooth decay are the best value-for-money.

Why are you doing this research?

We have lots of different ways to prevent and treat tooth decay in children, but some treatments need more appointments than others. Even though the NHS provides treatment for children under 16-years free-of-charge, there can still be costs to families, for example:

- Children missing school to go to appointments
- Parents taking time off work to take their child to appointments
- Bus/train fares and parking fees
- Arranging childcare for other children

Some treatments also cost the NHS more than others, which means that less money is available to spend on other types of healthcare.

We still do not know which treatments provide the most benefit to children, whilst also being good value-for-money, and which are not. This research will bring us closer to finding this out.

Why do you want to talk to my child and me?

I have approached you and your child as your child has attended the dental clinic and has experienced tooth decay. I am hoping to get about 20 children of different ages to take part.

Do we have to take part?

No! It is up to you and your child. If you both decide to take part, you:

- Will be asked to initial boxes to give your consent to take part
- Will be given a copy of your information sheet to keep

IRAS: 241330 STH19616 Stage 1 PIS Parent/Carer: Version 2(20th June 2018) 1
Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation

- Will be free at any time to stop taking part without giving a reason.

If you or your child decides to stop taking part, it will not affect the care your child receives.

What will happen if we agree to take part?

I will take you and your child to a quiet room, away from the busy dental clinic. I will give your child the questionnaire to complete, and will then ask them some questions about the questionnaire. The questionnaire and interview together should take no more than 10 minutes. The interview will be recorded and I will write some notes down.

Is there anything to be worried about if my child takes part?

There are no known risks to you or your child from taking part in the study. Your names will not appear in any report written about the study, so you do not need to worry that other people will know what you or your child have said. Your child's name will be removed from the interview recording, which will be kept securely on a password-protected computer in the university.

What are the possible benefits of taking part?

The study will not change the care or treatment your child receives at the dentist. The study will not benefit your child directly, but we hope that the study will help children with dental decay in the future. Your child will be given a £5 gift voucher to say thank you for taking part and for the time you have given.

What do I do next?

There is a form for you to fill in. If you and your child are happy to take part, please initial the boxes. If you would like more time to think about whether you want to take part, then you can take the forms away with you to complete later and bring them back for your child's next appointment.

If you would like to speak to me about your child's participation in the research, or about any other aspect of the project, please contact me by telephone: 0114 271 7877 or email hrogers1@sheffield.ac.uk.

What happens when the research stops?

When the study is finished I will look at all the information from your child and all the other children who have taken part. I will then write a report on my findings and send you a copy. Your child will continue their regular dental care as normal. The data collected throughout this research will be stored securely until the final report on the project has been written, before being destroyed. Information is stored in case more research can be done with the data collected.

What if there is a problem or something goes wrong?

If you or child are unhappy about anything, I will be happy to talk to you at any time. You and your child can also stop taking part at any time without having to give a reason.

What if I am not happy about the way the study has been conducted?

If you or your child are harmed by taking part, there are no special compensation arrangements. If you or your child are harmed due to someone's negligence, then you may have grounds for legal action. You may have to pay for this. Regardless of this, if you wish to complain, or have any concerns about any aspects of the way you have been approached or treated during the course of this study, the normal National Health Service complaints service is available to you.

If you have any complaints or concerns, please contact Helen Rogers on the number overleaf. If that is not satisfactory, please contact Mrs Tracey Plant, Clinical Hospital Manager, Charles Clifford Dental Hospital, Sheffield S10 2SZ or the Patient Services Team on 0114 271 2400 or email PST@sth.nhs.uk

Will anyone else know we are taking part?

All personal information recorded on paper will be kept under conditions of strict confidentiality. We will remove all personal details such as names and age from any electronic records, such as the voice recordings during the interviews so that your child cannot be recognised by the researchers. All the information from the study will be kept securely at the University of Sheffield. The reports will not name any participants.

Who is organising and funding the research?

The study is being organised by Helen Rogers, PhD student, Unit of Oral Health, Dentistry and Society, School of Clinical Dentistry, University of Sheffield. Helen is funded by the National Institute for Health Research, and is supported by an experienced supervisory team.

Who has reviewed the study?

Before any research goes ahead it is checked by an Ethics Committee. They make sure that the research is safe to do. This project has been checked and approved by South Yorkshire Research Ethics Committee.

Thank you for taking the time to read this. Please feel free to ask any questions if you need to.

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A quiz about teeth: what do you think?

Parent/Carer Consent Form

Participant Identifier	

Name of Researcher: Helen Rogers **Sponsor:** Sheffield Teaching Hospitals NHS Foundation Trust

If you agree to the following statements, please initial the boxes.

1. I have read the information provided for parents/carers about this project. I have had the chance to think about the information, ask questions and my questions have been answered. ☐
2. I understand that I can decide whether my child takes part in this project and that I can change my mind at any time without giving a reason. I understand this will not affect the dental care provided for my child. I understand this will not affect my legal rights, nor those of my child. ☐
3. I understand that personal information (my child's age, gender, ethnicity, contact details and postcode) that could identify my child will be collected during this study. This information will be stored in the School of Clinical Dentistry, University of Sheffield until the final report on this study has been written. It may be looked at by some members of the research team, the Sponsors or other responsible people to check the research is being done well. I agree for my child's information to be used in this way. ☐
4. I understand that the information collected about my child by the research team will be used to support other research in the future and may be shared with other researchers working with the sponsors, though your child's name would not be provided. ☐
5. OPTIONAL: I agree to be contacted by the research team in the future if they think my child may be suitable for taking part in other research projects. ☐
6. **I am happy for my child to take part in this project.** ☐

Please complete the details below:

Name of child:		**Parent/carer signature:**	
Name of parent/carer:		**Date:**	

Please leave this section blank for the researcher to complete:

Name of researcher:		**Date:**	
Researcher signature:			

IRAS: 241330 STH:19616 Stage 1 Parent/carer Consent Form: V2 (20th June 2018)
Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation

Funded by NHS National Institute for Health Research

H. Child information sheet and assent form for validation of the classification system

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A QUIZ ABOUT TEETH:
WHAT DO YOU THINK?

Who am I?
My name is Helen and I study at the University.

This is me:

What am I doing?
I am doing a project to find out what children think about a quiz about teeth

Can you help me?
I would like you to fill in a quiz for me. I will ask you some questions to find out what you think about the quiz. It will take about 10 minutes.

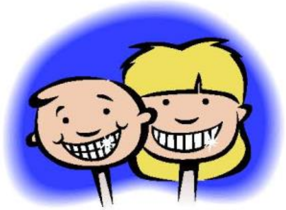
There are no right and wrong answers. It should be fun!

What happens after?
I will use your answers to make the quiz better. Then I will write a story to tell other people what I've found out.

After you have done the quiz and answered my questions, I will give you a small gift to say thank you for helping me.

IRAS: 241330 STH19616 Stage 1 PIS (5-10 yrs): Version 2 (20th June 2018)
Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation 1

What if you don't want to join in anymore?
If you don't want to do it anymore, you can stop at any time.
No one will be cross.



What do you do now?
There is a sheet to fill in before you start the quiz. You can get an adult to help you with this. Then we can start the quiz together!

If you have any questions or are unhappy, you can phone me on 0114 271 7877 or email hrogers1@sheffield.ac.uk or the Patient Services Team on 0114 271 2400 or email PST@sth.nhs.uk

Thank you for reading this.

Helen

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A quiz about teeth: what do you think?

Child Assent Form

Participant Identifier	
Name of child	

Name of Researcher: Helen Rogers **Sponsor:** Sheffield Teaching Hospitals NHS Foundation Trust

If you agree, please tick the boxes -

1. Somebody has read information to me about this study ☐
2. Somebody has told me what this study is about ☐
3. I understand what this study is about ☐
4. I have asked all the questions I want to ask ☐
5. My questions have all been answered well ☐
6. I know that I can stop taking part at any time ☐
7. **I would like to take part in this study** ☐

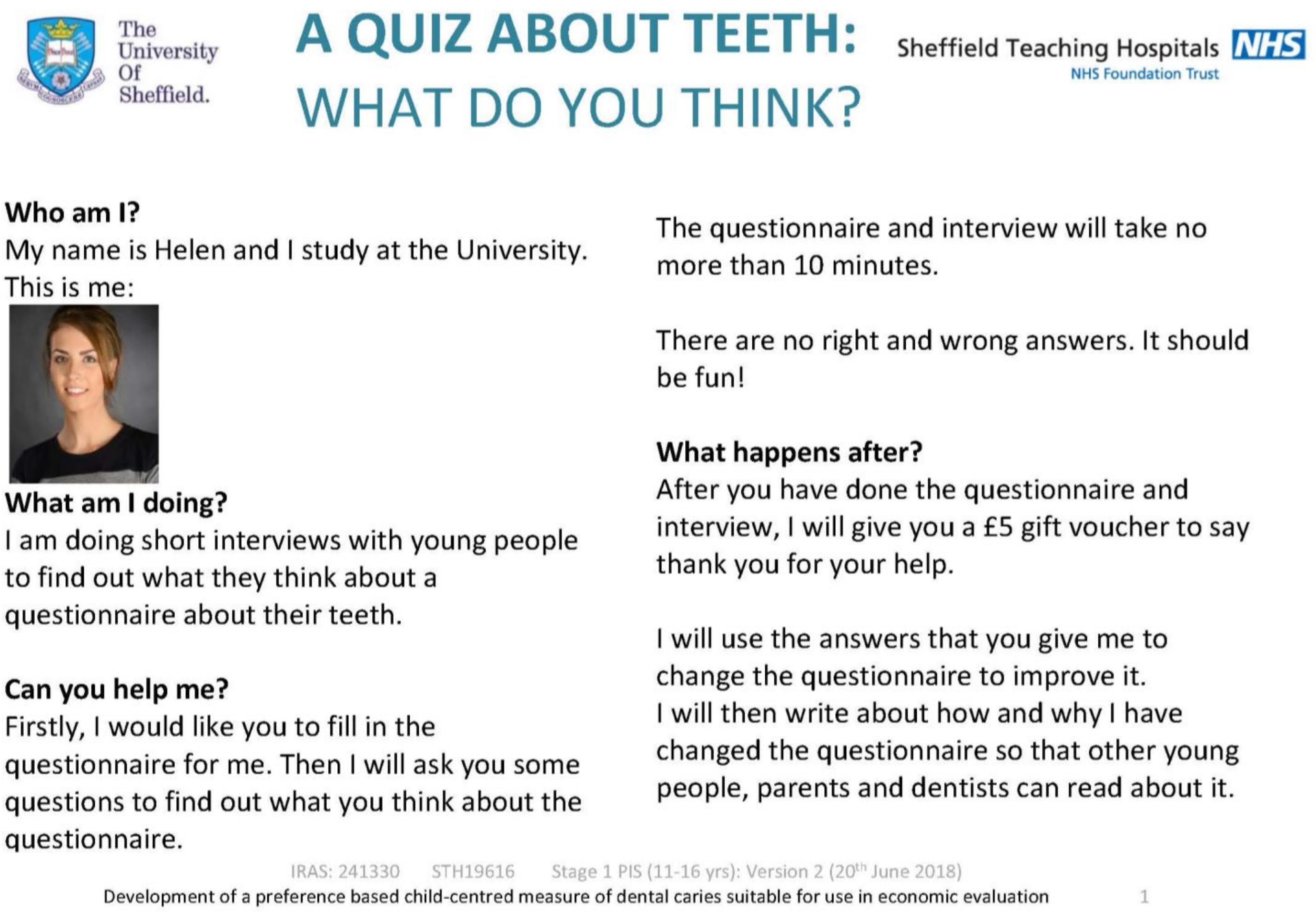
My name is:	
The date today is:	

That's it! You're ready to start the quiz!

Researcher signature:		**Date:**	

Funded by NHS National Institute for Health Research

I. Young person information sheet and assent form for validation of the classification system



The University Of Sheffield.

A QUIZ ABOUT TEETH: WHAT DO YOU THINK?

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Who am I?
My name is Helen and I study at the University.
This is me:

What am I doing?
I am doing short interviews with young people to find out what they think about a questionnaire about their teeth.

Can you help me?
Firstly, I would like you to fill in the questionnaire for me. Then I will ask you some questions to find out what you think about the questionnaire.

The questionnaire and interview will take no more than 10 minutes.

There are no right and wrong answers. It should be fun!

What happens after?
After you have done the questionnaire and interview, I will give you a £5 gift voucher to say thank you for your help.

I will use the answers that you give me to change the questionnaire to improve it.
I will then write about how and why I have changed the questionnaire so that other young people, parents and dentists can read about it.

IRAS: 241330 STH19616 Stage 1 PIS (11-16 yrs): Version 2 (20th June 2018)
Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation 1

What if you don't want to join in anymore?
It is fine for you to stop filling in the questionnaire or stop the interview at any time. No one will be unhappy with you.

What do you do now?
There is a form for you to fill in if you would like to take part. Please put your initials in the boxes, and write your name, signature and date in the spaces at the bottom.

If you have any questions or are unhappy, you can phone me on 0114 271 7877 or email hrogers1@sheffield.ac.uk or the Patient Services Team on 0114 271 2400 or email PST@sth.nhs.uk

Thank you for reading this.

Helen

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A quiz about teeth: what do you think?

Young Person Assent Form

Participant Identifier	

Name of Researcher: Helen Rogers **Sponsor:** Sheffield Teaching Hospitals NHS Foundation Trust

If you agree to the following statements, please initial the boxes.

1. I have read the information for young people about this project. I have had the chance to think about the information, ask questions and my questions have been answered. ☐
2. I understand that it is my choice to take part and that I can change my mind at any time without giving a reason. I understand it will not change my usual dental care. ☐
3. I understand that personal information about me collected during the study will be kept in the School of Clinical Dentistry, University of Sheffield until the research team have finished writing up their reports. It may be looked at by other responsible people to check the project is being done well. I agree for them to see my information. ☐
4. I understand that the information collected about me by the research team will be used to support other projects in the future and may be shared with other researchers, but they will not be given my name. ☐
5. OPTIONAL: I agree to be contacted by the research team in the future if they have other projects I could take part in. ☐
6. **I would like to take part in this project.** ☐

Please fill in your details below:

Full name:	
Signature:	
Date:	

Please leave this section blank for the researcher.

Researcher name:	
Signature:	
Date:	

Funded by NHS National Institute for Health Research

J. Topic guide for qualitative validation of the classification system

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Developing a caries-specific child-centred utility measure
A quiz about teeth: what do you think?

Topic Guide

Introduction

- Complete assent/consent forms for participant and parents/carers
- Complete socio-demographic data collection form
- Purpose is to talk to children and young people about a questionnaire which was made to help us find out what it's like to have holes in your teeth
- Children/young people will be asked to fill in the questionnaire. They will then be asked to 'think aloud' whilst completing it. They will also be asked some more questions about the questionnaire at the end
- Interview will last as long as they wish, but on average 10 minutes
- Use of a digital recorder by researcher, and the talk will later be written up
- Not a test, and no right or wrong answers (child/young person is the expert)
- Can ask for help with reading or filling in the quiz/questionnaire at any time
- Doesn't have to talk about anything they don't want to
- Participation is voluntary and children can change their mind and stop the interview at any point. Agree a stop signal (e.g. hand in the air, saying "stop")

Answers will be private

Think Aloud task prompts

- What are you thinking now?
- What are your thoughts about this page?
- Can you tell me more about that?

Questionnaire usability

- What did you think about the questionnaire?
- Go through each question and ask if participant found it easy, OK, or hard to answer and understand

- Can you see any problems with the questionnaire?

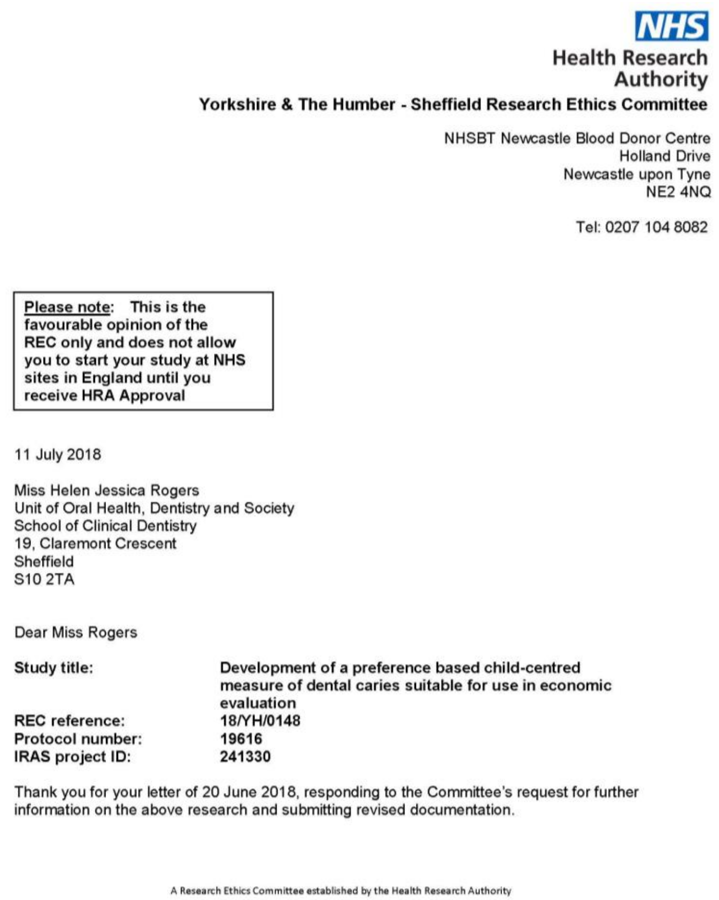
Removal and reintroduction of questions

- Were there any questions in the quiz that you didn’t like? Why didn’t you like them?
- Were there any questions in the quiz that you did like? Why did you like them?
- Were there any questions that you would want to take out of the quiz?
- Show participant questions that were removed from CARIES-QC
- Do you think any of these questions are important?
- Do you think any of these questions should be put into the questionnaire?

Next steps

- Thank the participant
- Reassurance again about confidentiality/privacy
- Explain that findings will ultimately be used to develop a questionnaire to help us find out which dental treatments are value-for-money. We will send all the participants a report so they know what we found out
- Ask the child and parent whether they would prefer the report to be sent to them via post or email
- Give the participant a gift voucher and ask them to sign reciept

K. Approvals from Research Ethics Committee and Health Research Authority



NHS
Health Research
Authority
Yorkshire & The Humber - Sheffield Research Ethics Committee

NHSBT Newcastle Blood Donor Centre
Holland Drive
Newcastle upon Tyne
NE2 4NQ

Tel: 0207 104 8082

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

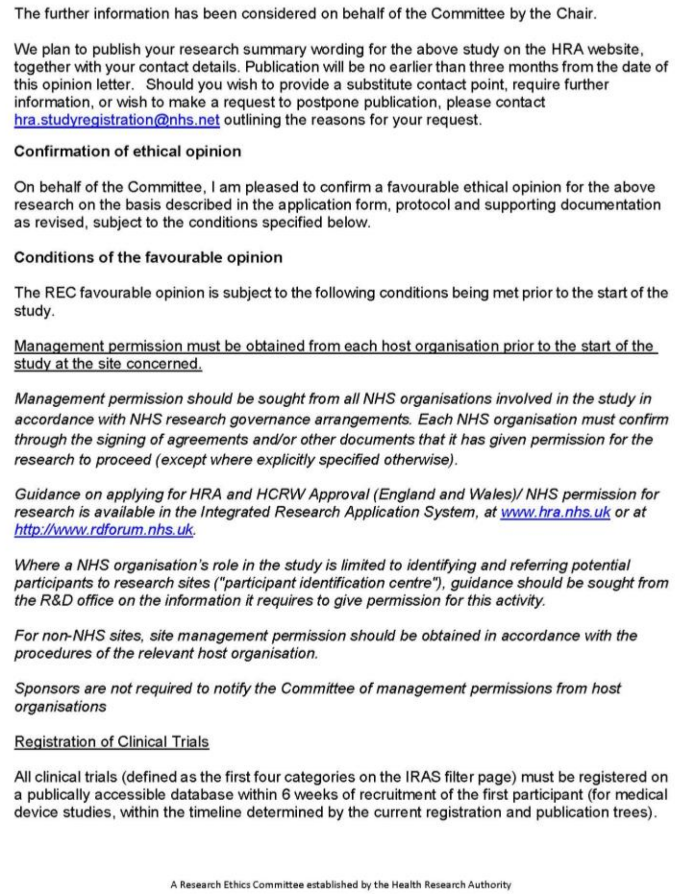
11 July 2018

Miss Helen Jessica Rogers
Unit of Oral Health, Dentistry and Society
School of Clinical Dentistry
19, Claremont Crescent
Sheffield
S10 2TA

Dear Miss Rogers

Study title:	**Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation**
REC reference:	**18/YH/0148**
Protocol number:	**19616**
IRAS project ID:	**241330**

Thank you for your letter of 20 June 2018, responding to the Committee's request for further information on the above research and submitting revised documentation.



The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact hra.studyregistration@nhs.net outlining the reasons for your request.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System, at www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Stage 1 Poster V2]	2	20 June 2018
Copies of advertisement materials for research participants [Stage 3 Poster V2]	2	20 June 2018
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Insurance Certificate]		07 February 2018
Interview schedules or topic guides for participants [Stage 1 Socio-dem form V2]	2	20 June 2018
Interview schedules or topic guides for participants [Stage 3 Topic Guide V2]	2	20 June 2018
Interview schedules or topic guides for participants [Stage 1 Topic Guide V2]	2	20 June 2018
Interview schedules or topic guides for participants [Safeguarding Protocols]	1	25 September 2017
Letter from funder [Letter of Intent DRF]		08 September 2016
Letters of invitation to participant [Stage 3 Parent opt-in covering letter V1 New Document]	1	20 June 2018
Other [Protocol V2 tracked changes]	2	20 June 2018
Other [Response to REC and HRA]	1	20 June 2018
Participant consent form [Stage 1 Assent 5-10 V2 tracked changes]	2	20 June 2018
Participant consent form [Stage 1 Assent 11-16 V2 tracked changes]	2	20 June 2018
Participant consent form [Stage 1 Consent V2 tracked changes]	2	20 June 2018
Participant consent form [Stage 3 Consent V1 New document]	1	20 June 2018
Participant consent form [Stage 3 Assent 11-16 V1 New Document]	1	20 June 2018
Participant information sheet (PIS) [Stage 1 PIS Parents V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 1 PIS 5-10 V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 1 PIS 11-16 V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 3 PIS 11-16 V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 3 PIS Parents V2 tracked changes]	2	20 June 2018
Referee's report or other scientific critique report [NIHR DRF Feedback]		01 June 2016
Summary CV for supervisor (student research) [H Rodd Short CV]		01 February 2018
Summary CV for supervisor (student research) [Donna Rowen Short CV]		01 February 2018
Summary CV for supervisor (student research) [Z Marshman Short CV]		01 February 2018

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

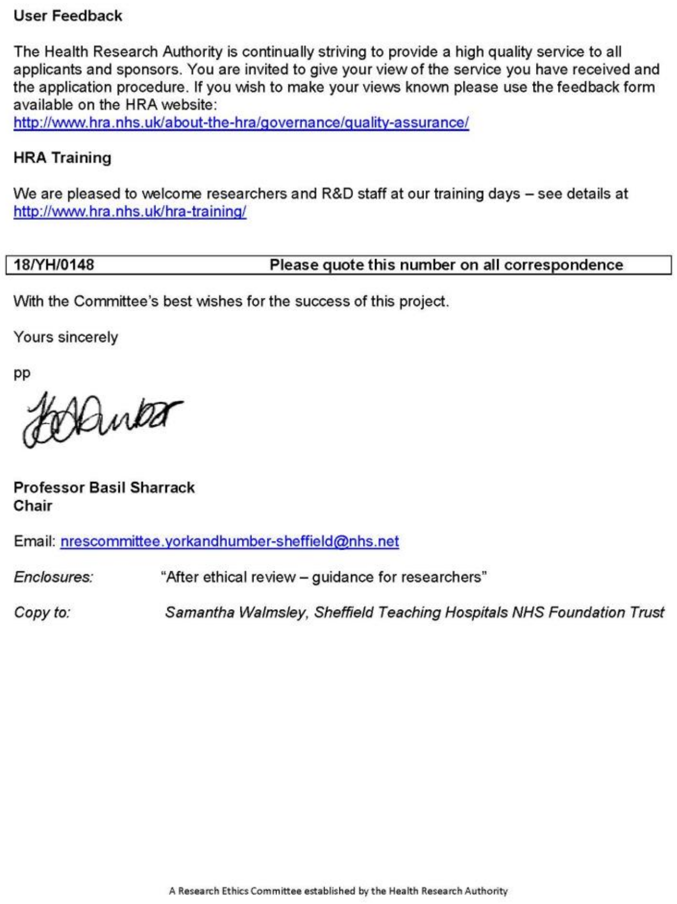
After ethical review

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.



User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

18/YH/0148	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely

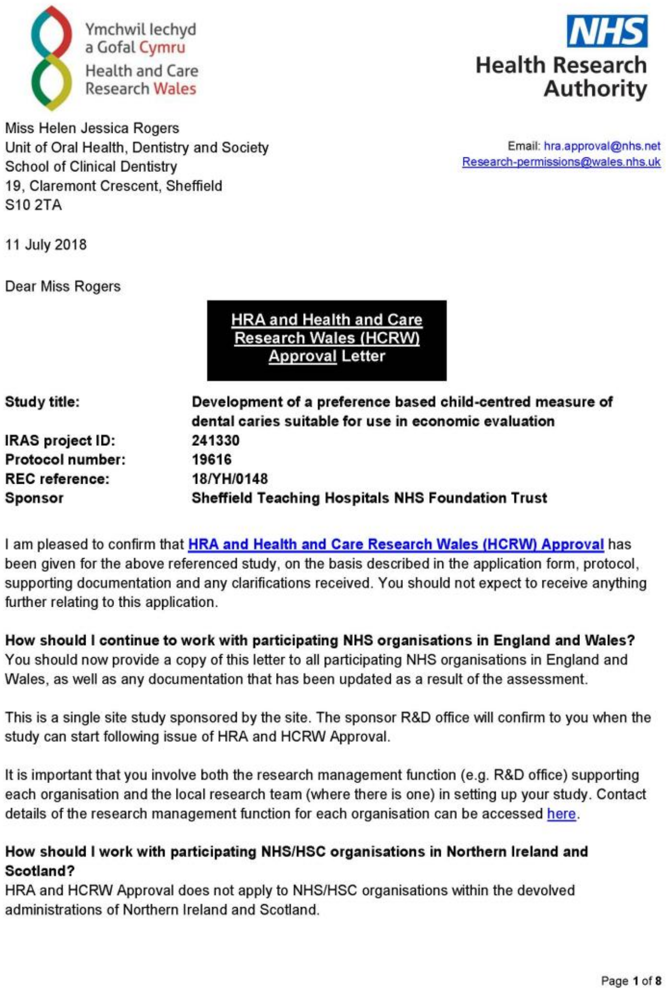
pp

Professor Basil Sharrack
Chair

Email: nrescommittee.yorkandhumber-sheffield@nhs.net

Enclosures: "After ethical review – guidance for researchers"

Copy to: *Samantha Walmsley, Sheffield Teaching Hospitals NHS Foundation Trust*



Ymchwil Iechyd
a Gofal Cymru
Health and Care
Research Wales

NHS
Health Research
Authority

Miss Helen Jessica Rogers
Unit of Oral Health, Dentistry and Society
School of Clinical Dentistry
19, Claremont Crescent, Sheffield
S10 2TA

Email: hra.approval@nhs.net
Research-permissions@wales.nhs.uk

11 July 2018

Dear Miss Rogers

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation**
IRAS project ID:	**241330**
Protocol number:	**19616**
REC reference:	**18/YH/0148**
Sponsor	**Sheffield Teaching Hospitals NHS Foundation Trust**

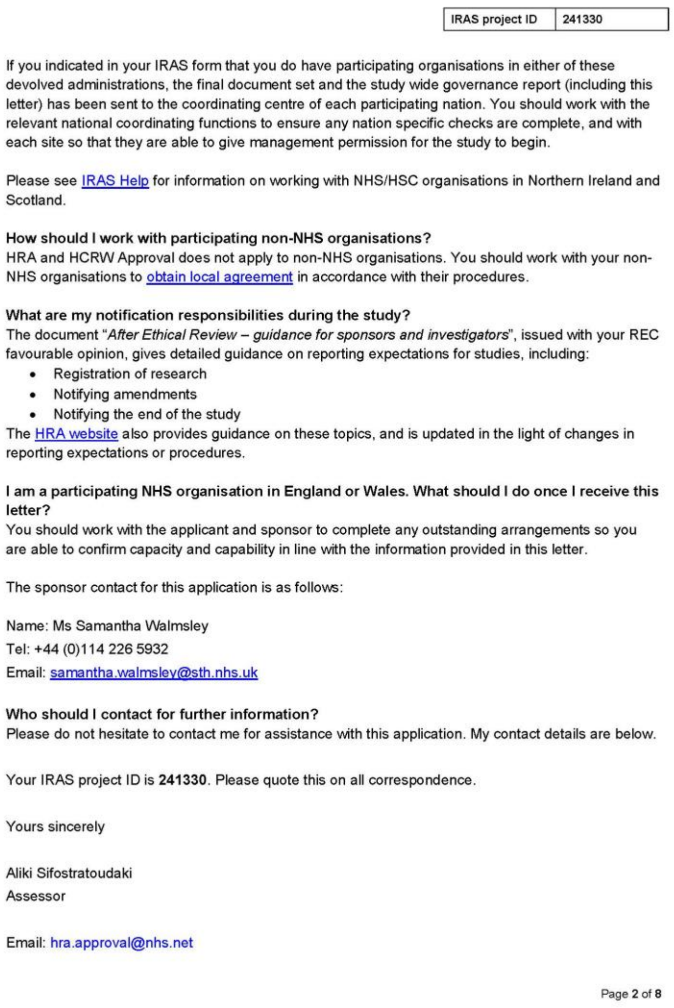
I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

How should I continue to work with participating NHS organisations in England and Wales?
You should now provide a copy of this letter to all participating NHS organisations in England and Wales, as well as any documentation that has been updated as a result of the assessment.

This is a single site study sponsored by the site. The sponsor R&D office will confirm to you when the study can start following issue of HRA and HCRW Approval.

It is important that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details of the research management function for each organisation can be accessed here.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within the devolved administrations of Northern Ireland and Scotland.



If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) has been sent to the coordinating centre of each participating nation. You should work with the relevant national coordinating functions to ensure any nation specific checks are complete, and with each site so that they are able to give management permission for the study to begin.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?
The document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

I am a participating NHS organisation in England or Wales. What should I do once I receive this letter?
You should work with the applicant and sponsor to complete any outstanding arrangements so you are able to confirm capacity and capability in line with the information provided in this letter.

The sponsor contact for this application is as follows:

Name: Ms Samantha Walmsley
Tel: +44 (0)114 226 5932
Email: samantha.walmsley@sth.nhs.uk

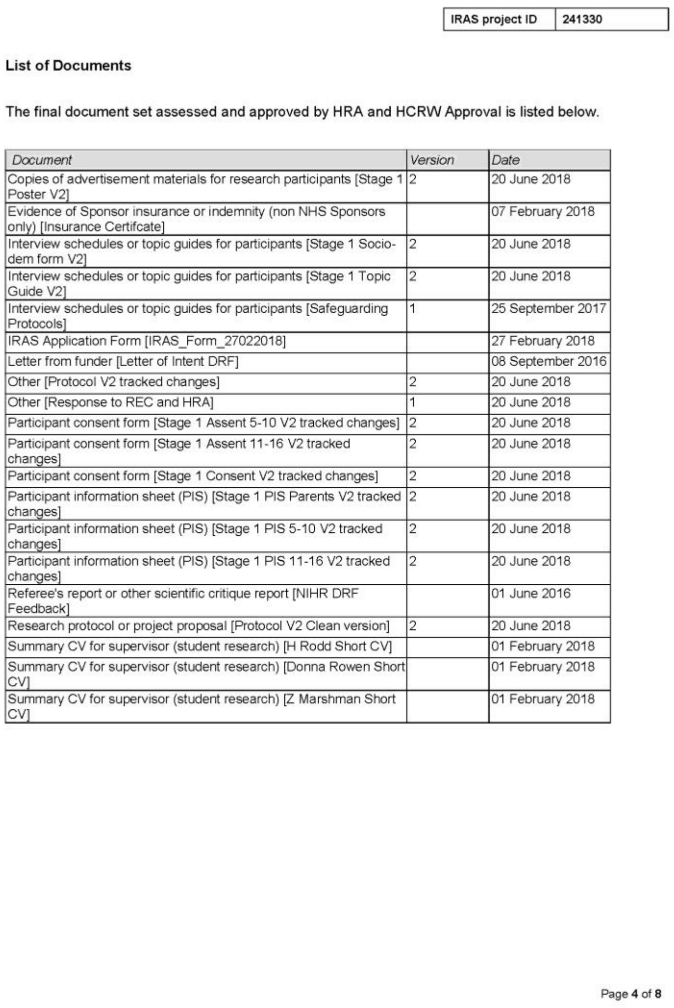
Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **241330**. Please quote this on all correspondence.

Yours sincerely

Aliki Sifostratoudaki
Assessor

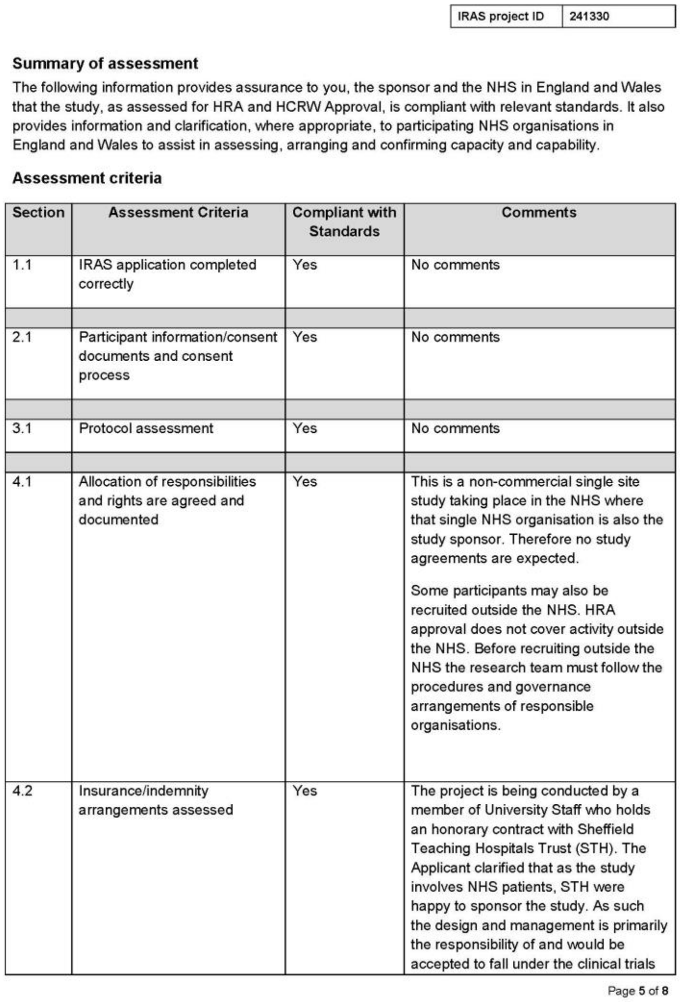
Email: hra.approval@nhs.net

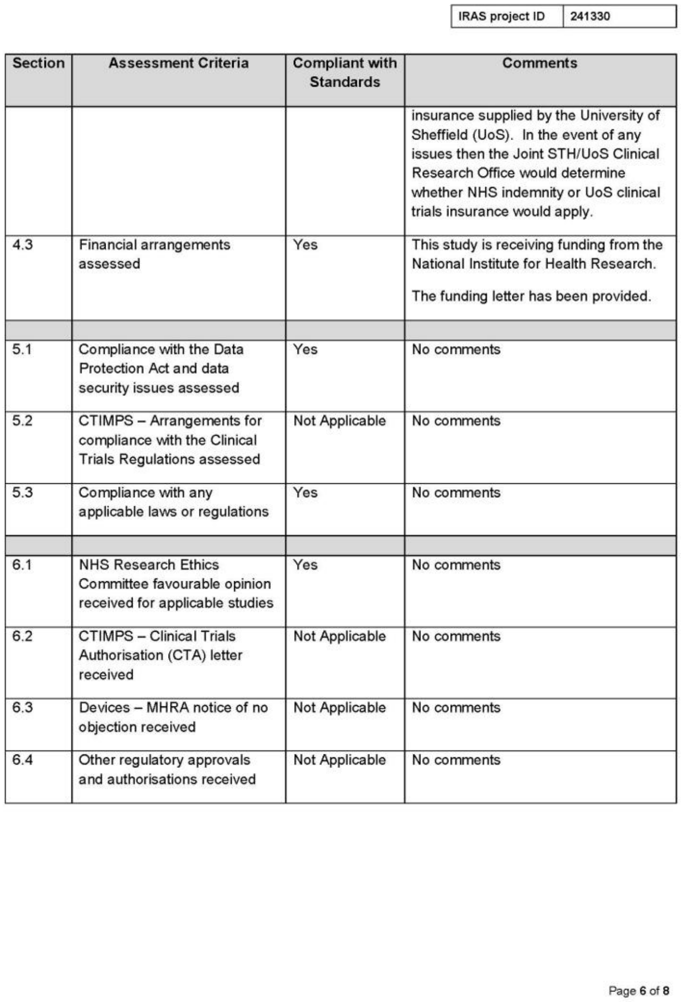


List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Stage 1 Poster V2]	2	20 June 2018
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Insurance Certifcate]		07 February 2018
Interview schedules or topic guides for participants [Stage 1 Socio-dem form V2]	2	20 June 2018
Interview schedules or topic guides for participants [Stage 1 Topic Guide V2]	2	20 June 2018
Interview schedules or topic guides for participants [Safeguarding Protocols]	1	25 September 2017
IRAS Application Form [IRAS_Form_27022018]		27 February 2018
Letter from funder [Letter of Intent DRF]		08 September 2016
Other [Protocol V2 tracked changes]	2	20 June 2018
Other [Response to REC and HRA]	1	20 June 2018
Participant consent form [Stage 1 Assent 5-10 V2 tracked changes]	2	20 June 2018
Participant consent form [Stage 1 Assent 11-16 V2 tracked changes]	2	20 June 2018
Participant consent form [Stage 1 Consent V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 1 PIS Parents V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 1 PIS 5-10 V2 tracked changes]	2	20 June 2018
Participant information sheet (PIS) [Stage 1 PIS 11-16 V2 tracked changes]	2	20 June 2018
Referee's report or other scientific critique report [NIHR DRF Feedback]		01 June 2016
Research protocol or project proposal [Protocol V2 Clean version]	2	20 June 2018
Summary CV for supervisor (student research) [H Rodd Short CV]		01 February 2018
Summary CV for supervisor (student research) [Donna Rowen Short CV]		01 February 2018
Summary CV for supervisor (student research) [Z Marshman Short CV]		01 February 2018



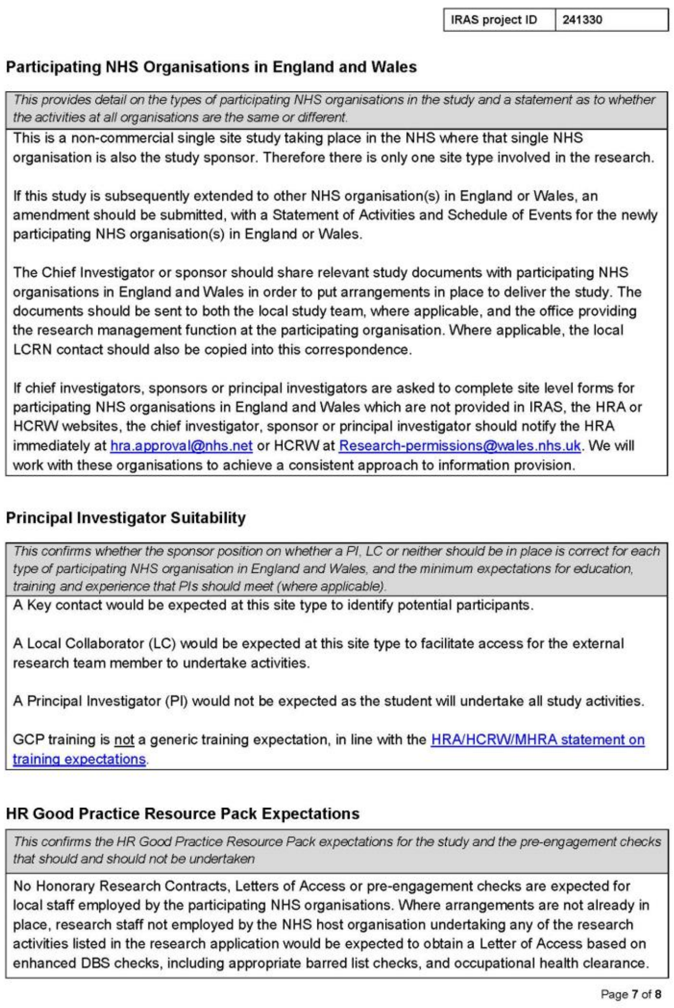


Summary of assessment

The following information provides assurance to you, the sponsor and the NHS in England and Wales that the study, as assessed for HRA and HCRW Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England and Wales to assist in assessing, arranging and confirming capacity and capability.

Assessment criteria

Section	Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	No comments
2.1	Participant information/consent documents and consent process	Yes	No comments
3.1	Protocol assessment	Yes	No comments
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	This is a non-commercial single site study taking place in the NHS where that single NHS organisation is also the study sponsor. Therefore no study agreements are expected. Some participants may also be recruited outside the NHS. HRA approval does not cover activity outside the NHS. Before recruiting outside the NHS the research team must follow the procedures and governance arrangements of responsible organisations.
4.2	Insurance/indemnity arrangements assessed	Yes	The project is being conducted by a member of University Staff who holds an honorary contract with Sheffield Teaching Hospitals Trust (STH). The Applicant clarified that as the study involves NHS patients, STH were happy to sponsor the study. As such the design and management is primarily the responsibility of and would be accepted to fall under the clinical trials insurance supplied by the University of Sheffield (UoS). In the event of any issues then the Joint STH/UoS Clinical Research Office would determine whether NHS indemnity or UoS clinical trials insurance would apply.
4.3	Financial arrangements assessed	Yes	This study is receiving funding from the National Institute for Health Research. The funding letter has been provided.
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	No comments
5.3	Compliance with any applicable laws or regulations	Yes	No comments
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	No comments
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	No comments
6.3	Devices – MHRA notice of no objection received	Not Applicable	No comments
6.4	Other regulatory approvals and authorisations received	Not Applicable	No comments



Participating NHS Organisations in England and Wales

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.

This is a non-commercial single site study taking place in the NHS where that single NHS organisation is also the study sponsor. Therefore there is only one site type involved in the research.

If this study is subsequently extended to other NHS organisation(s) in England or Wales, an amendment should be submitted, with a Statement of Activities and Schedule of Events for the newly participating NHS organisation(s) in England or Wales.

The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England and Wales in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the research management function at the participating organisation. Where applicable, the local LCRN contact should also be copied into this correspondence.

If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England and Wales which are not provided in IRAS, the HRA or HCRW websites, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net or HCRW at Research-permissions@wales.nhs.uk. We will work with these organisations to achieve a consistent approach to information provision.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and Wales, and the minimum expectations for education, training and experience that PIs should meet (where applicable).

A Key contact would be expected at this site type to identify potential participants.

A Local Collaborator (LC) would be expected at this site type to facilitate access for the external research team member to undertake activities.

A Principal Investigator (PI) would not be expected as the student will undertake all study activities.

GCP training is not a generic training expectation, in line with the HRA/HCRW/MHRA statement on training expectations.

HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken

No Honorary Research Contracts, Letters of Access or pre-engagement checks are expected for local staff employed by the participating NHS organisations. Where arrangements are not already in place, research staff not employed by the NHS host organisation undertaking any of the research activities listed in the research application would be expected to obtain a Letter of Access based on enhanced DBS checks, including appropriate barred list checks, and occupational health clearance.

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in England and Wales to aid study set-up.

The applicant has indicated that they intend to apply for inclusion on the NIHR CRN Portfolio.

L. Project-specific safeguarding protocol

Project-specific Safeguarding Protocol

Developing a caries-specific child-centred utility measure:

Safeguarding protocol for study participants

1. If child is not brought for clinic appointment: Follow unit's existing policy for risk assessment, contact family, and arrange further appointment or discharge back to referrer with letter, as appropriate
2. If clinician's observations or child/parent comments are made during clinic appointment that give rise to concern: Follow unit's existing processes, speak to consultant on clinic, address concerns with family, initiate System 1 checks, speak to safeguarding team or make immediate referral to social care as appropriate
3. If child participant contacts researchers directly (phone or email) with a matter that raises concern regarding safeguarding: assess whether the child is in immediate danger, discuss with consultant and follow existing safeguarding pathways. Inform child that you have concerns and who you will be speaking to about it.
4. If child participant contacts researchers directly (phone or email) with a matter that does not relate to safeguarding, but may relate to bullying at school, well-being or questions about their treatment: *reassure the child that their comments will be kept private (unless they relate to safeguarding) and discuss with senior colleague to see how best to answer the child's queries. This may be to direct them to a bullying website for help, or to answer their queries about treatment or the research project*

Contacts

1. For safeguarding advice call: 0114 2053535
2. Or ring Sara Thomas/Sharon Clarke for advice on: 65446 (Sheffield Hospitals advice)
3. If can't get through to these or child is out of area then ring social services: Rotherham – 01709823987; Doncaster – 01302734739; Derby – 01332717118; Barnsley – 01226772400; Chesterfield – 08456058058; Bassetlaw Office (Nottinghamshire) - 08449808080

M.Results from factor analysis

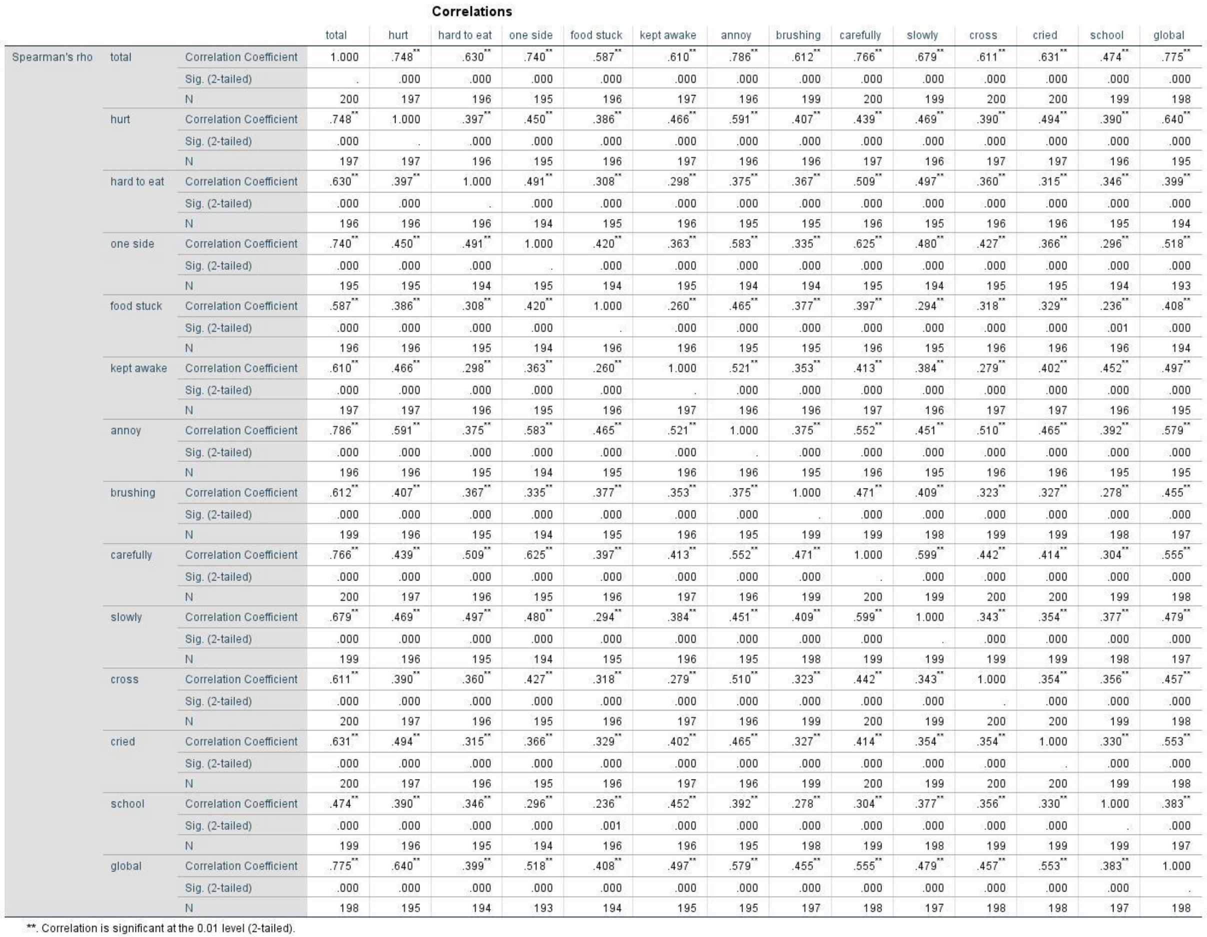
Table 10-3: Results from Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.914
Bartlett's Test of Sphericity	Approx. Chi-Square	910.797
	df	66
	Sig.	0

Table 10-4: Explanation of total variance

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.585	46.54	46.54	5.585	46.54	46.54
2	0.963	8.025	54.566			
3	0.845	7.039	61.605			
4	0.75	6.248	67.853			
5	0.685	5.708	73.561			
6	0.633	5.275	78.836			
7	0.576	4.798	83.634			
8	0.504	4.202	87.836			
9	0.459	3.822	91.658			
10	0.397	3.308	94.966			
11	0.311	2.594	97.56			
12	0.293	2.44	100			
Extraction Method: Principal Component Analysis.						

N. Correlations between items

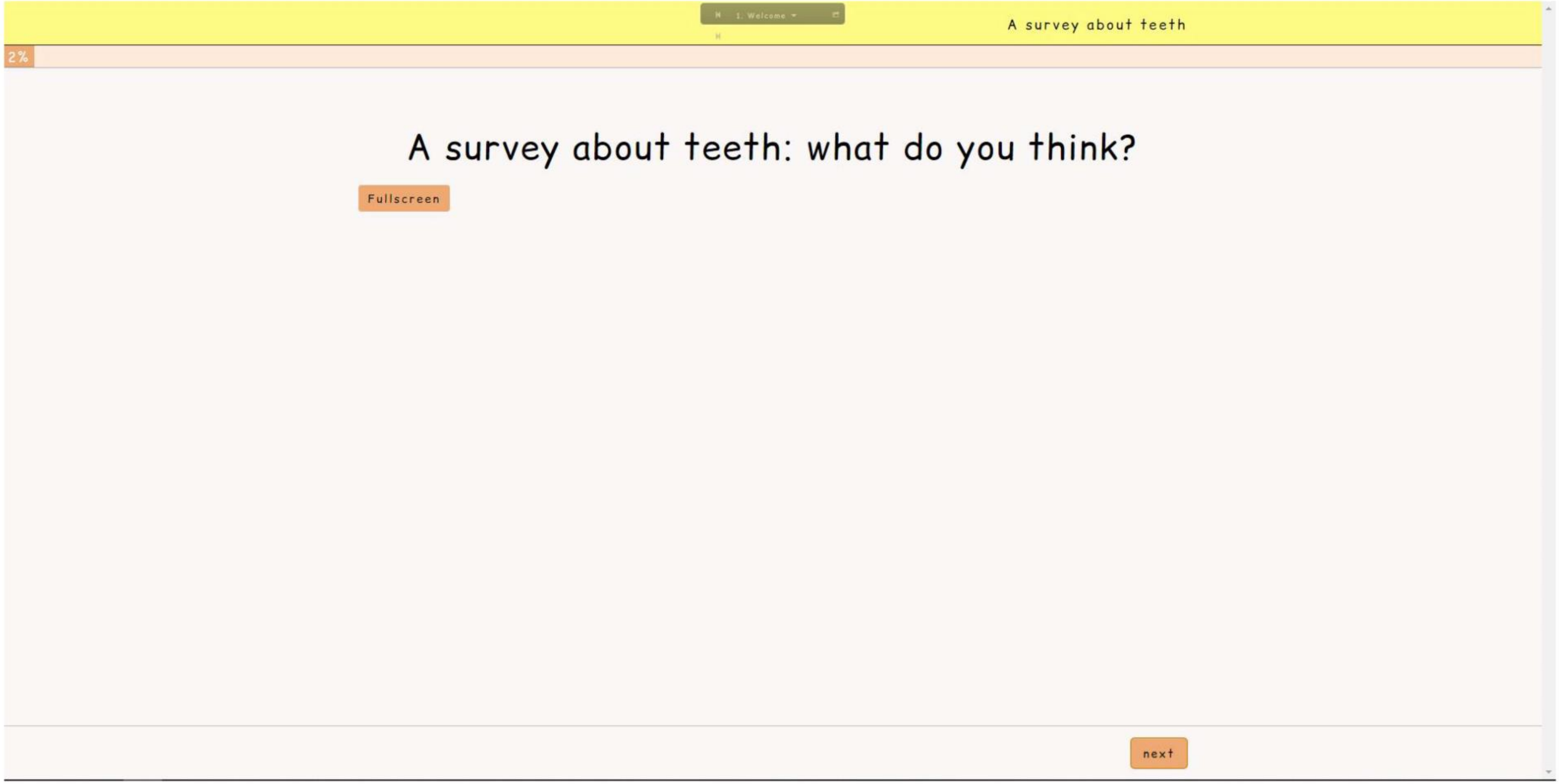


Correlations

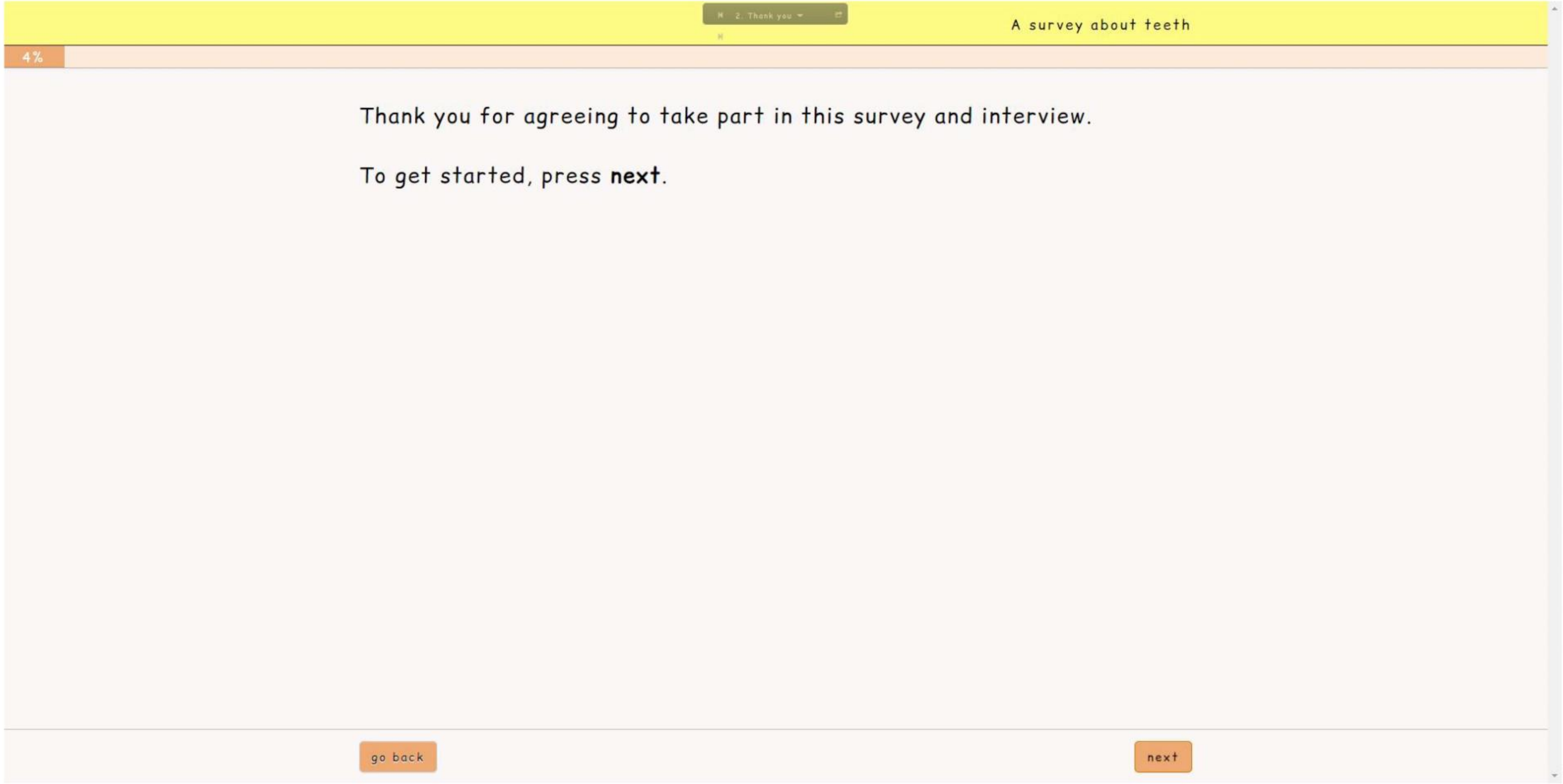
			total	hurt	hard to eat	one side	food stuck	kept awake	annoy	brushing	carefully	slowly	cross	cried	school	global
Spearman's rho	total	Correlation Coefficient	1.000	.748**	.630**	.740**	.587**	.610**	.786**	.612**	.766**	.679**	.611**	.631**	.474**	.775**
		Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
		N	200	197	196	195	196	197	196	199	200	199	200	200	199	198
	hurt	Correlation Coefficient	.748**	1.000	.397**	.450**	.386**	.466**	.591**	.407**	.439**	.469**	.390**	.494**	.390**	.640**
		Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
		N	197	197	196	195	196	197	196	196	197	196	197	197	196	195
	hard to eat	Correlation Coefficient	.630**	.397**	1.000	.491**	.308**	.298**	.375**	.367**	.509**	.497**	.360**	.315**	.346**	.399**
		Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
		N	196	196	196	194	195	196	195	195	196	195	196	196	195	194
	one side	Correlation Coefficient	.740**	.450**	.491**	1.000	.420**	.363**	.583**	.335**	.625**	.480**	.427**	.366**	.296**	.518**
		Sig. (2-tailed)	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
		N	195	195	194	195	194	195	194	194	195	194	195	195	194	193
	food stuck	Correlation Coefficient	.587**	.386**	.308**	.420**	1.000	.260**	.465**	.377**	.397**	.294**	.318**	.329**	.236**	.408**
		Sig. (2-tailed)	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.001	.000
		N	196	196	195	194	196	196	195	195	196	195	196	196	196	194
	kept awake	Correlation Coefficient	.610**	.466**	.298**	.363**	.260**	1.000	.521**	.353**	.413**	.384**	.279**	.402**	.452**	.497**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000
		N	197	197	196	195	196	197	196	196	197	196	197	197	196	195
	annoy	Correlation Coefficient	.786**	.591**	.375**	.583**	.465**	.521**	1.000	.375**	.552**	.451**	.510**	.465**	.392**	.579**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000
		N	196	196	195	194	195	196	196	195	196	195	196	196	195	195
	brushing	Correlation Coefficient	.612**	.407**	.367**	.335**	.377**	.353**	.375**	1.000	.471**	.409**	.323**	.327**	.278**	.455**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000
		N	199	196	195	194	195	196	195	199	199	198	199	199	198	197
	carefully	Correlation Coefficient	.766**	.439**	.509**	.625**	.397**	.413**	.552**	.471**	1.000	.599**	.442**	.414**	.304**	.555**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000
		N	200	197	196	195	196	197	196	199	200	199	200	200	199	198
	slowly	Correlation Coefficient	.679**	.469**	.497**	.480**	.294**	.384**	.451**	.409**	.599**	1.000	.343**	.354**	.377**	.479**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000
		N	199	196	195	194	195	196	195	198	199	199	199	199	198	197
	cross	Correlation Coefficient	.611**	.390**	.360**	.427**	.318**	.279**	.510**	.323**	.442**	.343**	1.000	.354**	.356**	.457**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000
		N	200	197	196	195	196	197	196	199	200	199	200	200	199	198
	cried	Correlation Coefficient	.631**	.494**	.315**	.366**	.329**	.402**	.465**	.327**	.414**	.354**	.354**	1.000	.330**	.553**
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000
		N	200	197	196	195	196	197	196	199	200	199	200	200	199	198
	school	Correlation Coefficient	.474**	.390**	.346**	.296**	.236**	.452**	.392**	.278**	.304**	.377**	.356**	.330**	1.000	.383**
		Sig. (2-tailed)	.000	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000	.	.000
		N	199	196	195	194	196	196	195	198	199	198	199	199	199	197
	global	Correlation Coefficient	.775**	.640**	.399**	.518**	.408**	.497**	.579**	.455**	.555**	.479**	.457**	.553**	.383**	1.000
		Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.
		N	198	195	194	193	194	195	195	197	198	197	198	198	197	198

**. Correlation is significant at the 0.01 level (2-tailed).

O. Survey developed for DCE vs BWS study

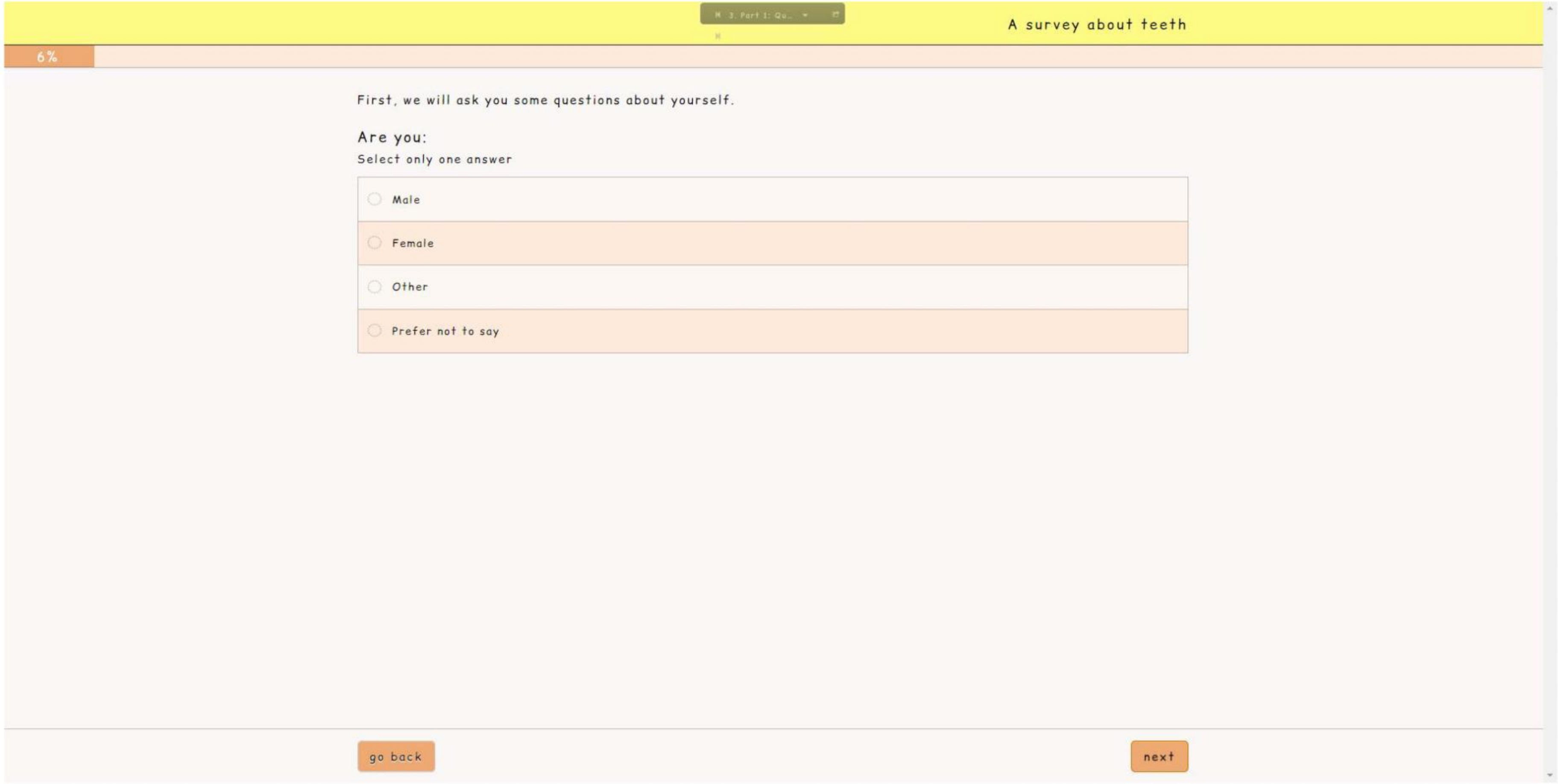


A survey about teeth: what do you think?



Thank you for agreeing to take part in this survey and interview.

To get started, press **next**.



A survey about teeth

6%

First, we will ask you some questions about yourself.

Are you:

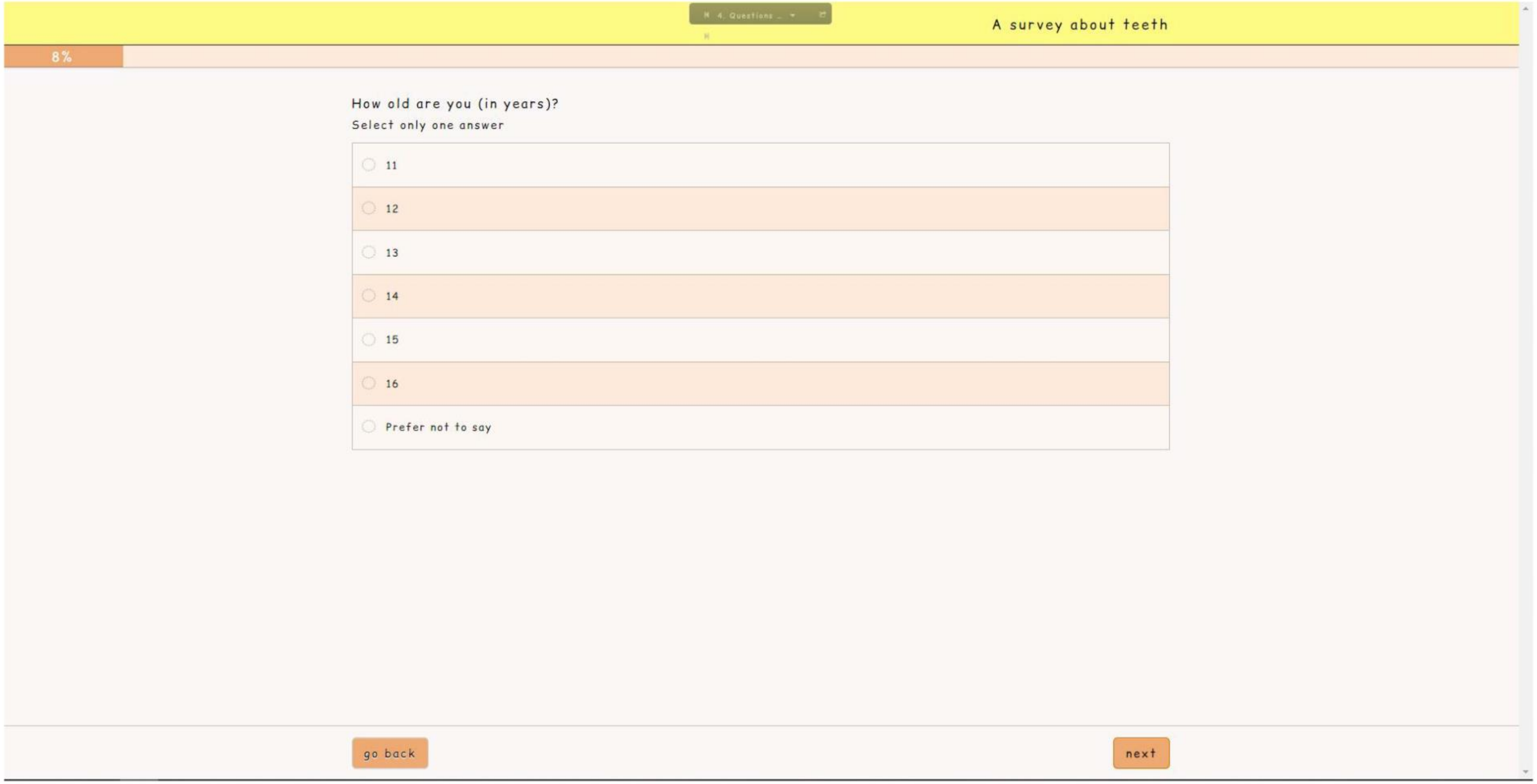
Select only one answer

○ Male

○ Female

○ Other

○ Prefer not to say



How old are you (in years)?

Select only one answer

○ 11

○ 12

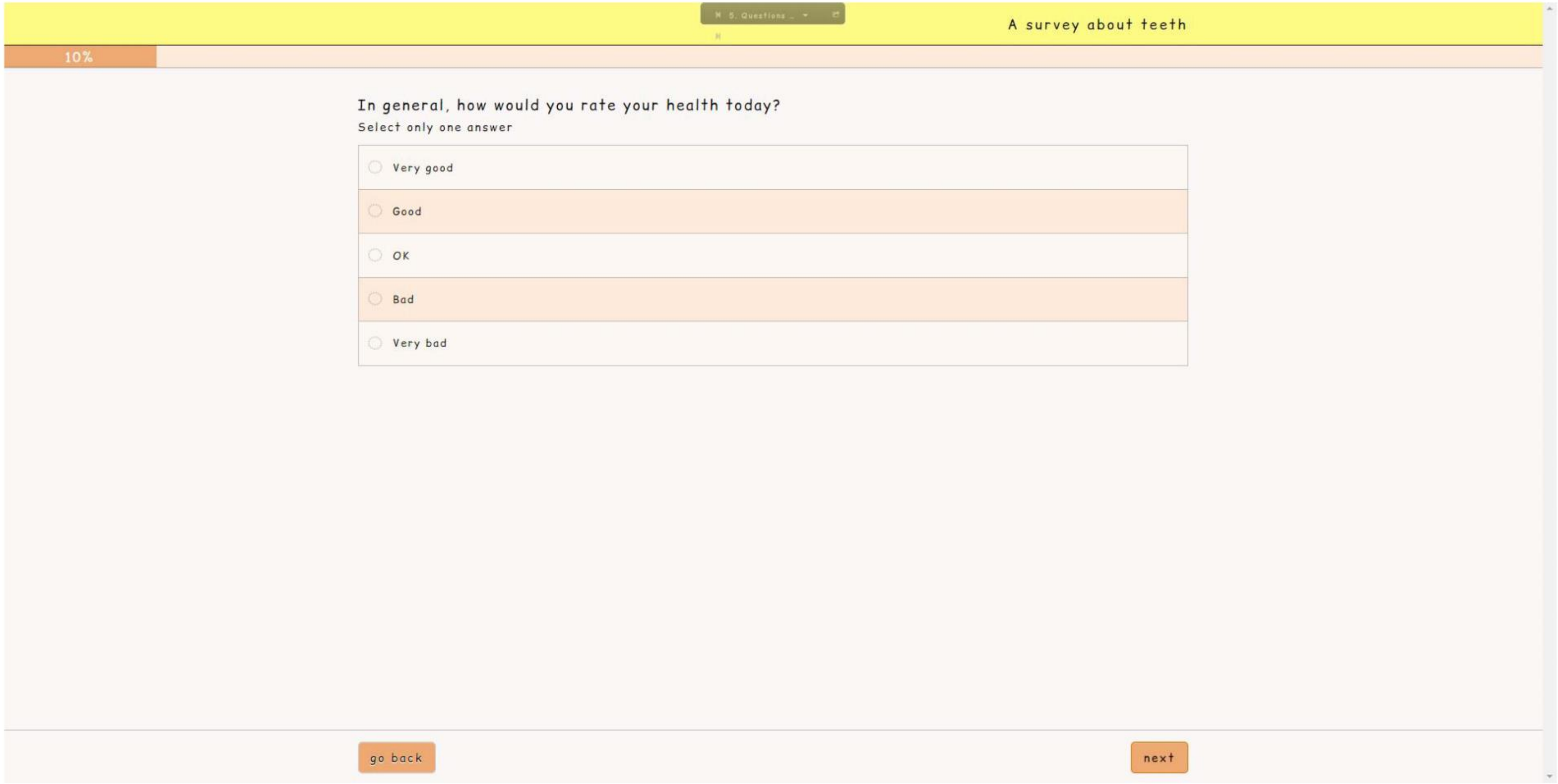
○ 13

○ 14

○ 15

○ 16

○ Prefer not to say



In general, how would you rate your health today?

Select only one answer

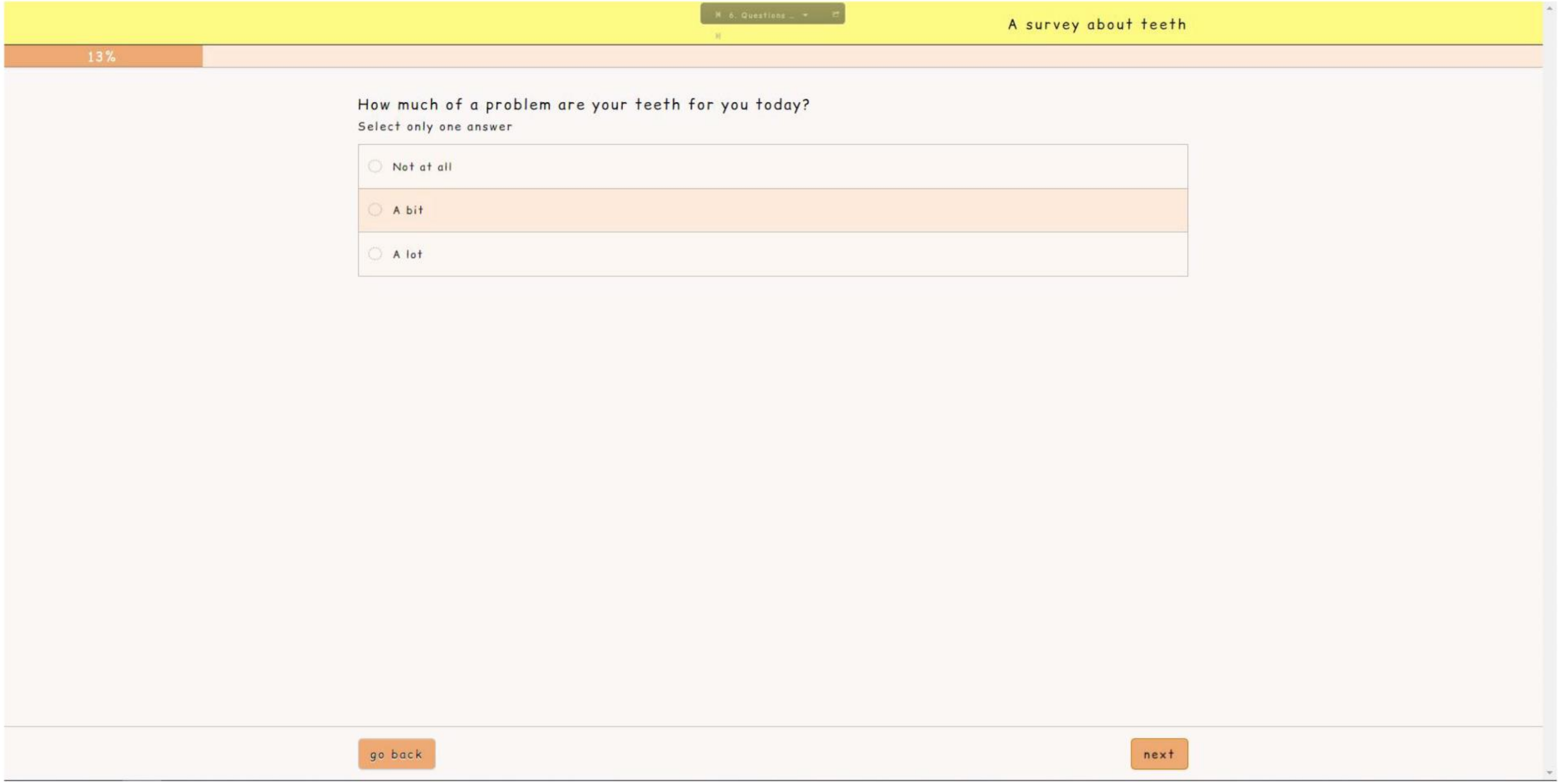
○ Very good

○ Good

○ OK

○ Bad

○ Very bad



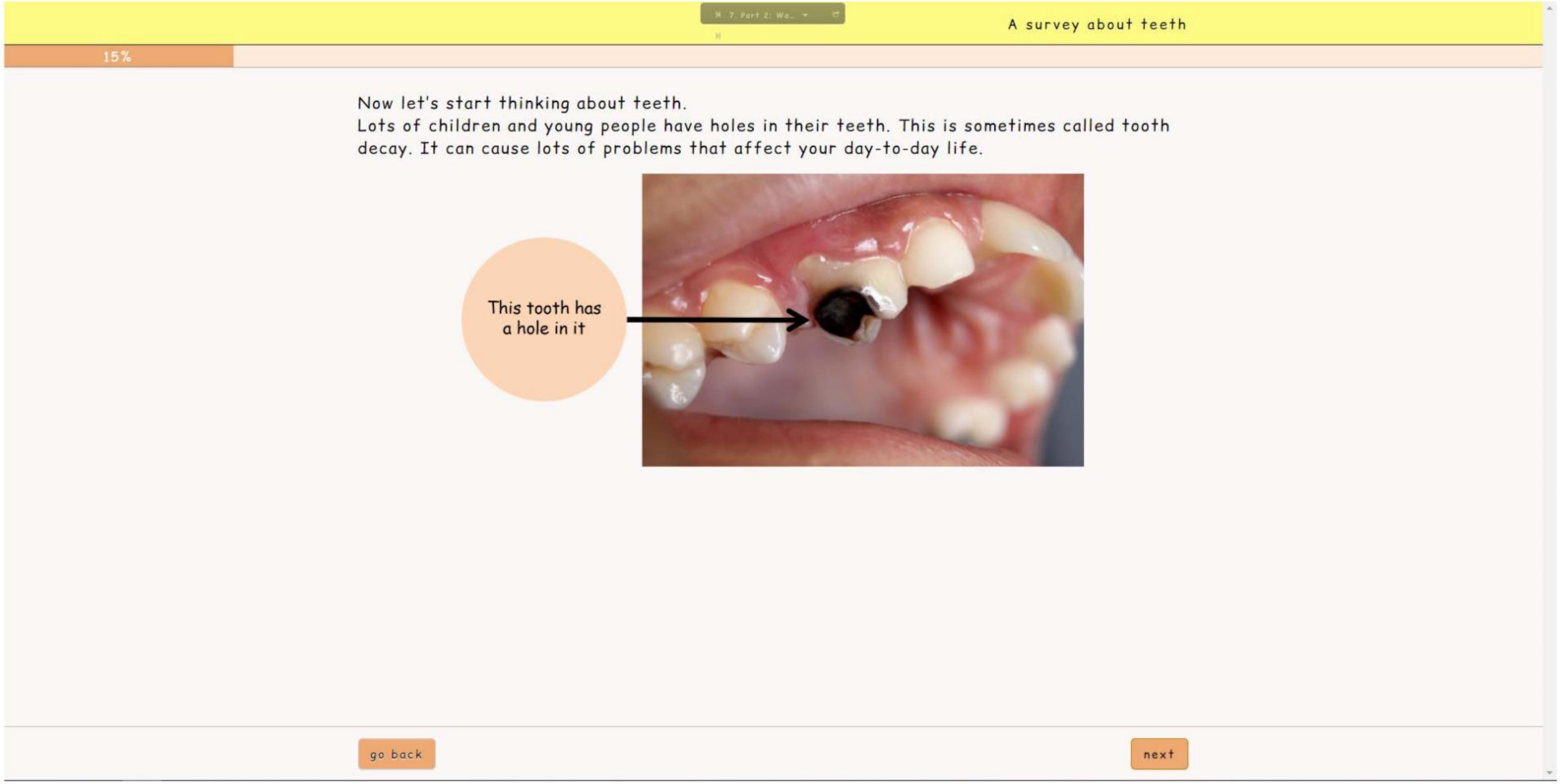
How much of a problem are your teeth for you today?

Select only one answer

○ Not at all

○ A bit

○ A lot



A survey about teeth

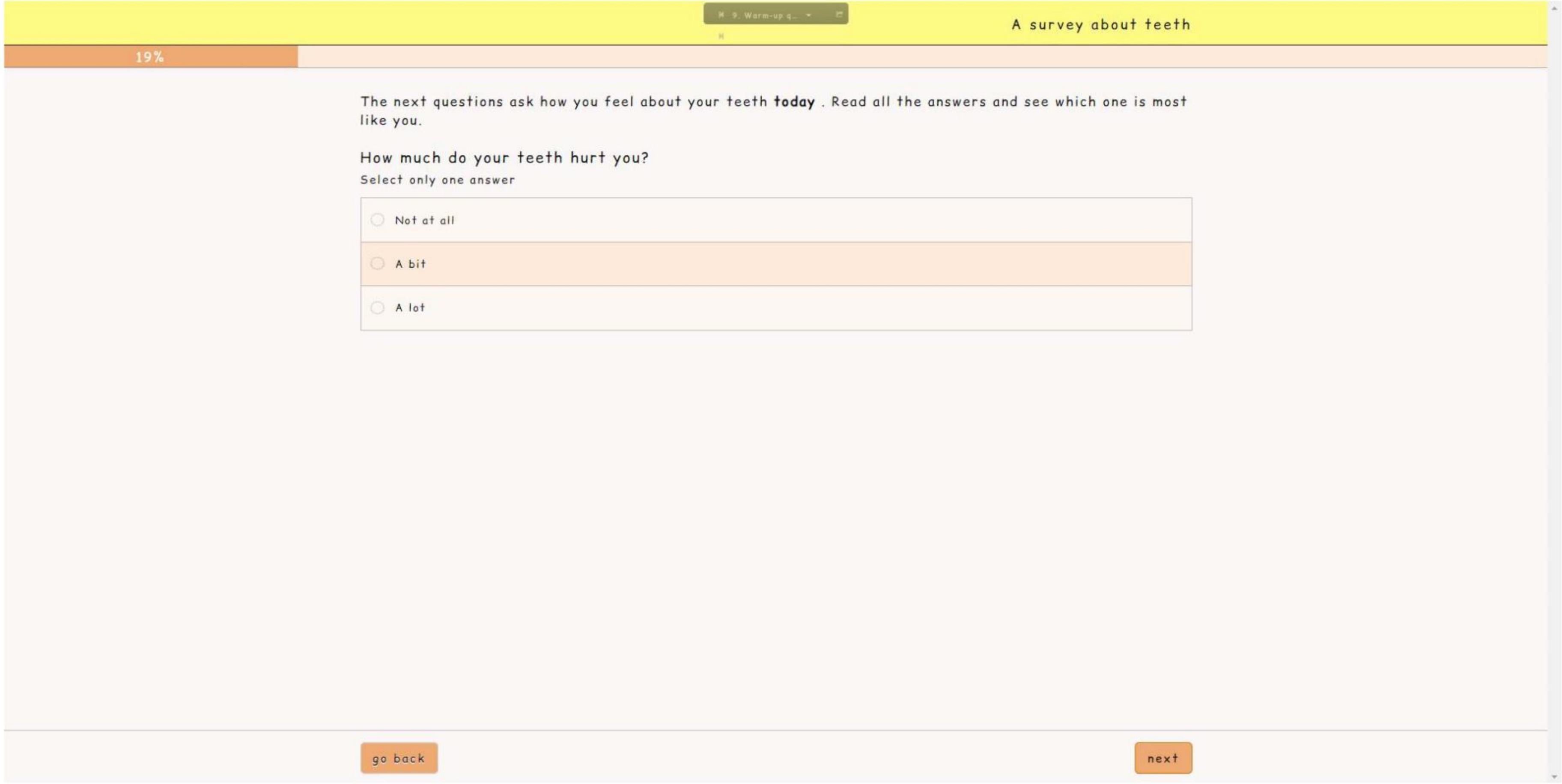
Now let's start thinking about teeth.
Lots of children and young people have holes in their teeth. This is sometimes called tooth decay. It can cause lots of problems that affect your day-to-day life.

This tooth has a hole in it

Tooth decay (holes) can make your teeth hurt. This can be very annoying and can make you feel frustrated.

Sometimes they can cause a lot of pain that wakes you up in the night or makes it hard to get to sleep.

Having holes in your teeth can affect what you eat, as it can be difficult to eat some types of food. Sometimes these problems can make you feel very sad and can make you cry.



The next questions ask how you feel about your teeth **today** . Read all the answers and see which one is most like you.

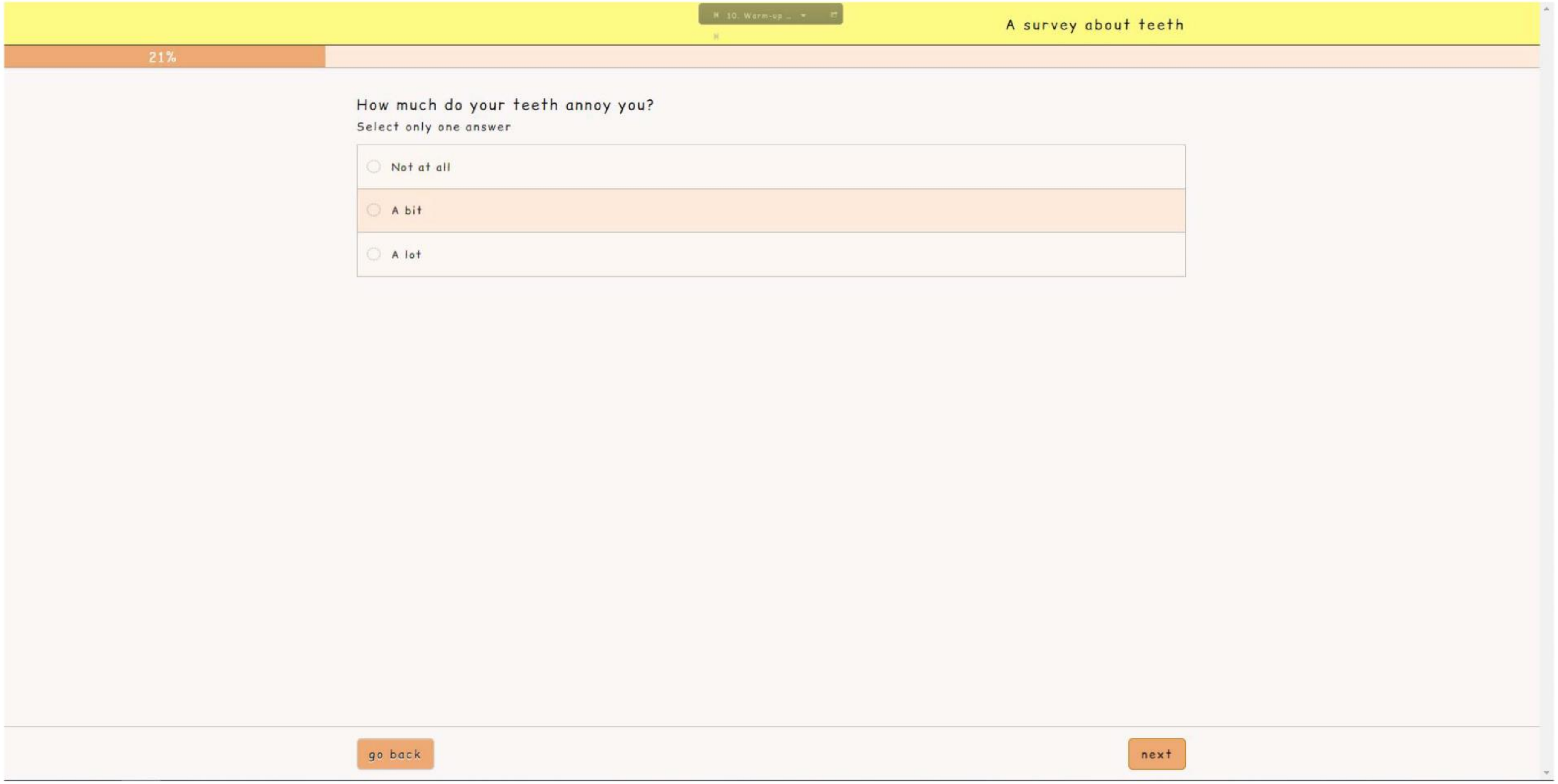
How much do your teeth hurt you?

Select only one answer

○ Not at all

○ A bit

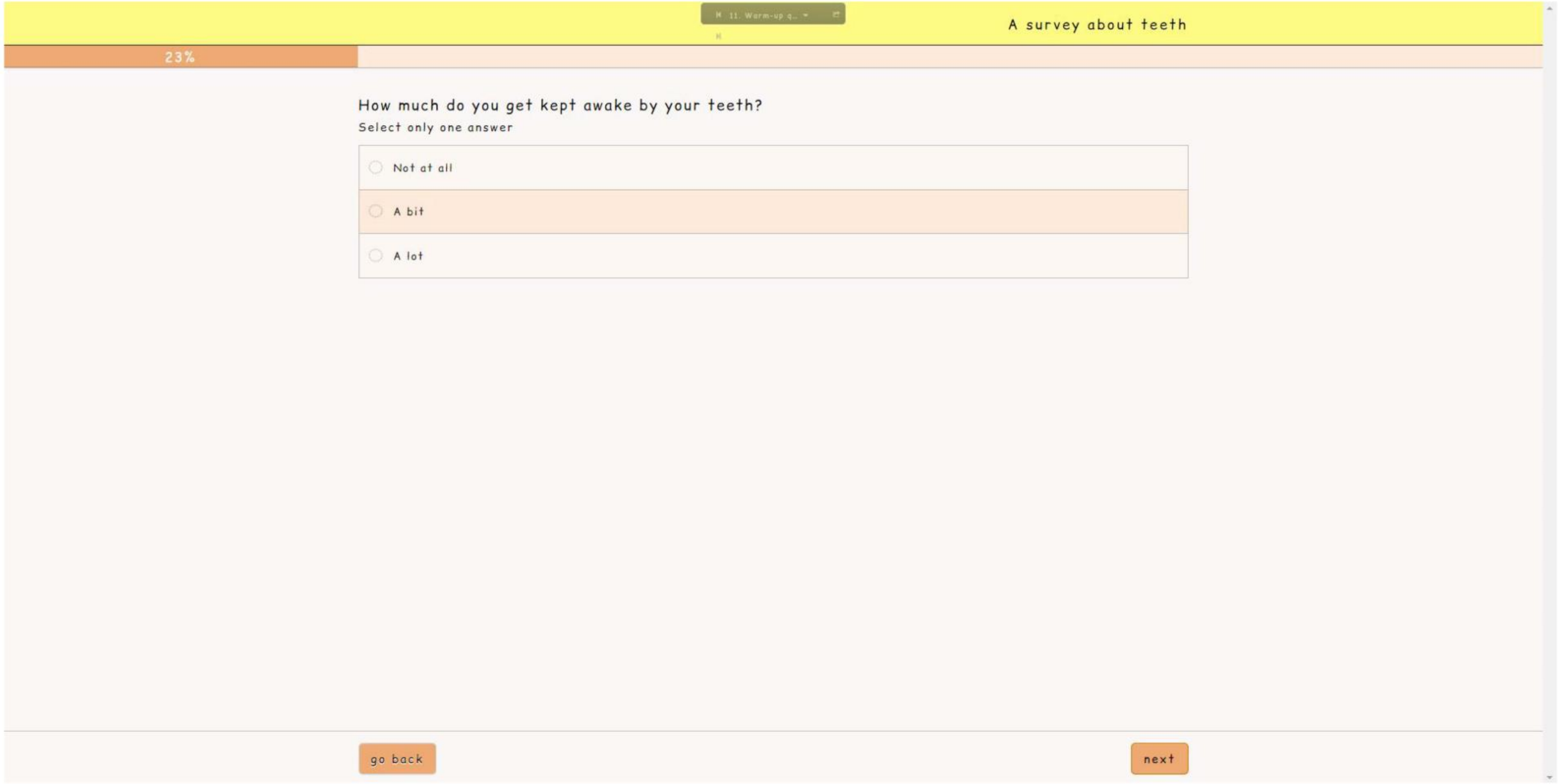
○ A lot



How much do your teeth annoy you?

Select only one answer

○ Not at all
○ A bit
○ A lot



A survey about teeth

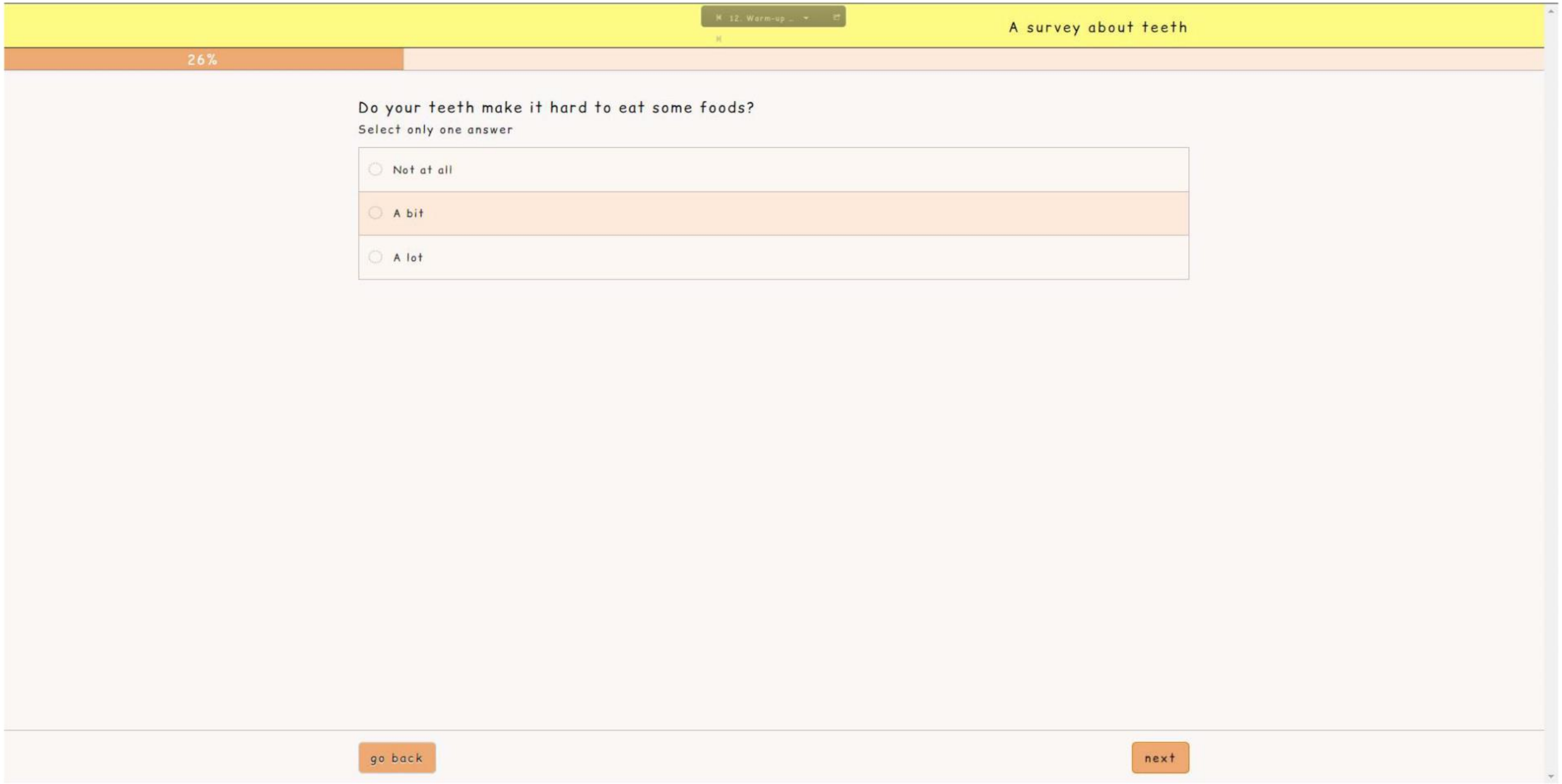
How much do you get kept awake by your teeth?

Select only one answer

○ Not at all

○ A bit

○ A lot



Do your teeth make it hard to eat some foods?

Select only one answer

○ Not at all

○ A bit

○ A lot

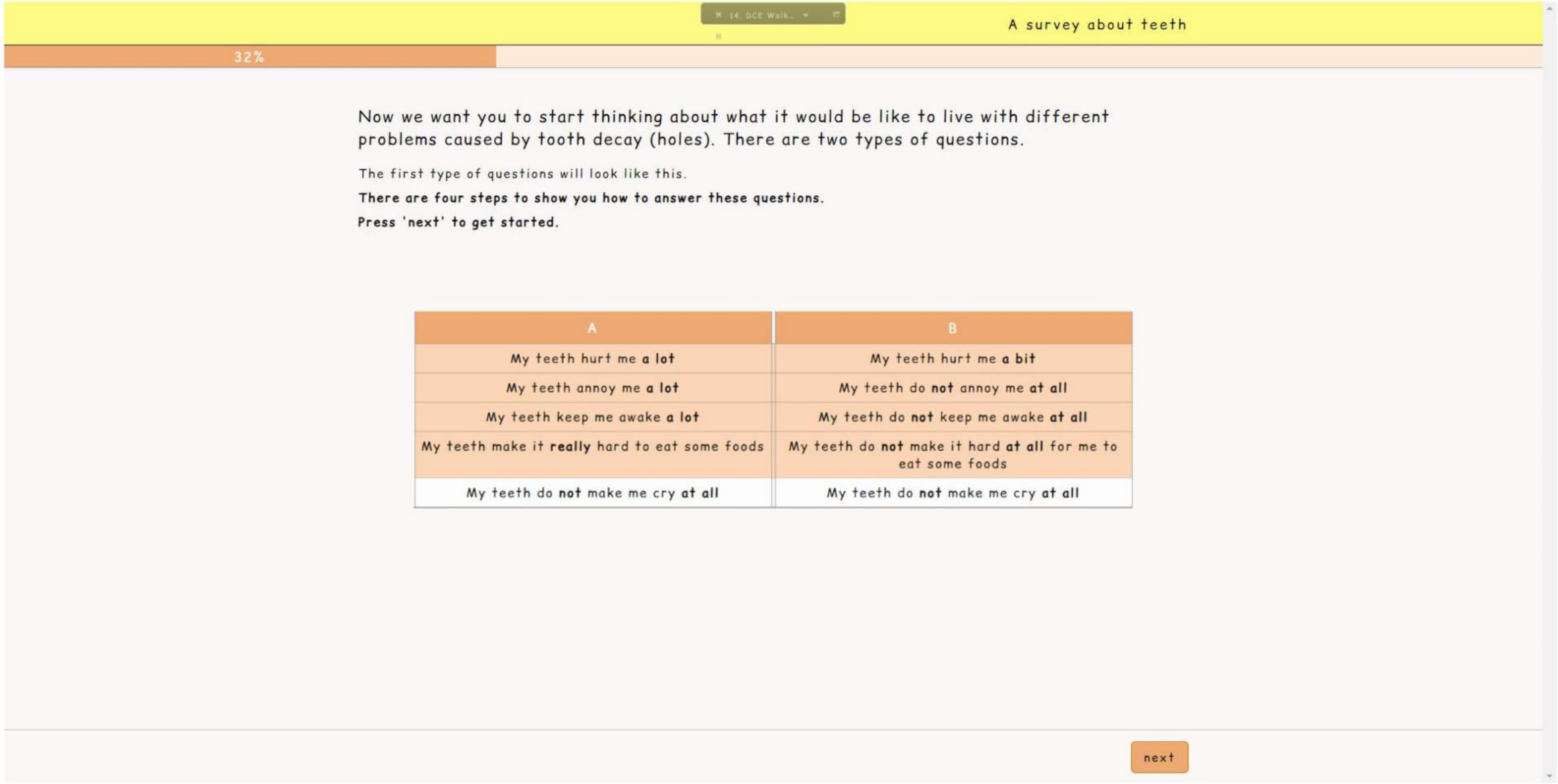
How much have you cried because of your teeth?

Select only one answer

○ Not at all

○ A bit

○ A lot



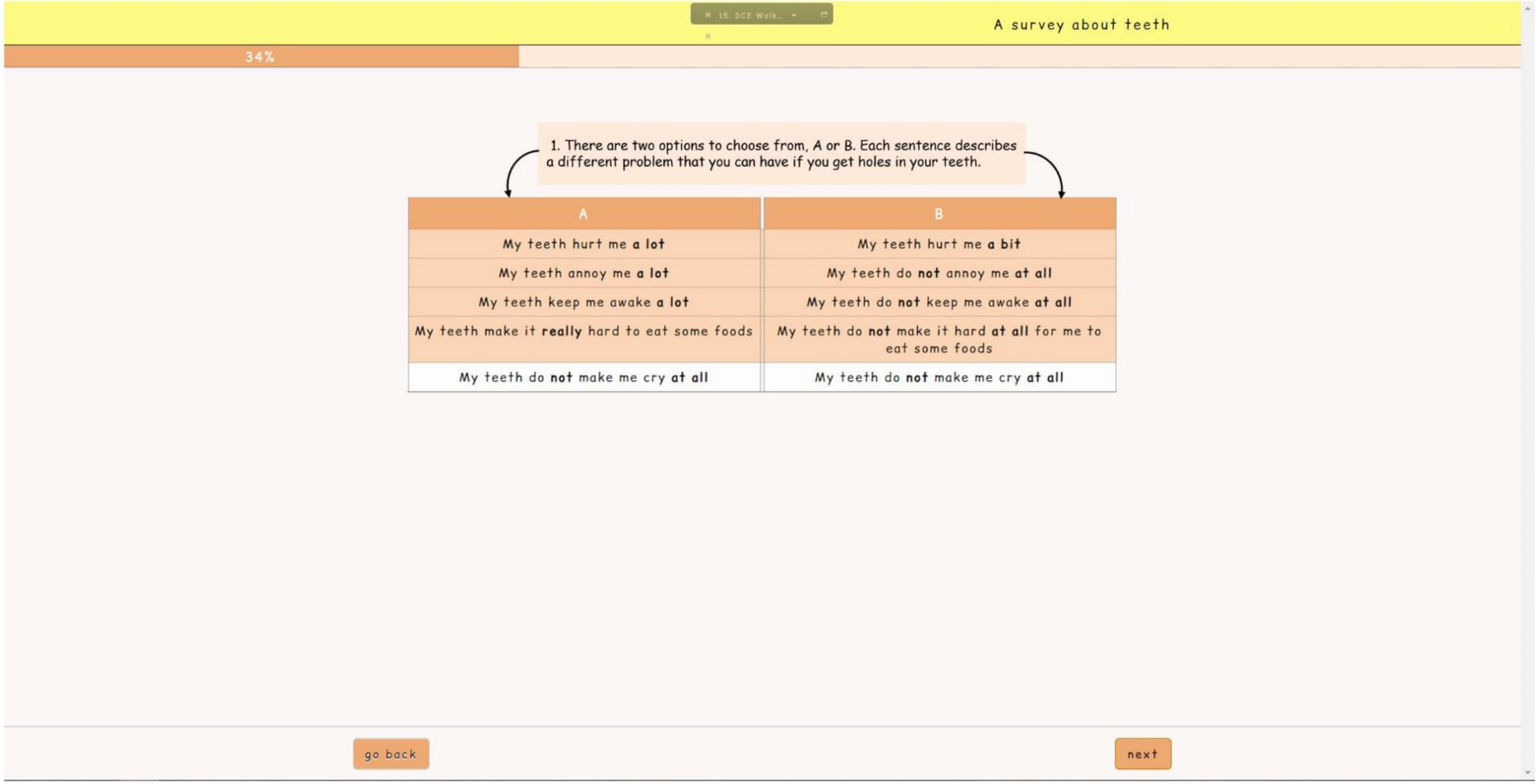
Now we want you to start thinking about what it would be like to live with different problems caused by tooth decay (holes). There are two types of questions.

The first type of questions will look like this.

There are four steps to show you how to answer these questions.

Press 'next' to get started.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**



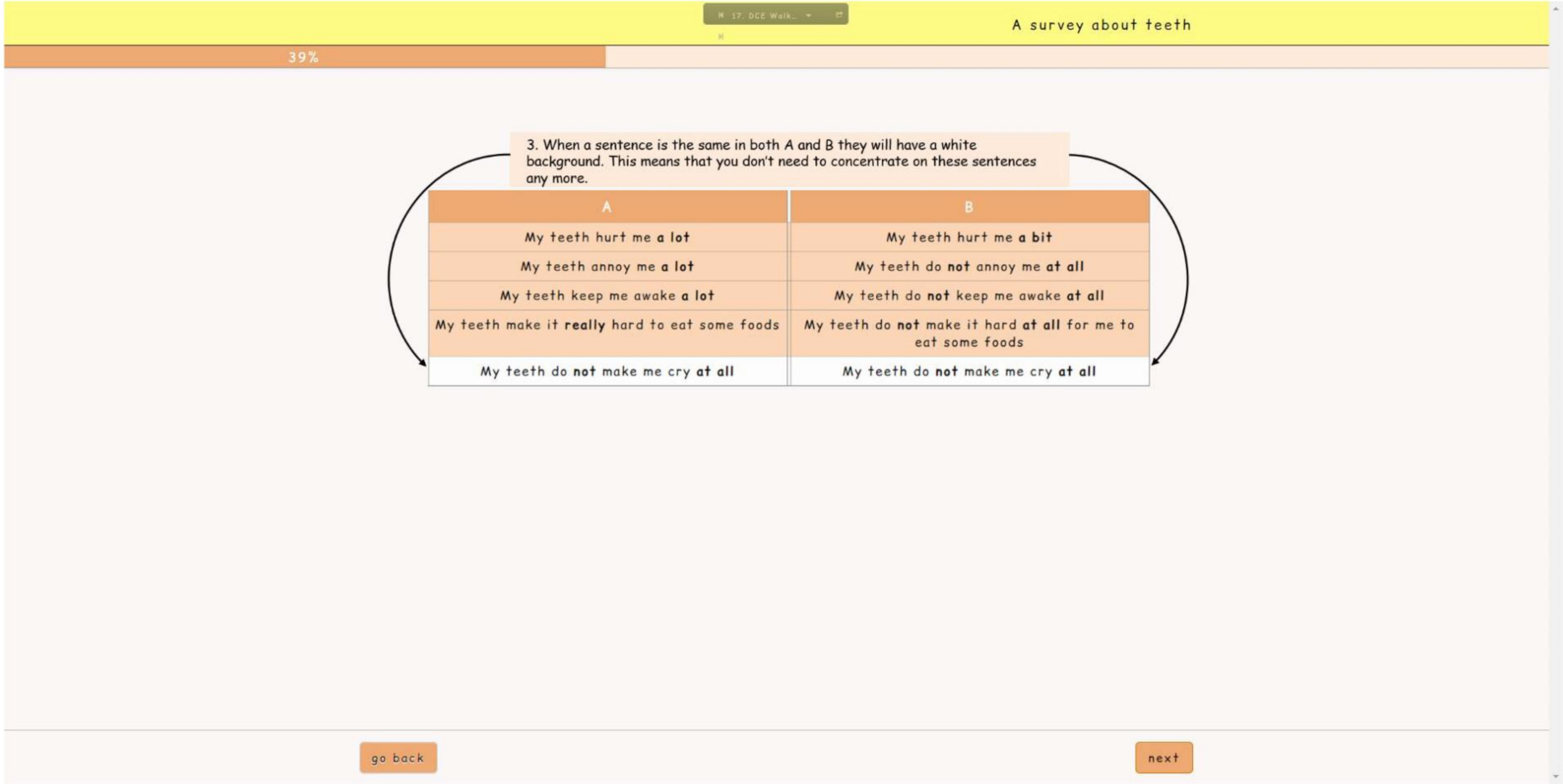
1. There are two options to choose from, A or B. Each sentence describes a different problem that you can have if you get holes in your teeth.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**

A survey about teeth

2. The words **in bold** describe how bad the problem is.

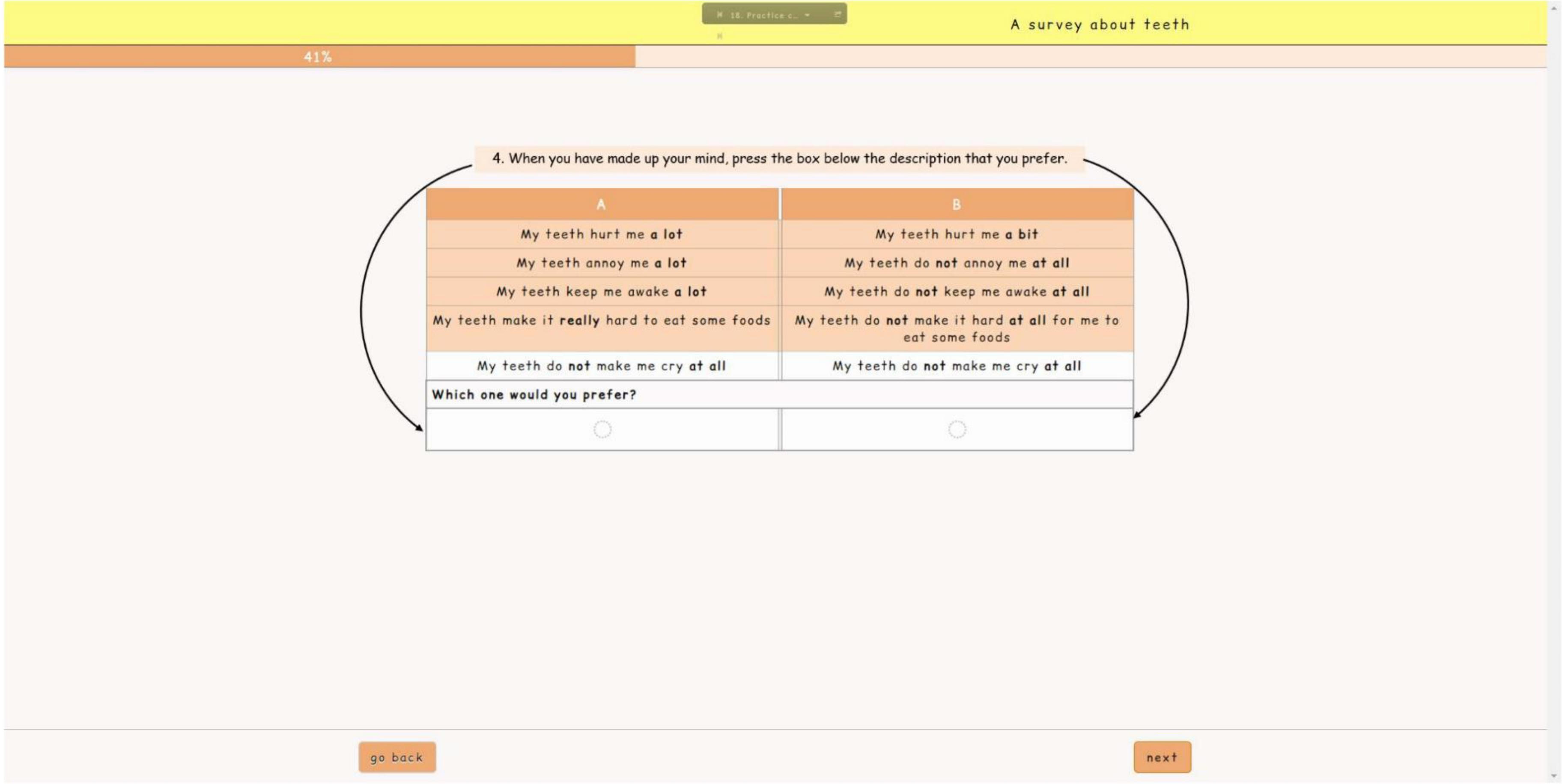
A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**



A survey about teeth

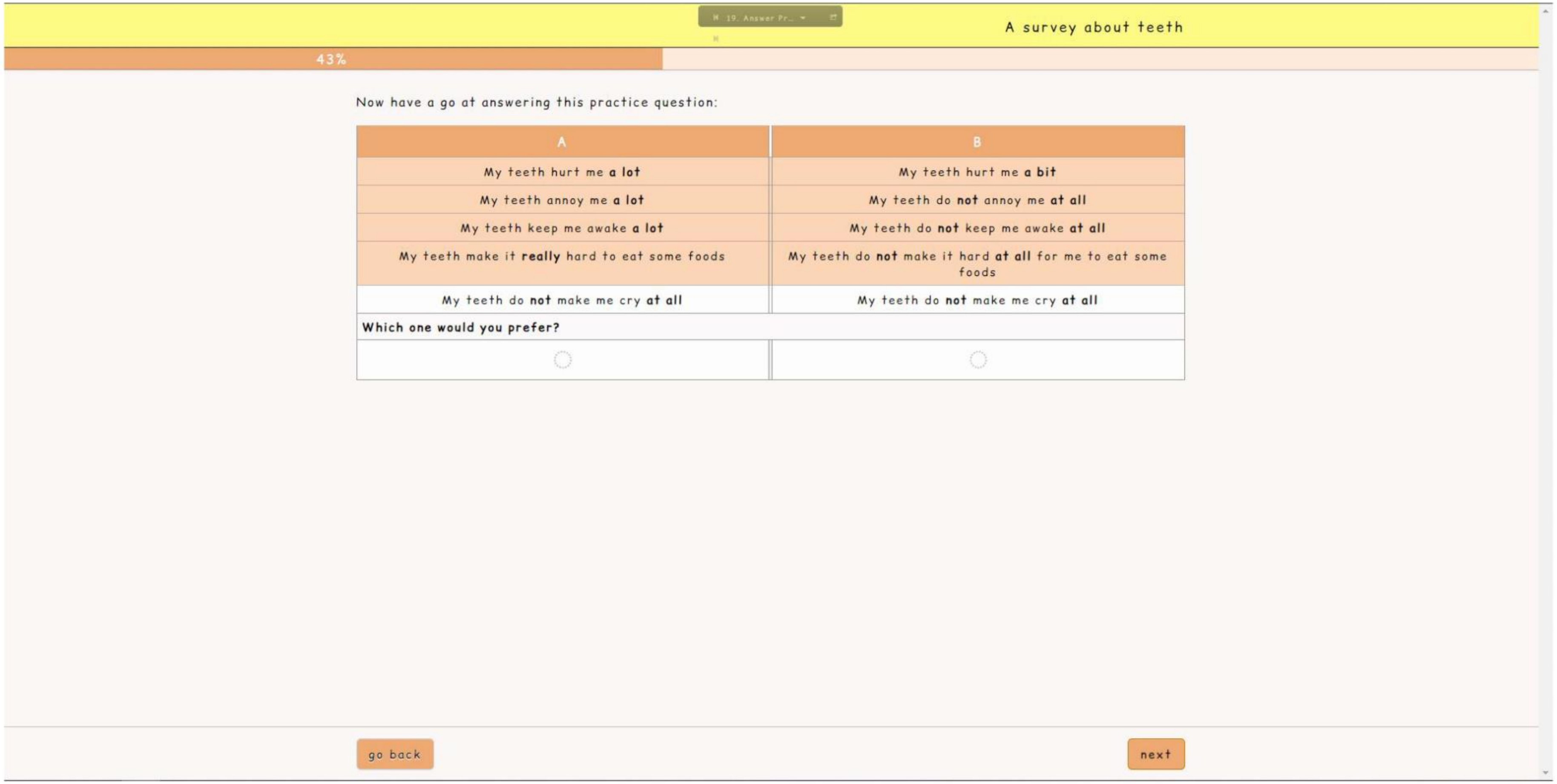
3. When a sentence is the same in both A and B they will have a white background. This means that you don't need to concentrate on these sentences any more.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**



4. When you have made up your mind, press the box below the description that you prefer.

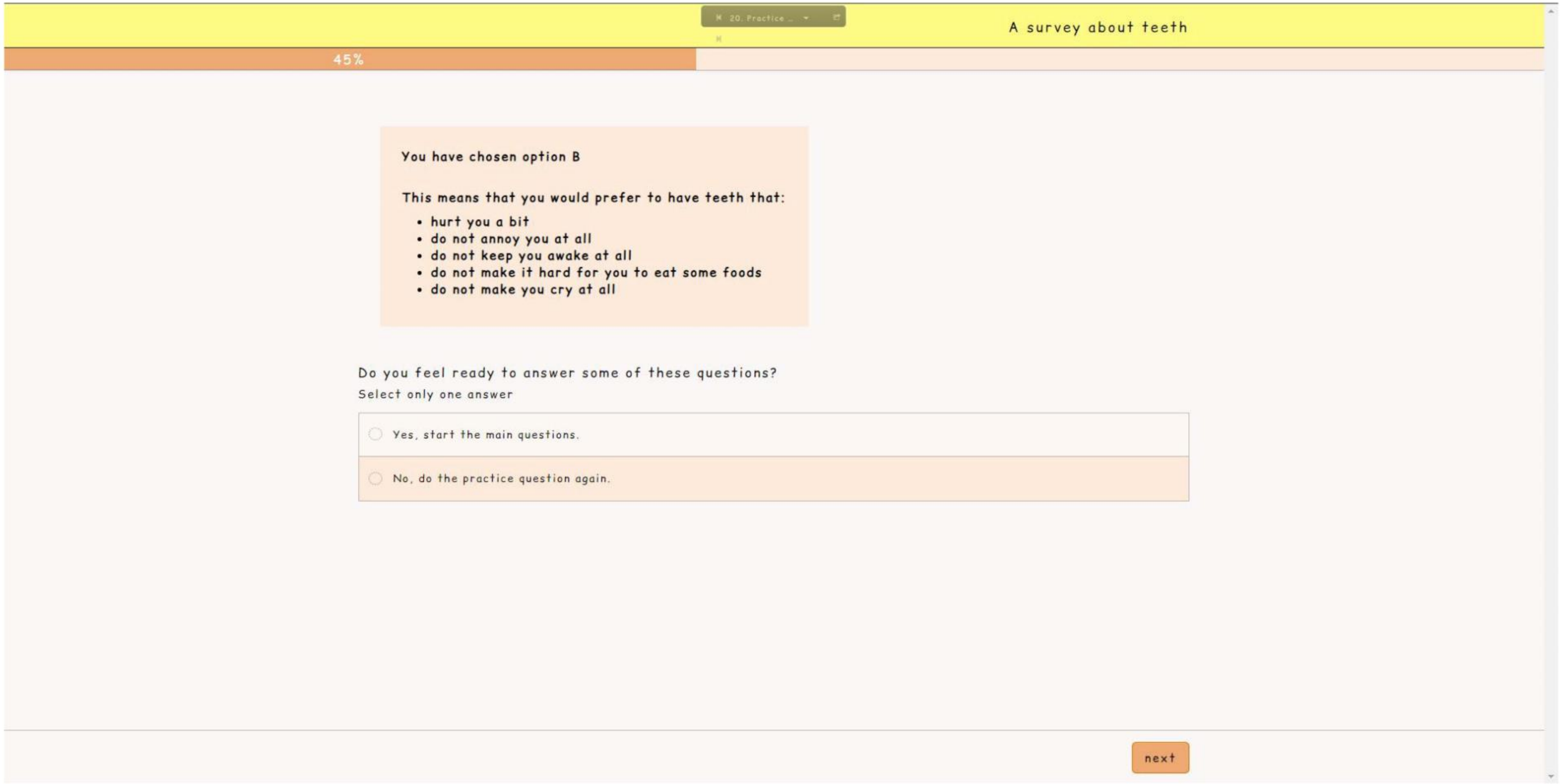
A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**
Which one would you prefer?	
○	○



A survey about teeth

Now have a go at answering this practice question:

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**
Which one would you prefer?	
○	○



You have chosen option B

This means that you would prefer to have teeth that:

- hurt you a bit
- do not annoy you at all
- do not keep you awake at all
- do not make it hard for you to eat some foods
- do not make you cry at all

Do you feel ready to answer some of these questions?

Select only one answer

○ Yes, start the main questions.

○ No, do the practice question again.

A survey about teeth

Please read the descriptions below:

A	B
My teeth hurt me **a bit**	My teeth hurt me **a lot**
My teeth annoy me **a lot**	My teeth annoy me **a bit**
My teeth keep me awake **a lot**	My teeth keep me awake **a bit**
My teeth make it **a bit** hard to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth make me cry **a bit**	My teeth make me cry **a bit**
Which one would you prefer?	
○	○

Please read the descriptions below:

A	B
My teeth do **not** hurt me **at all**	My teeth hurt me **a lot**
My teeth do **not** annoy me **at all**	My teeth annoy me **a lot**
My teeth do **not** keep me awake **at all**	My teeth keep me awake **a lot**
My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make it **really** hard to eat some foods
My teeth do **not** make me cry **at all**	My teeth make me cry **a lot**

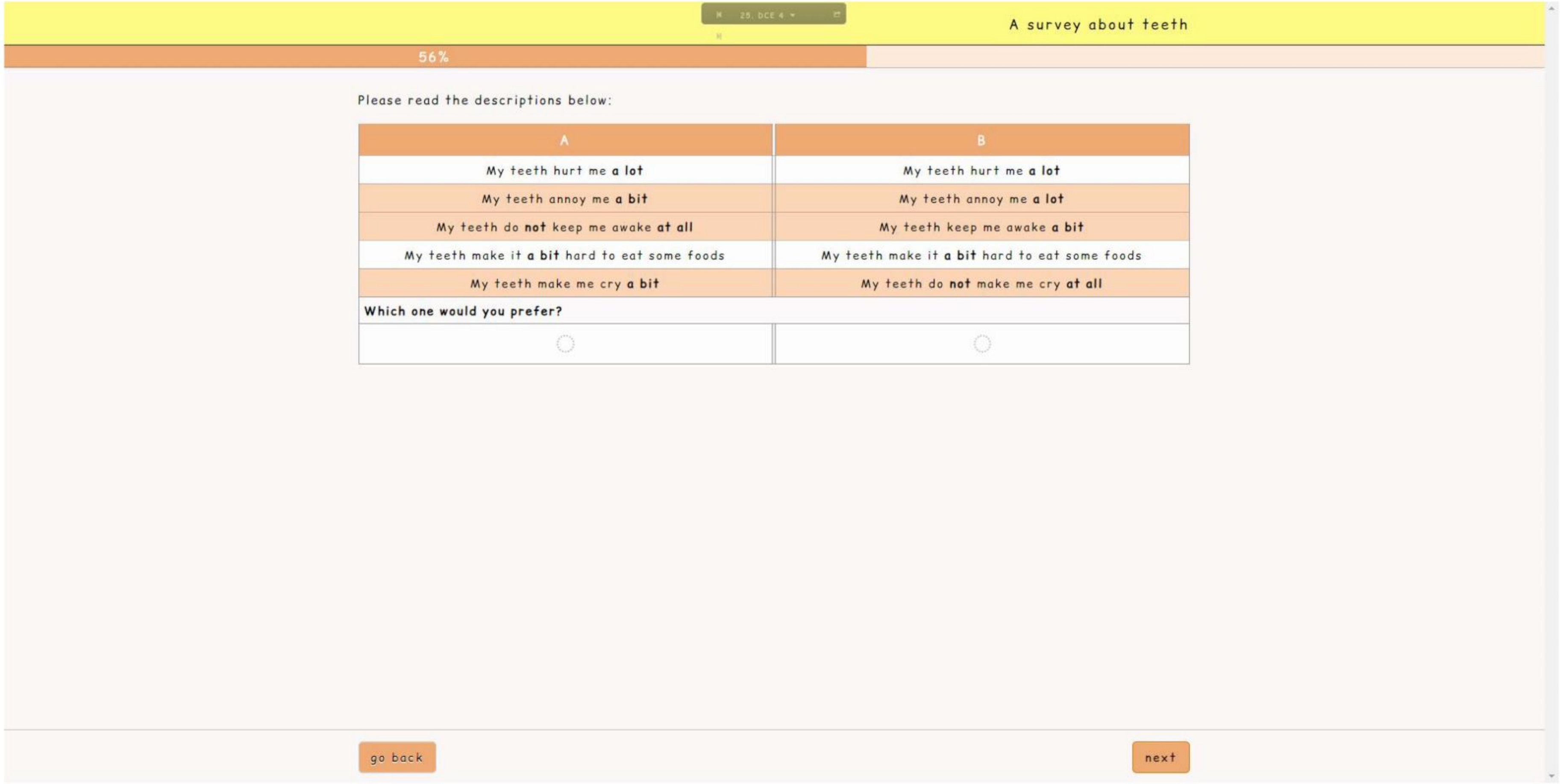
Which one would you prefer?

○	○

A survey about teeth

Please read the descriptions below:

A	B
My teeth do **not** hurt me **at all**	My teeth hurt me **a lot**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth do **not** keep me awake **at all**	My teeth keep me awake **a bit**
My teeth make it **a bit** hard to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth do **not** make me cry **at all**	My teeth make me cry **a lot**
Which one would you prefer?	
○	○



A survey about teeth

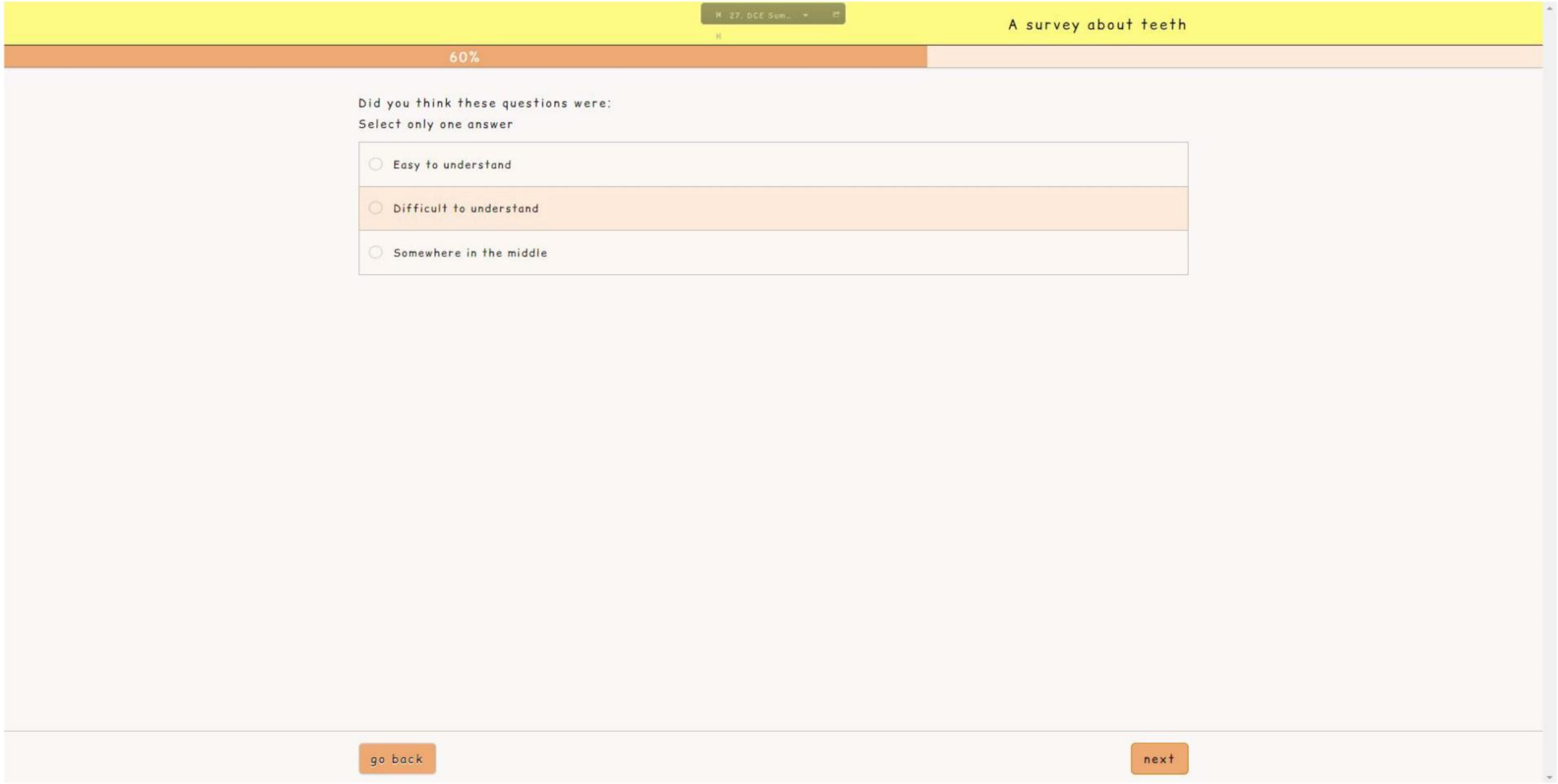
Please read the descriptions below:

A	B
My teeth hurt me **a lot**	My teeth hurt me **a lot**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth do **not** keep me awake **at all**	My teeth keep me awake **a bit**
My teeth make it **a bit** hard to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
Which one would you prefer?	
○	○

A survey about teeth

Please read the descriptions below:

A	B
My teeth do **not** hurt me **at all**	My teeth hurt me **a bit**
My teeth annoy me **a bit**	My teeth annoy me **a bit**
My teeth do **not** keep me awake **at all**	My teeth keep me awake **a bit**
My teeth make it **a bit** hard to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
Which one would you prefer?	
○	○



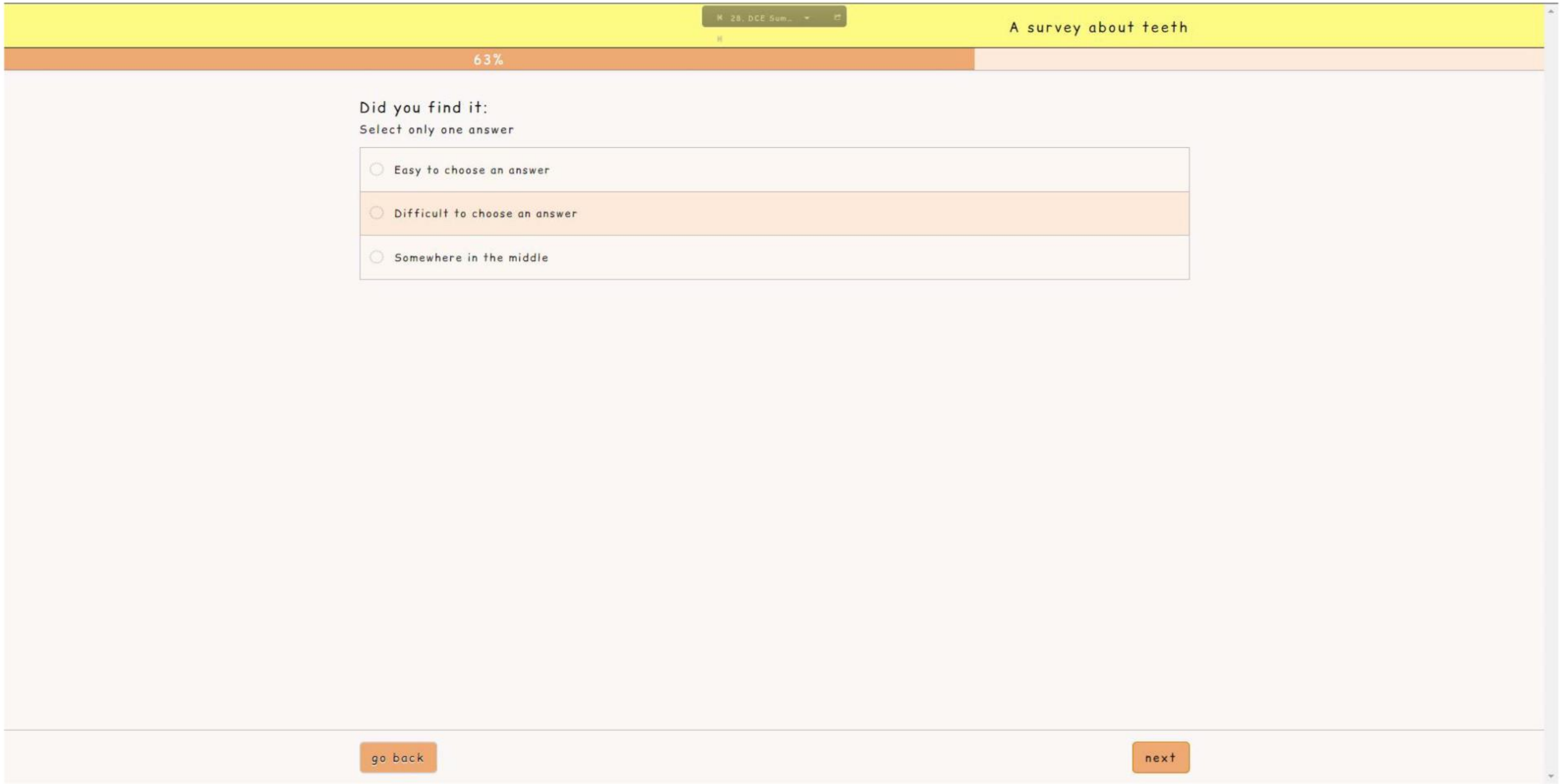
Did you think these questions were:

Select only one answer

○ Easy to understand

○ Difficult to understand

○ Somewhere in the middle



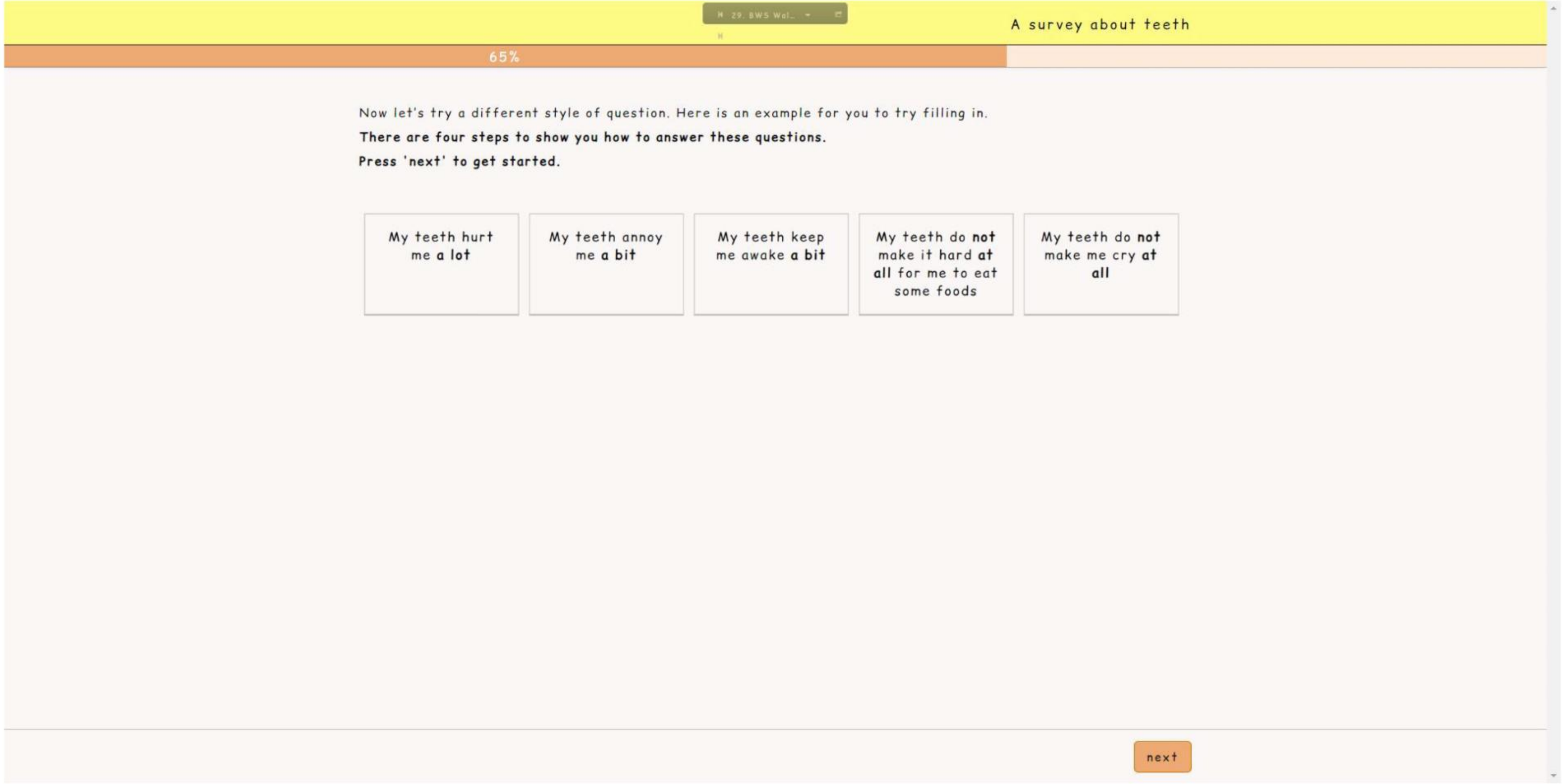
Did you find it:

Select only one answer

○ Easy to choose an answer

○ Difficult to choose an answer

○ Somewhere in the middle

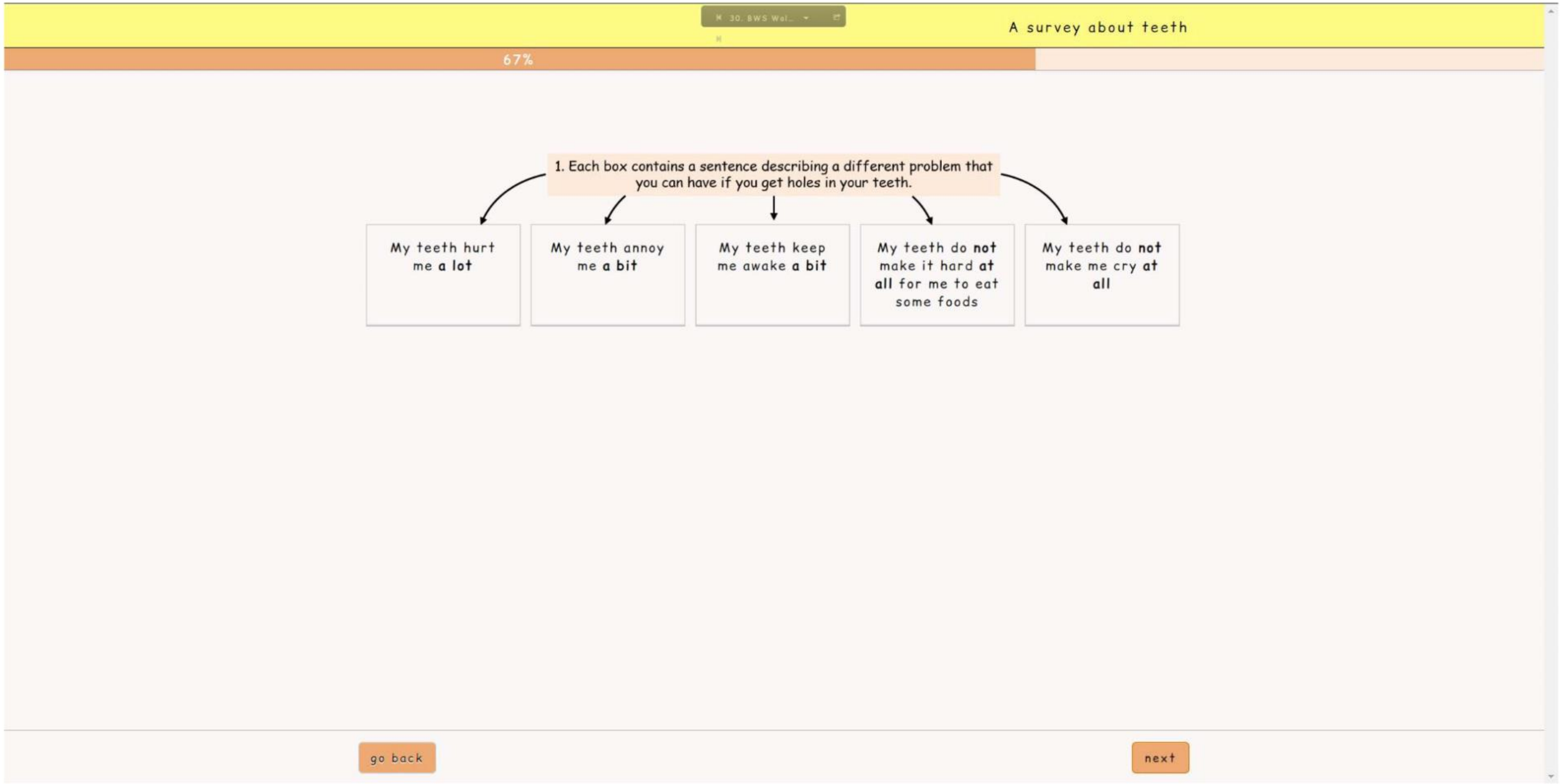


Now let's try a different style of question. Here is an example for you to try filling in.

There are four steps to show you how to answer these questions.

Press 'next' to get started.

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**



A survey about teeth

1. Each box contains a sentence describing a different problem that you can have if you get holes in your teeth.

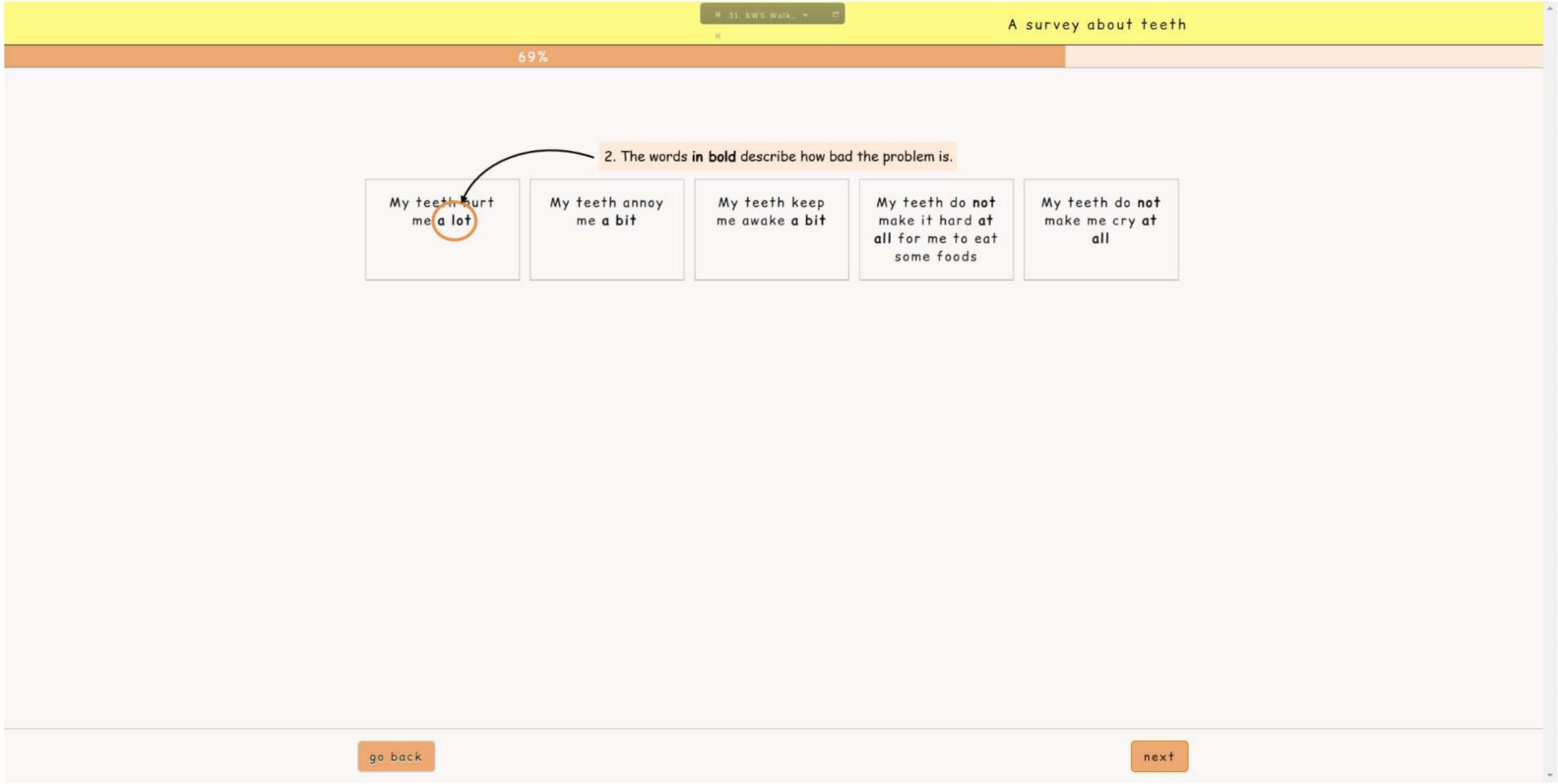
My teeth hurt me **a lot**

My teeth annoy me **a bit**

My teeth keep me awake **a bit**

My teeth do **not** make it hard **at all** for me to eat some foods

My teeth do **not** make me cry **at all**



2. The words **in bold** describe how bad the problem is.

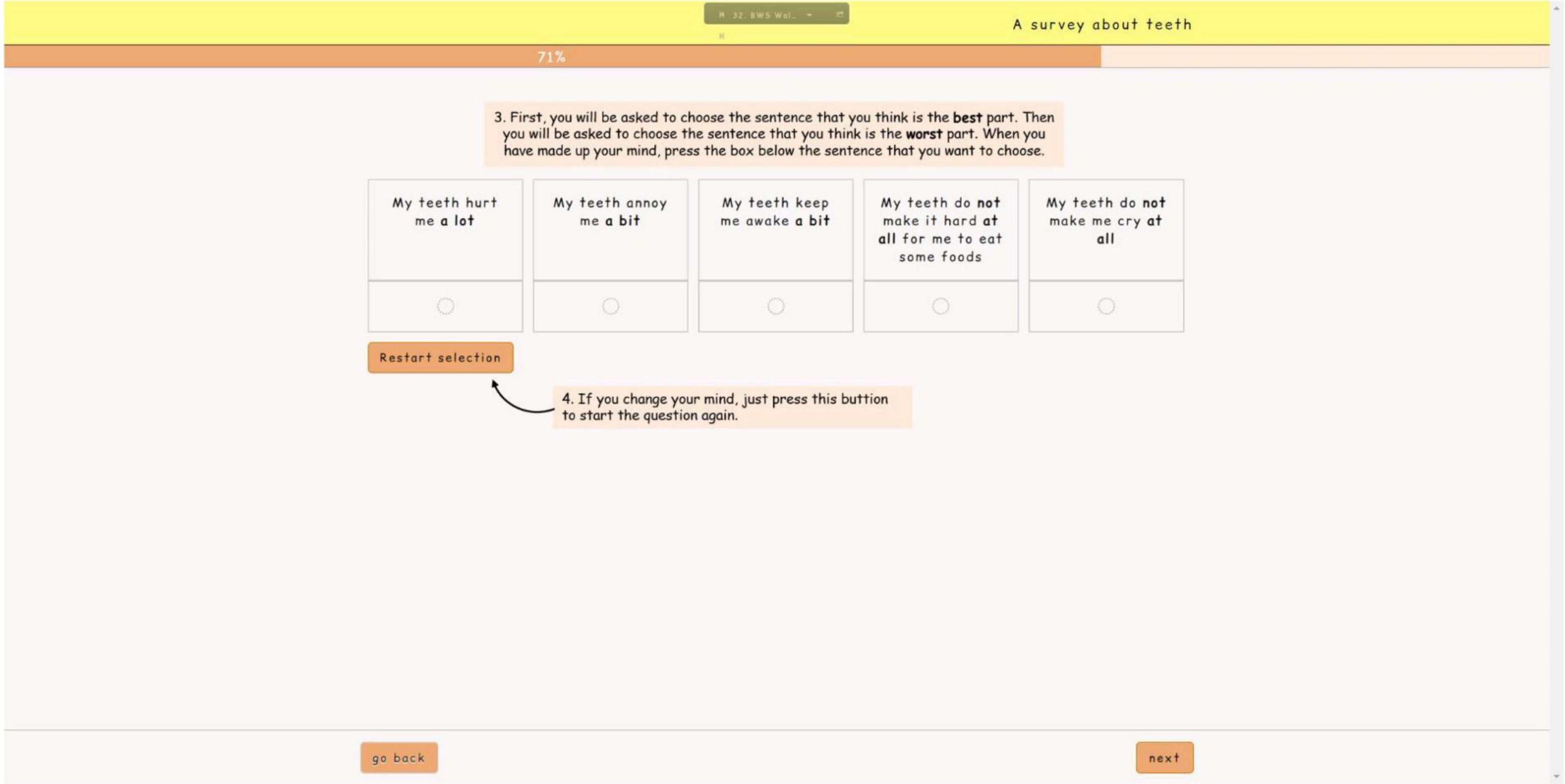
My teeth hurt me **a lot**

My teeth annoy me **a bit**

My teeth keep me awake **a bit**

My teeth do **not** make it hard **at all** for me to eat some foods

My teeth do **not** make me cry **at all**



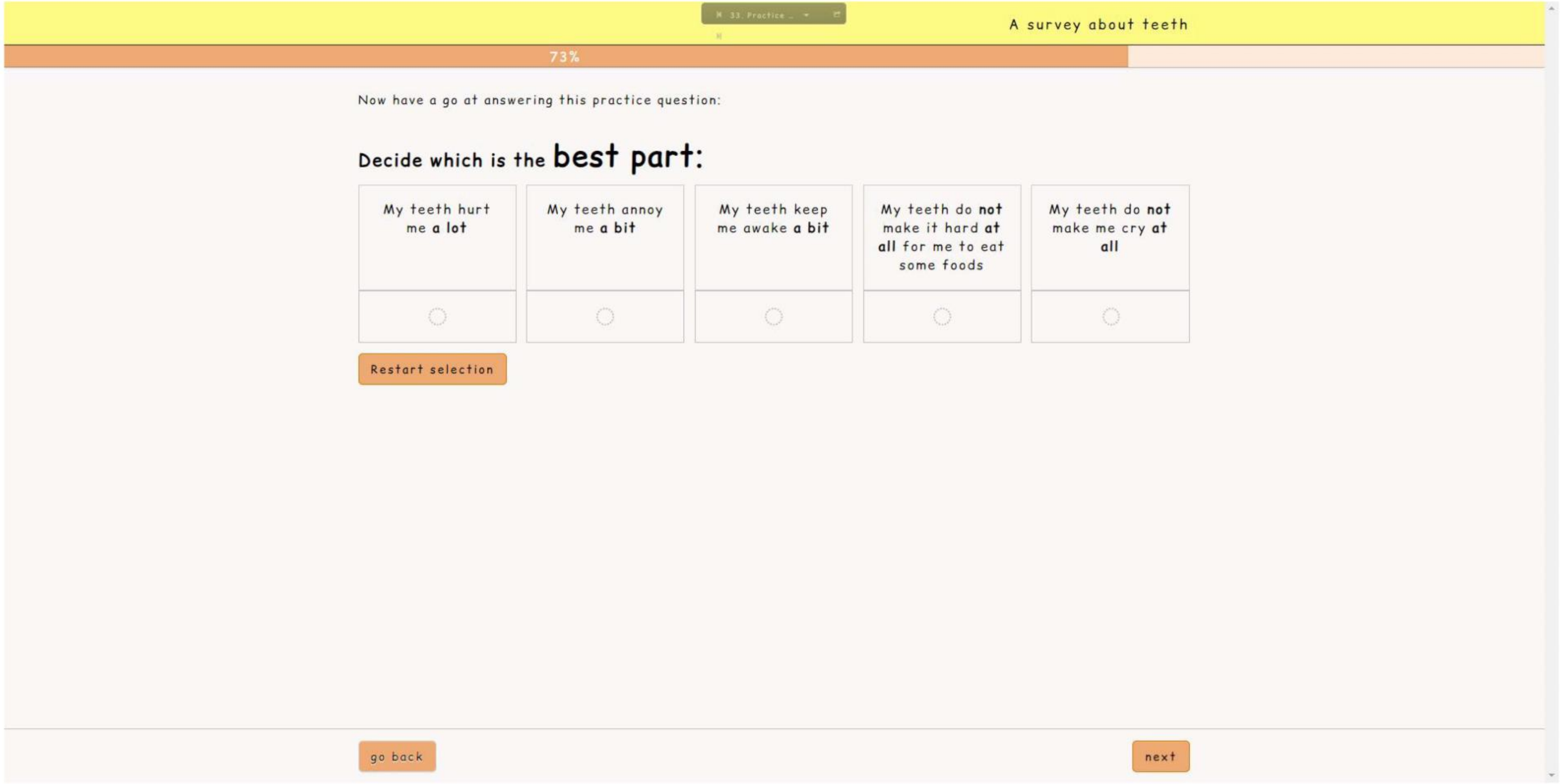
A survey about teeth

71%

3. First, you will be asked to choose the sentence that you think is the **best** part. Then you will be asked to choose the sentence that you think is the **worst** part. When you have made up your mind, press the box below the sentence that you want to choose.

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○

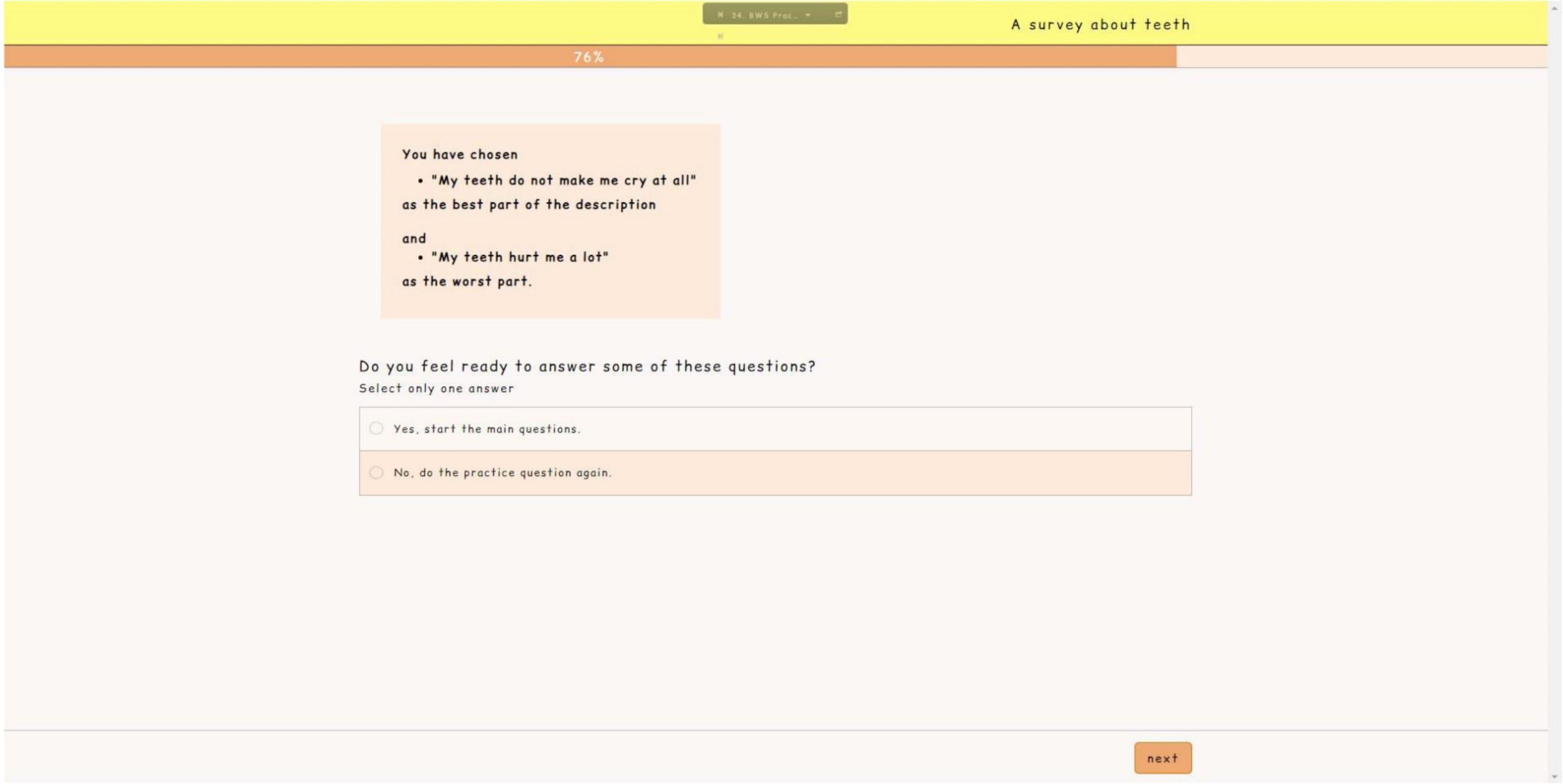
4. If you change your mind, just press this buttion to start the question again.



Now have a go at answering this practice question:

Decide which is the best part:

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○



You have chosen

- "My teeth do not make me cry at all"

as the best part of the description

and

- "My teeth hurt me a lot"

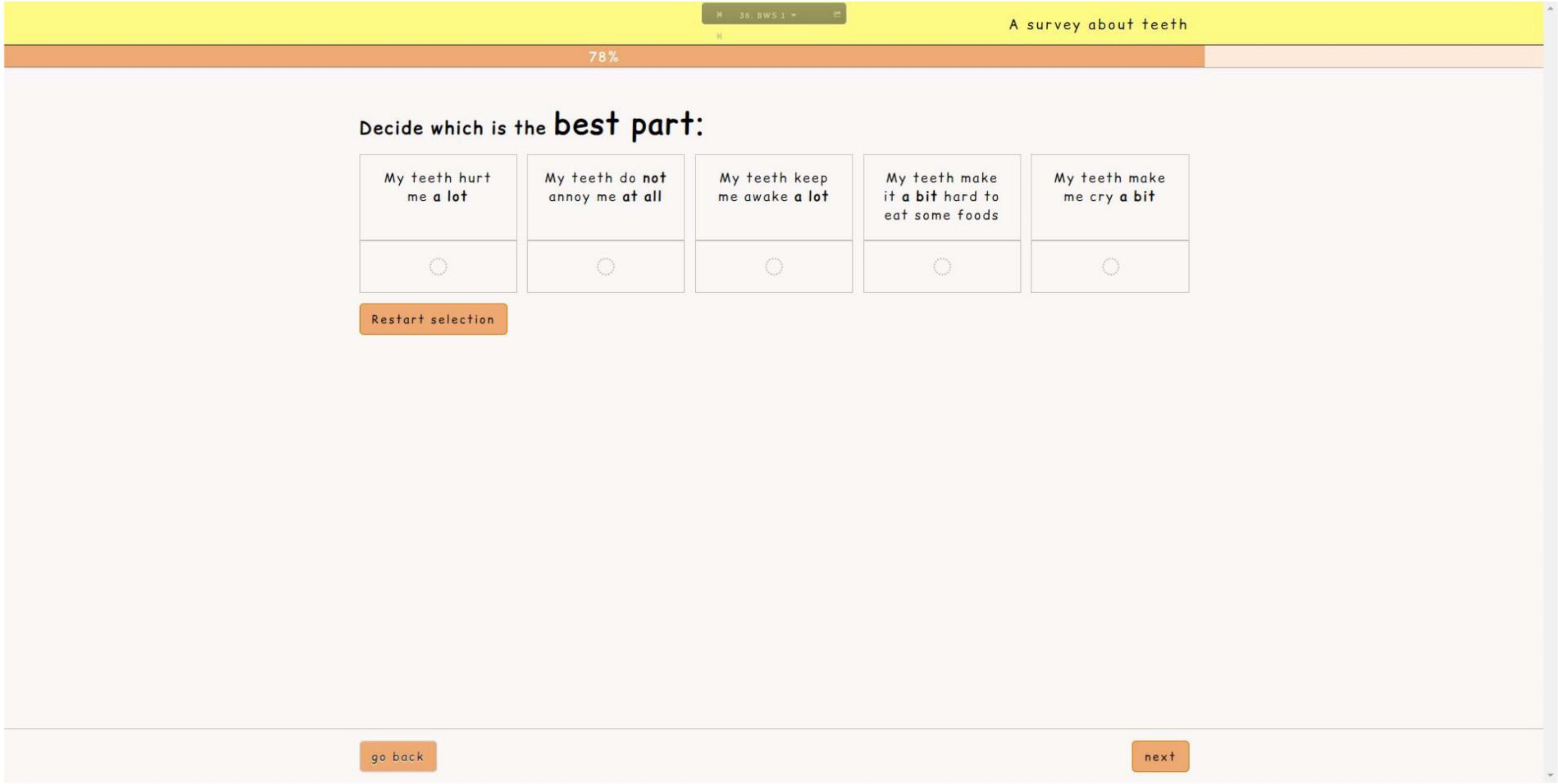
as the worst part.

Do you feel ready to answer some of these questions?

Select only one answer

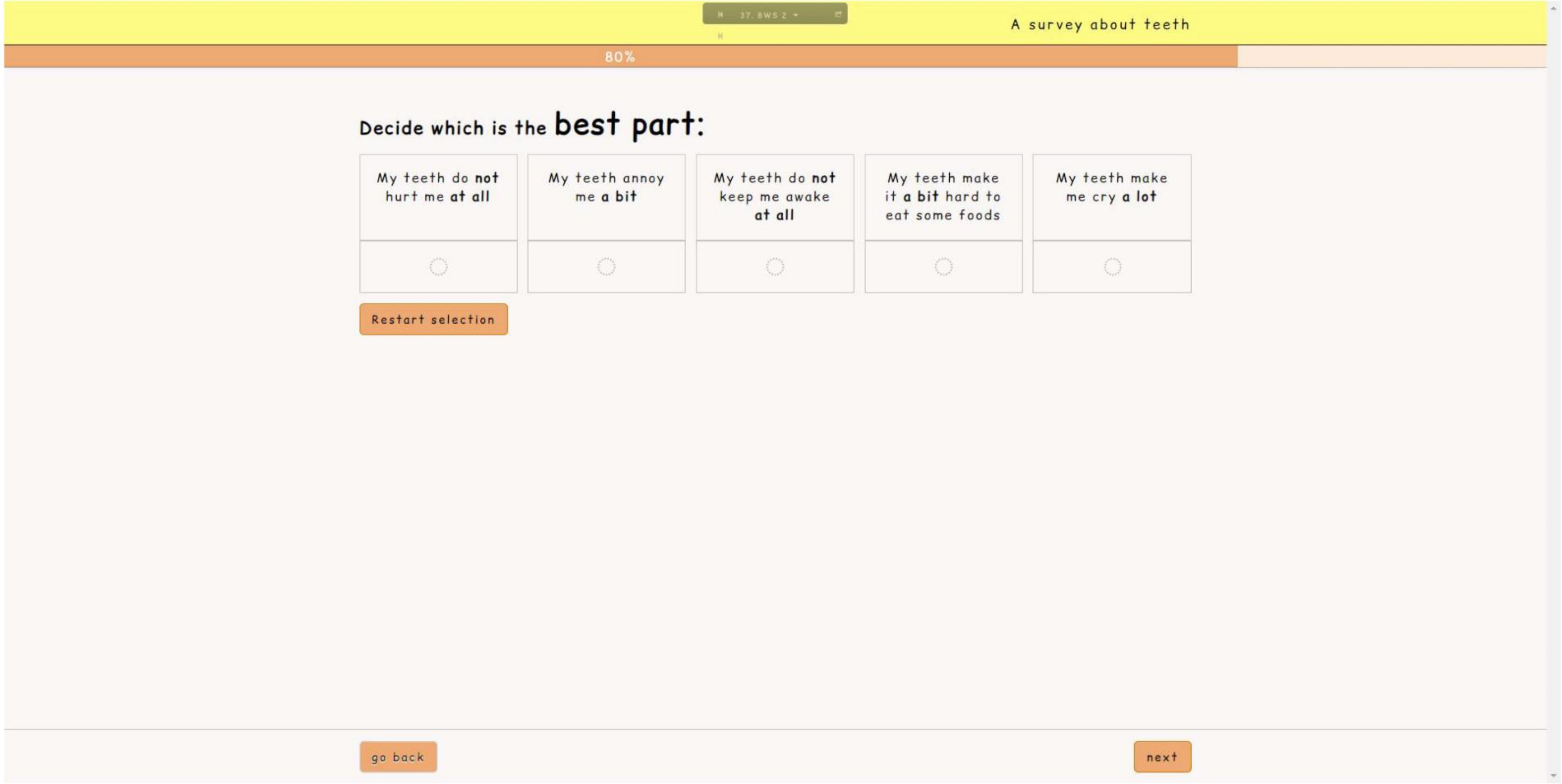
○ Yes, start the main questions.

○ No, do the practice question again.



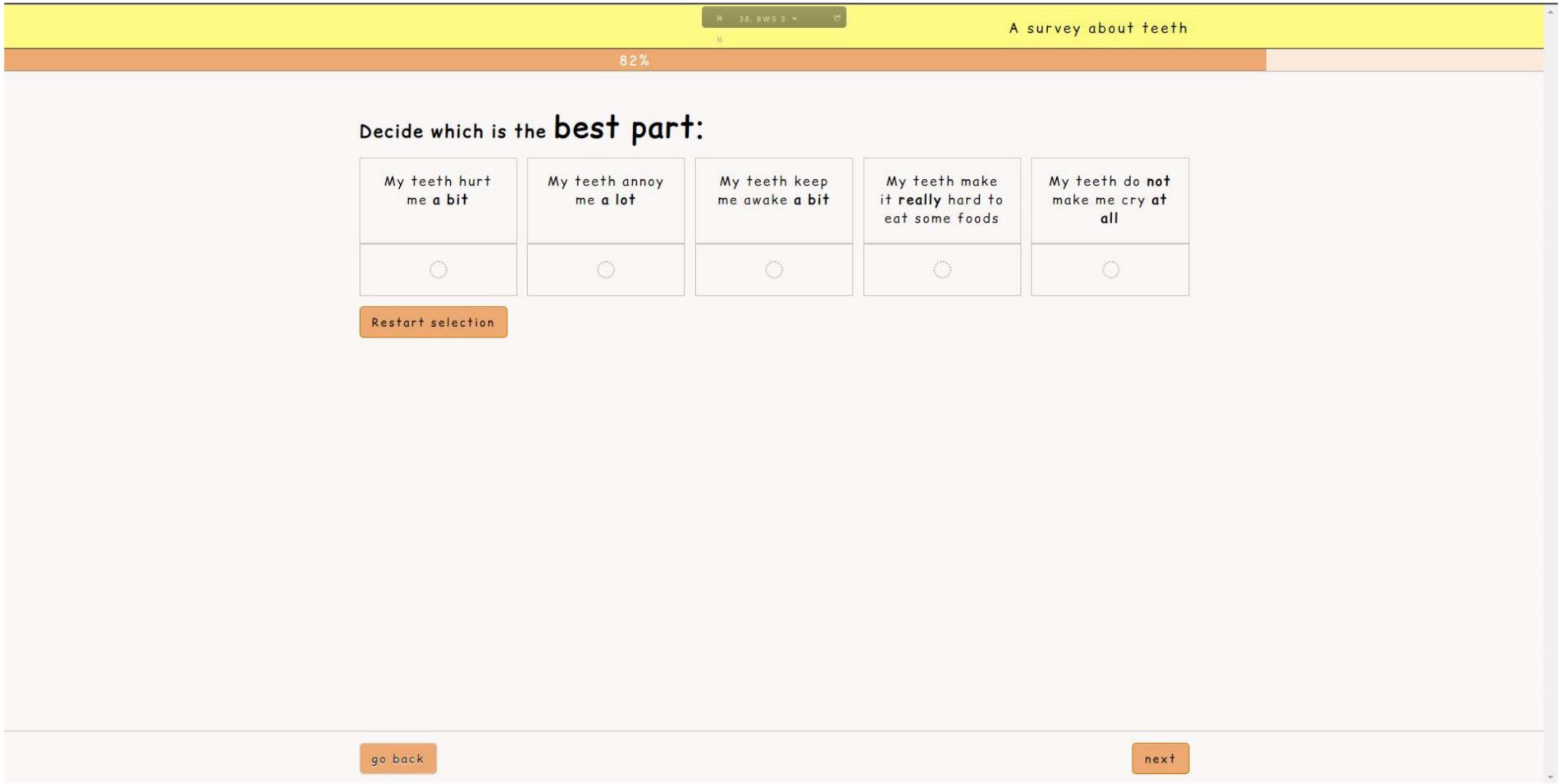
Decide which is the **best part:**

My teeth hurt me **a lot**	My teeth do **not** annoy me **at all**	My teeth keep me awake **a lot**	My teeth make it **a bit** hard to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○



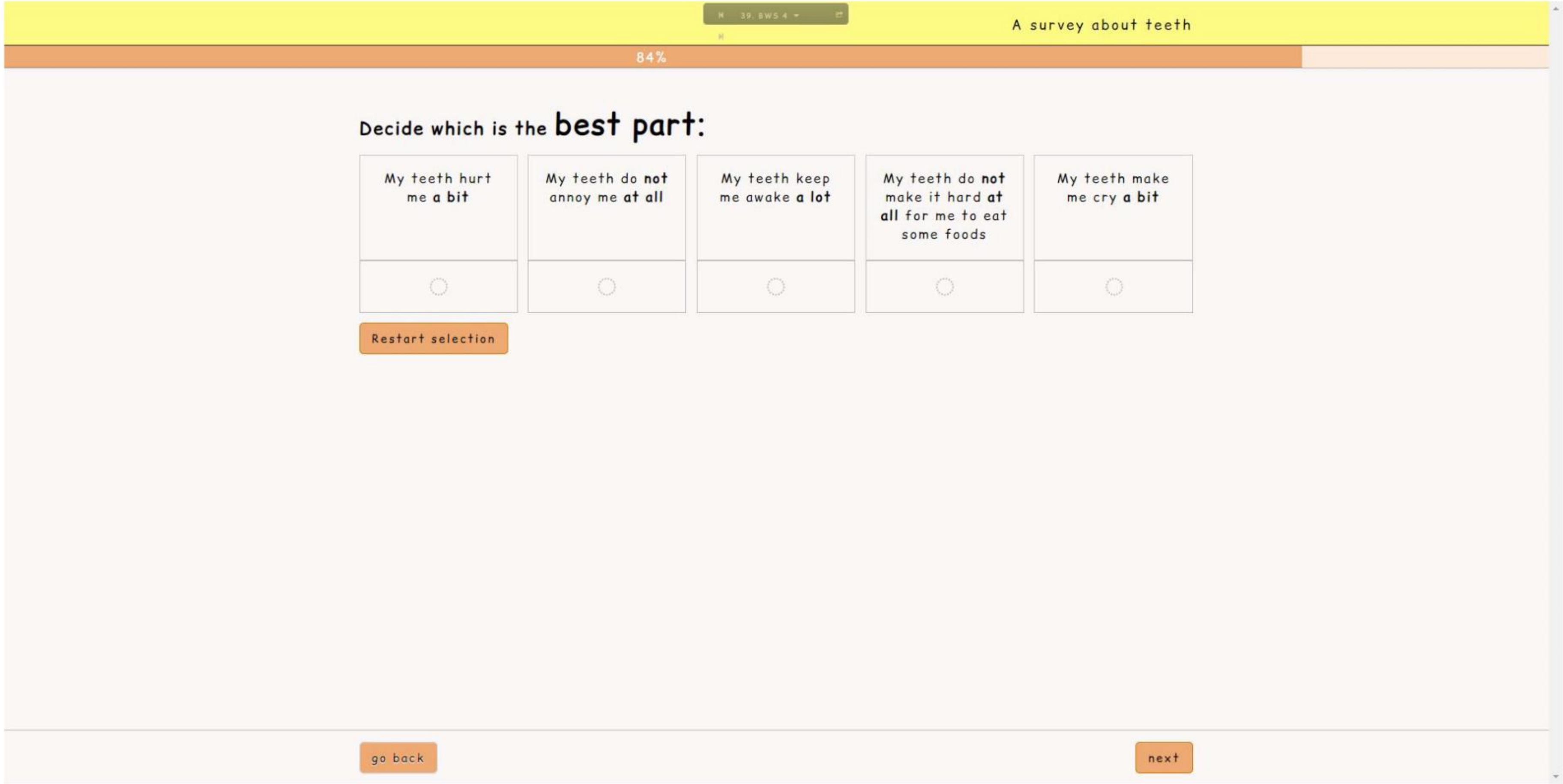
Decide which is the **best part:**

My teeth do **not** hurt me **at all**	My teeth annoy me **a bit**	My teeth do **not** keep me awake **at all**	My teeth make it **a bit** hard to eat some foods	My teeth make me cry **a lot**
○	○	○	○	○



Decide which is the **best part:**

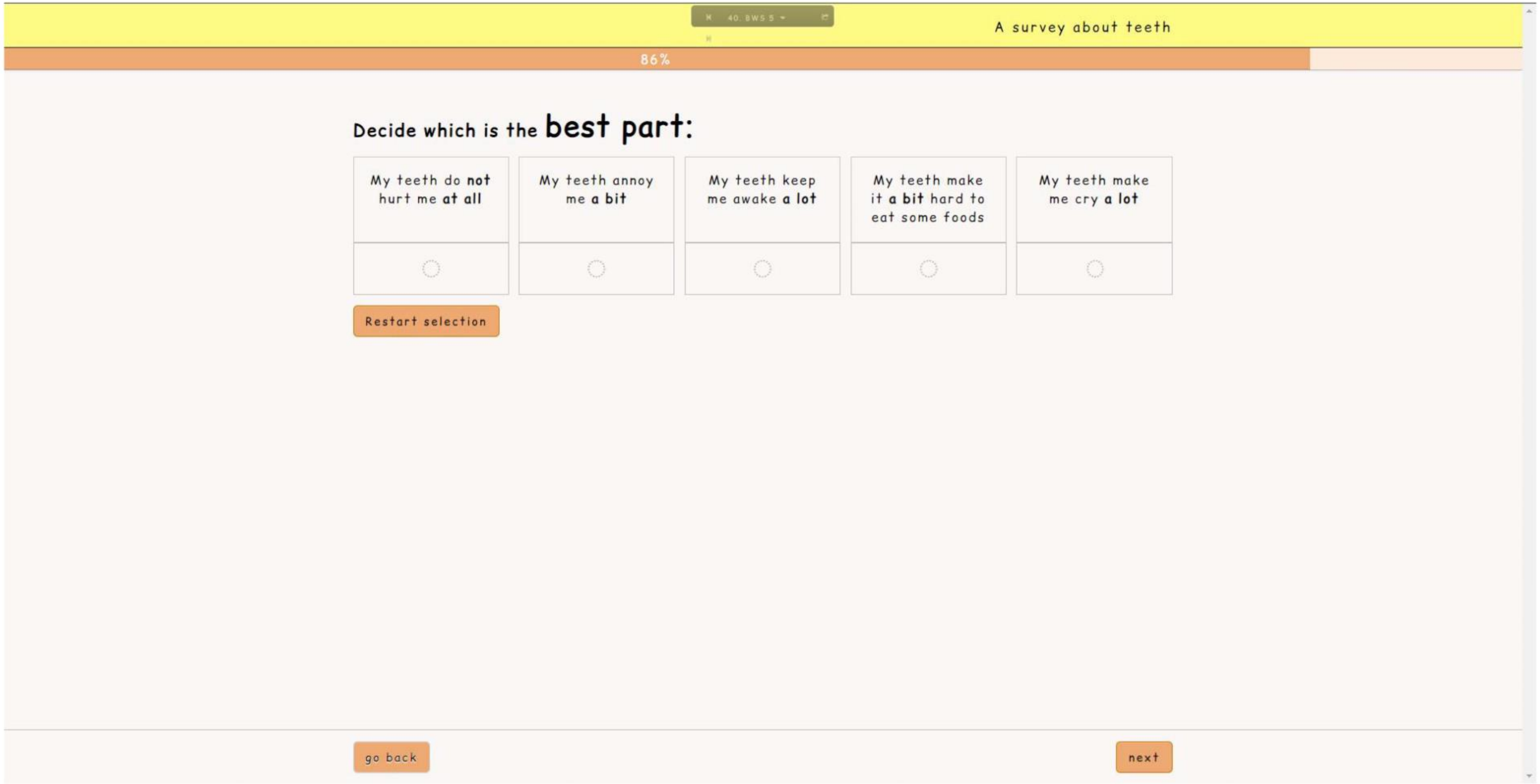
My teeth hurt me **a bit**	My teeth annoy me **a lot**	My teeth keep me awake **a bit**	My teeth make it **really** hard to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○



A survey about teeth

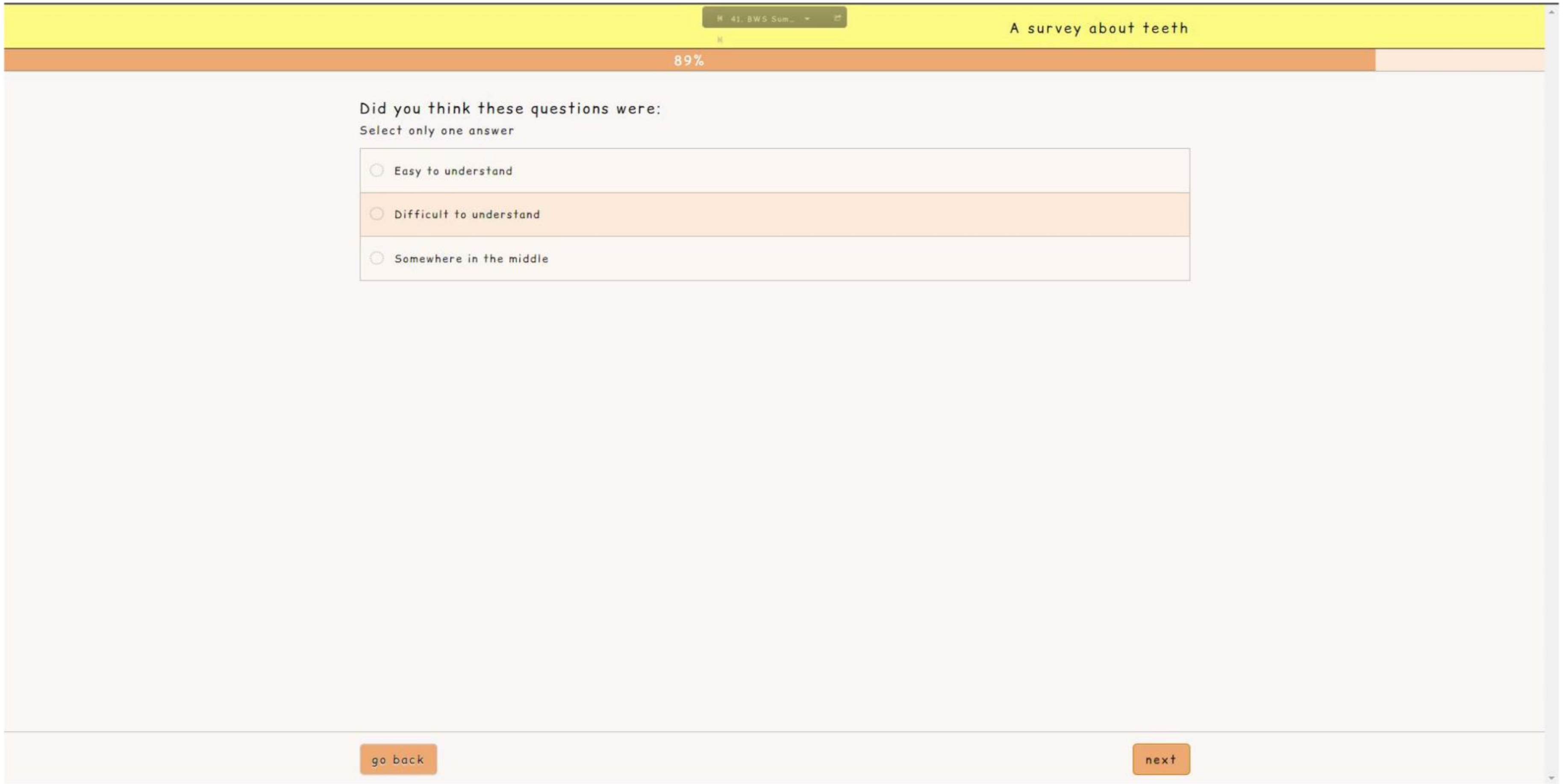
Decide which is the **best part:**

My teeth hurt me **a bit**	My teeth do **not** annoy me **at all**	My teeth keep me awake **a lot**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○



Decide which is the **best part:**

My teeth do **not** hurt me **at all**	My teeth annoy me **a bit**	My teeth keep me awake **a lot**	My teeth make it **a bit** hard to eat some foods	My teeth make me cry **a lot**
○	○	○	○	○



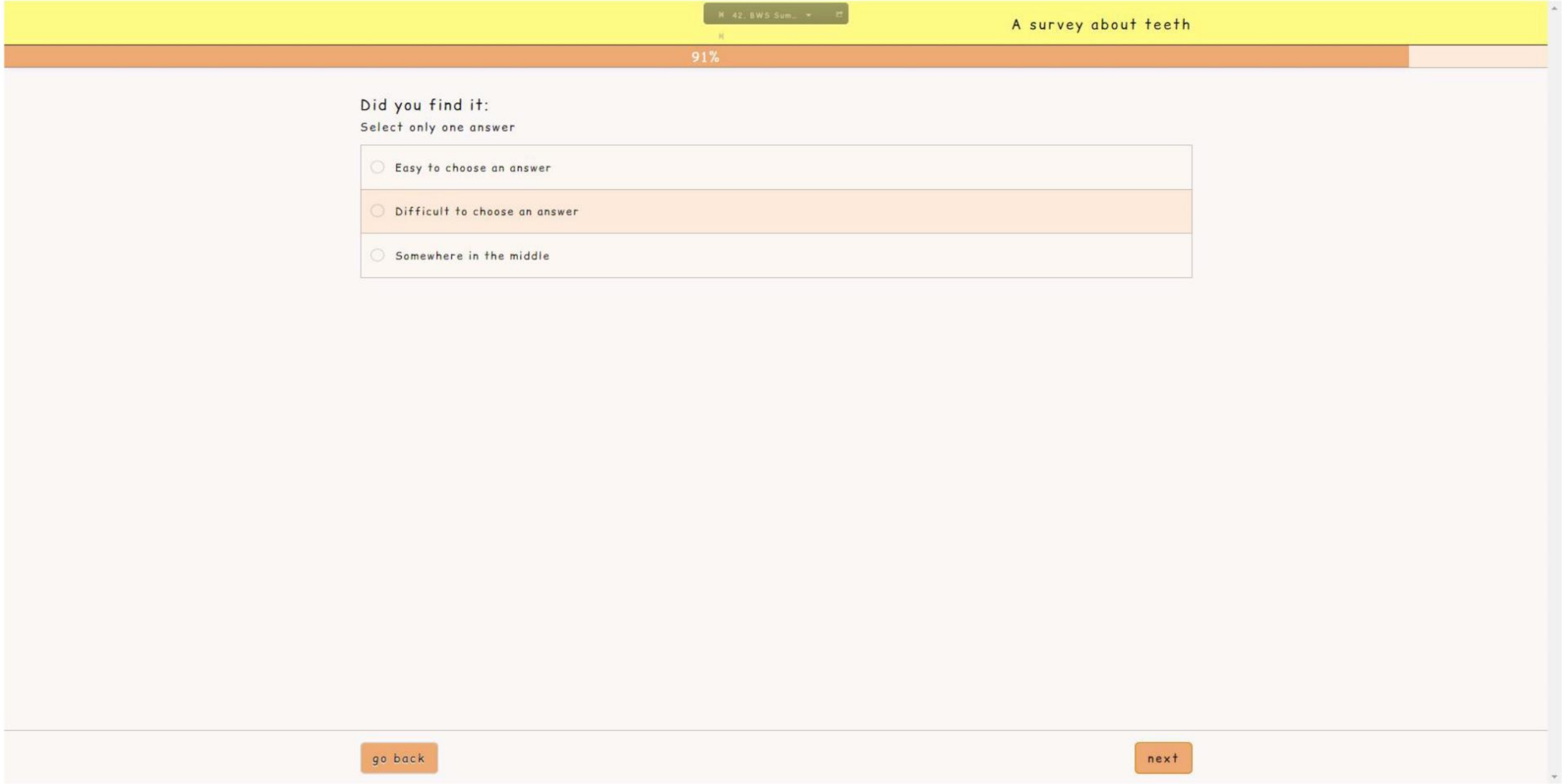
Did you think these questions were:

Select only one answer

○ Easy to understand

○ Difficult to understand

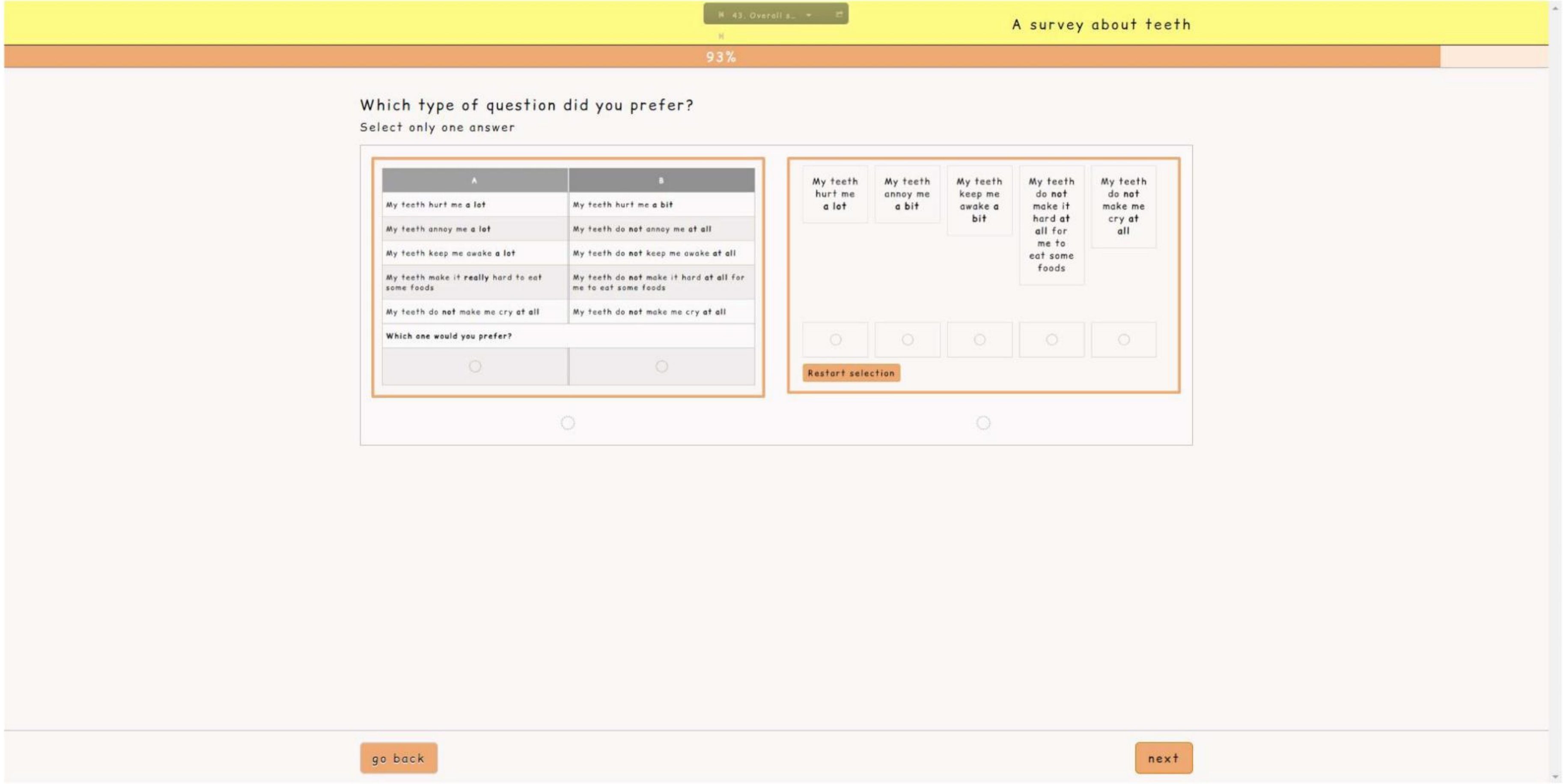
○ Somewhere in the middle



Did you find it:

Select only one answer

- ○ Easy to choose an answer
- ○ Difficult to choose an answer
- ○ Somewhere in the middle



Which type of question did you prefer?

Select only one answer

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth do **not** make me cry **at all**	My teeth do **not** make me cry **at all**
Which one would you prefer?	
○	○

○

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at** **all** for me to eat some foods	My teeth do **not** make me cry **at** **all**
○	○	○	○	○

○

Thank You for taking part in this survey

P. Parental information sheet and consent form for survey and interview

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A survey about teeth: what do you think?

Information sheet for parents/carers

Hello, and thank you for reading this information sheet. My name is Helen Rogers and I am a research student in the University of Sheffield, School of Clinical Dentistry. Before you decide whether you are happy for your child to take part in this project, it is important that you know why the research is being done and what it will involve. Please take time to read this leaflet and talk with others if you wish.

What are you researching?

I am carrying out a research survey to find out what children think about their teeth. The survey was adapted from a questionnaire that was made in the University of Sheffield to find out how tooth decay affects children and young people. I have made the questionnaire shorter, and have turned the questions into two different types of simple tasks which can be completed on a tablet (e.g. iPad) or computer.

This part of my project is trying to find out what children think of the survey and which type of task they prefer. I also hope to find out which age group of children can complete the tasks without any help. The information I find out will be used to help me to improve my survey. The final survey will be used to help us find out which treatments for children with tooth decay are the best value-for-money.

Why are you doing this research?

We have lots of different ways to prevent and treat tooth decay in children, but some treatments need more appointments than others. Even though the NHS provides treatment for children under 16-years free-of-charge, there can still be costs to families, such as:

- Children missing school to go to dental appointments
- Parents taking time off work to take their child to appointments
- Bus/train fares and parking fees
- Arranging childcare for other children

Some treatments also cost the NHS more than others, which means that less money is free to spend on other types of healthcare.

We still do not know which dental treatments are the best value-for-money, and which are not. This research will help us to find this out.

Why do you want to talk to my child?

I am hoping to get about 30 to 35 children of different ages to take part in this stage of my project. Your child's class has been selected at random and these letters have been sent to the parents/carers of all the children in this class.

Does my child have to take part?

No! It is up to you and your child.

If you are happy for your child to take part, please complete the attached consent form and return to your child's school. Your child will then be given all the information they need to decide whether they would like to take part. Your child's lessons will not be disrupted if they take part in this study. Your child will be free to stop taking part at any time, without giving a reason.

If you do not return the consent form by the date given on the attached letter from your child's school, we will assume you do not wish your child to take part.

If you or your child decides not to take part, it will not affect the dental care your child receives in any way.

What will happen if my child agrees to take part?

I will arrange a suitable time for your child to complete the survey in a quiet room within the school. Your child will be asked to complete the survey on the computer and I will encourage them to think aloud when they do each task. Afterwards, I will ask them some questions about the survey and how easy or difficult it was. The survey and interview together should take no more than 15 minutes. I will use a voice recorder during the interview and will write some notes down.

Is there anything to be worried about if my child takes part?

There are no known risks to you or your child from taking part in the study. Your names will not appear in any report written about the study, so you do not need to worry that other people will know what you or your child have said. Your child's name will be removed from the interview recording, which will be kept securely on a password-protected computer in the university. All the information, including your child's age, ethnicity, gender, your contact details and postcode, will be safely stored at the university until the last report on this study has been written.

What are the possible benefits of taking part?

The study will not change the care or treatment your child receives at the dentist. The study will not benefit your child directly, but we hope that the study will help children with dental decay in the future. Your child will be given a £5 gift voucher to say thank you for taking part and for the time given.

What do I do next?

If you are happy for your child to take part, please complete and return the consent form before the date given on the letter from your child's school. If you do not wish your child to take part, you do not need to do anything.

If you would like to speak to me about your child's participation in the research, or about any other aspect of the research, please contact me by telephone: 0114 271 7877 or email hrogers1@sheffield.ac.uk.

What happens when the research stops?

When the study is finished I will look at all the information that I have gained from your child and all the other children who have taken part. I will then write a report on my findings and send you a copy. Your child will continue their regular dental care as normal. The data collected will be stored securely until the last report on the project has been written, before being destroyed. Information is stored in this way in case more research can be done with the data collected.

What if there is a problem or something goes wrong?

If you or child are unhappy about anything, I will be happy to talk to you at any time. You or ~~and~~ your child can also stop taking part at any time without having to give a reason.

What if I am not happy about the way the study has been conducted?

If you or your child are harmed by taking part, there are no special compensation arrangements. If you or your child are harmed due to someone's negligence, then you may have grounds for legal action. You may have to pay for this. Regardless of this, if you wish to complain, or have any concerns about any aspects of the way you have been approached or treated during the course of this study, please contact Kathryn Hurrell-Gillingham, Research Manager for the School of Clinical Dentistry, University of Sheffield, on 0114 2717998 or email k.hurrell-gillingham@sheffield.ac.uk.

Will anyone else know we are taking part?

All personal information recorded on paper will be kept under conditions of strict confidentiality. We will remove all personal details such as names and age from any electronic records, such as the voice recordings during the interviews so that your child cannot be recognised by the researchers. All the information from the study will be kept securely at the University of Sheffield. The reports will not name any participants.

Who is organising and funding the research?

The study is being organised by Helen Rogers, PhD Student, Unit of Oral Health, Dentistry and Society, School of Clinical Dentistry, University of Sheffield. Helen is funded by the National Institute for Health Research, and is supported by an experienced supervisory team.

Who has reviewed the study?

Before any research goes ahead it is checked by an Ethics Committee. They make sure that the research is safe to do. This project has been checked and approved by South Yorkshire Research Ethics Committee.

Thank you for taking the time to read this. Please feel free to ask any questions if you need to.

Q. Adolescent information sheet and assent form for survey and interview

The University Of Sheffield.

A SURVEY ABOUT TEETH: WHAT DO YOU THINK?

Sheffield Teaching Hospitals
NHS Foundation Trust

Who am I?
My name is Helen and I study at the University.
This is me:

What am I doing?
I have made a survey about teeth for young people to complete on a tablet or computer. I am asking young people at two schools in this area to help me find out how to make the survey better. To do this, I am asking children aged 11- to 16-years to fill in my survey, and asking them questions to find out what they think about it.

Can you help me?
Firstly, I would like you to fill in the survey for me.
The first part of the survey involves some questions about you. The second part of the survey contains ten questions about your teeth. The questions are asked in two different ways. There are some practice questions to complete first so that you understand how to answer them.

I will ask you some questions as you complete the survey, and some more at the end to find out what you think about the survey. There are no right or wrong answers.

The survey and interview will take about 20 minutes to complete.

What happens after?
After you have done the survey and interview, I will give you a £5 gift voucher to say thank you for your help.

I will use the answers that you give me to change the survey to make it better. The improved survey will then be given to lots of young people around the country to complete. This will help us to understand more about how young people feel about their teeth.

I will then write about what I have found out, so that other young people, parents and dentists can read about it.

What if you don't want to join in anymore?
It is fine for you to stop filling in the survey or stop the interview at any time. No one will be unhappy with you.

What do you do now?
There is a form for you to fill in if you would like to take part. Please put your initials in the boxes, and write your name, signature and date in the spaces at the bottom.

If you have any questions or are unhappy, you can phone me on 0114 271 7877 or email hrogers1@sheffield.ac.uk.

Thank you for reading this.

Helen

The University Of Sheffield.

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

A survey about teeth: what do you think?

Young Person Assent Form

For office use only:

School	
Participant Identifier	

Name of Researcher: Helen Rogers

Sponsor: Sheffield Teaching Hospitals NHS Foundation Trust

If you agree to the following statements, please initial the boxes.

1. I have read the information provided for young people about this project. I have had the chance to think about the information, ask questions and my questions have been answered. ☐
2. I understand that it is my choice to take part and that I can change my mind at any time without giving a reason. I understand it will not change anything at school, my usual dental care or my legal rights. ☐
3. I understand that personal information that identifies me will be collected during the study. It will be kept in the School of Clinical Dentistry, University of Sheffield and may be looked at by other members of the research team, or other responsible people to check the project is being done well. I agree for them to see my information. ☐
4. I understand that my school will share information about me and that my parents or carers may share information about me with the research team. ☐
5. I understand that the information collected about me by the research team will be used to support other projects in the future and may be shared with other researchers but they will not be given my name. ☐
6. OPTIONAL: I agree to be contacted by the research team in the future if there are any other projects I may be able to take part in. ☐
7. **I would like to take part in this project.** ☐

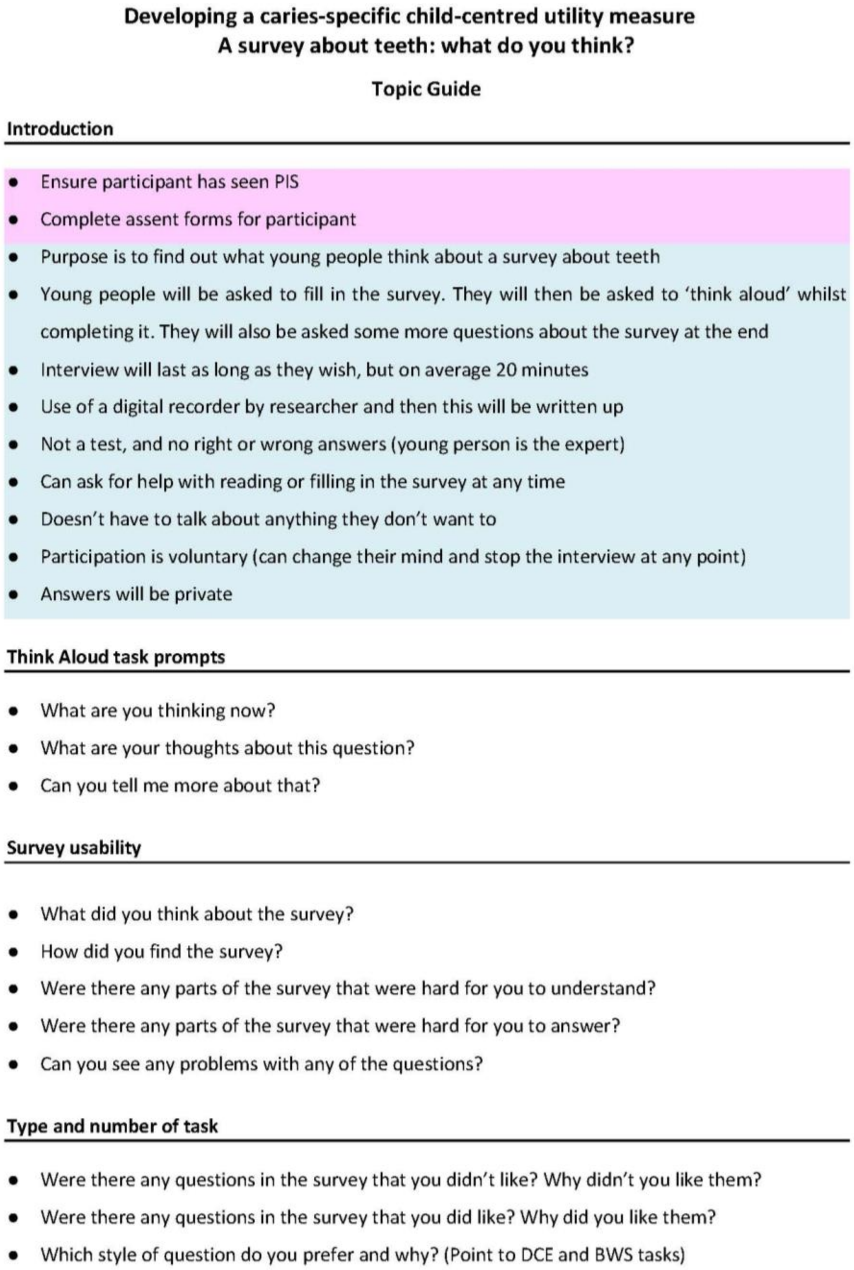
Full name:	
Signature:	
Date:	

Name of researcher:		**Date:**	
Researcher signature:			

IRAS: 241330 STH:19616 Stage 3 Young Person Assent Form: V1 (20th June 2018)
Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation

Funded by NHS National Institute for Health Research

R. Topic guide for qualitative DCE vs BWS interview



Developing a caries-specific child-centred utility measure
A survey about teeth: what do you think?

Topic Guide

Introduction

- Ensure participant has seen PIS
- Complete assent forms for participant
- Purpose is to find out what young people think about a survey about teeth
- Young people will be asked to fill in the survey. They will then be asked to 'think aloud' whilst completing it. They will also be asked some more questions about the survey at the end
- Interview will last as long as they wish, but on average 20 minutes
- Use of a digital recorder by researcher and then this will be written up
- Not a test, and no right or wrong answers (young person is the expert)
- Can ask for help with reading or filling in the survey at any time
- Doesn't have to talk about anything they don't want to
- Participation is voluntary (can change their mind and stop the interview at any point)
- Answers will be private

Think Aloud task prompts

- What are you thinking now?
- What are your thoughts about this question?
- Can you tell me more about that?

Survey usability

- What did you think about the survey?
- How did you find the survey?
- Were there any parts of the survey that were hard for you to understand?
- Were there any parts of the survey that were hard for you to answer?
- Can you see any problems with any of the questions?

Type and number of task

- Were there any questions in the survey that you didn't like? Why didn't you like them?
- Were there any questions in the survey that you did like? Why did you like them?
- Which style of question do you prefer and why? (Point to DCE and BWS tasks)

- How many of these type of questions do you think you could manage before you lose concentration or get bored?

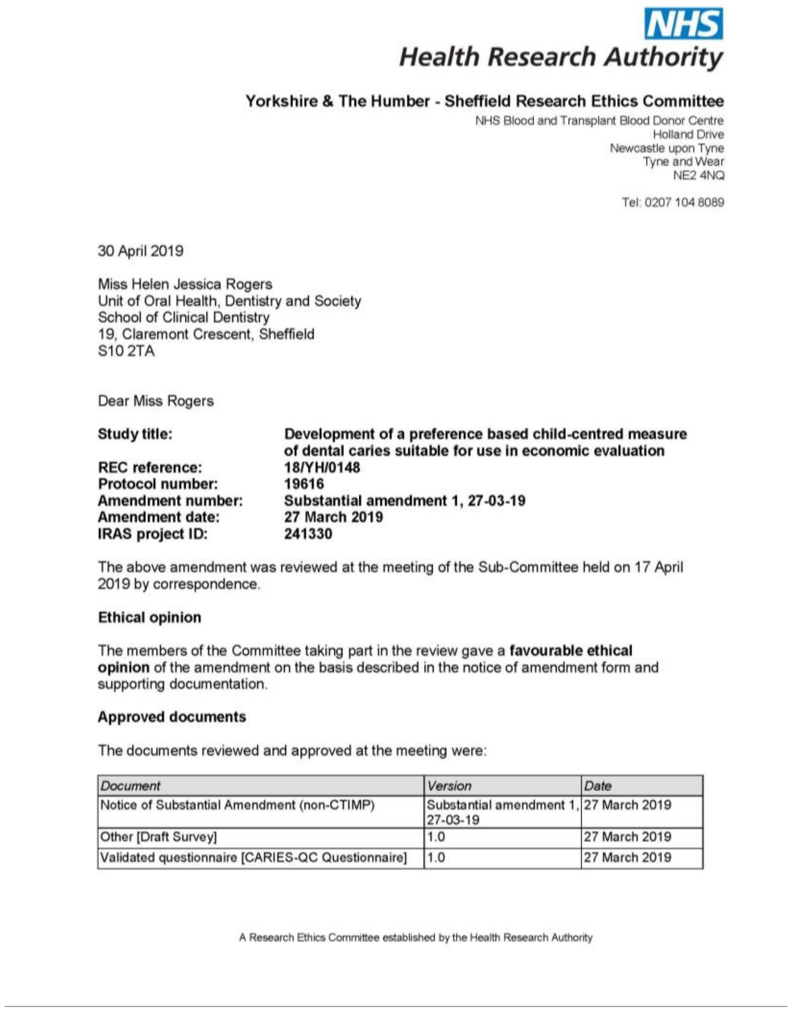
Practice questions and information about caries

- During the survey you were asked to complete two practice questions. What did you think about these questions? Did you think they were helpful or not?
- Do you think more practice questions would help you or not? Do you think a different practice question would be better? Would a practice question that is not about teeth be useful or not?
- In the survey, you were given some information about tooth decay (holes). What did you think about this information? Did you find it useful? Is there anything you would want to change or add to this information?

Next steps

- Thank the participant
- Reassurance again about confidentiality/privacy
- Explain that findings will be used to develop a survey that will help us find out which dental treatments are value-for-money
- Inform the child that they will receive a copy of the final report through their school
- Give the participant a gift voucher and ask them to sign reciept

S. Favourable ethical opinion for first substantial amendment



NHS
Health Research Authority

Yorkshire & The Humber - Sheffield Research Ethics Committee
NHS Blood and Transplant Blood Donor Centre
Holland Drive
Newcastle upon Tyne
Tyne and Wear
NE2 4NQ

Tel: 0207 104 8089

30 April 2019

Miss Helen Jessica Rogers
Unit of Oral Health, Dentistry and Society
School of Clinical Dentistry
19, Claremont Crescent, Sheffield
S10 2TA

Dear Miss Rogers

Study title:	**Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation**
REC reference:	**18/YH/0148**
Protocol number:	**19616**
Amendment number:	**Substantial amendment 1, 27-03-19**
Amendment date:	**27 March 2019**
IRAS project ID:	**241330**

The above amendment was reviewed at the meeting of the Sub-Committee held on 17 April 2019 by correspondence.

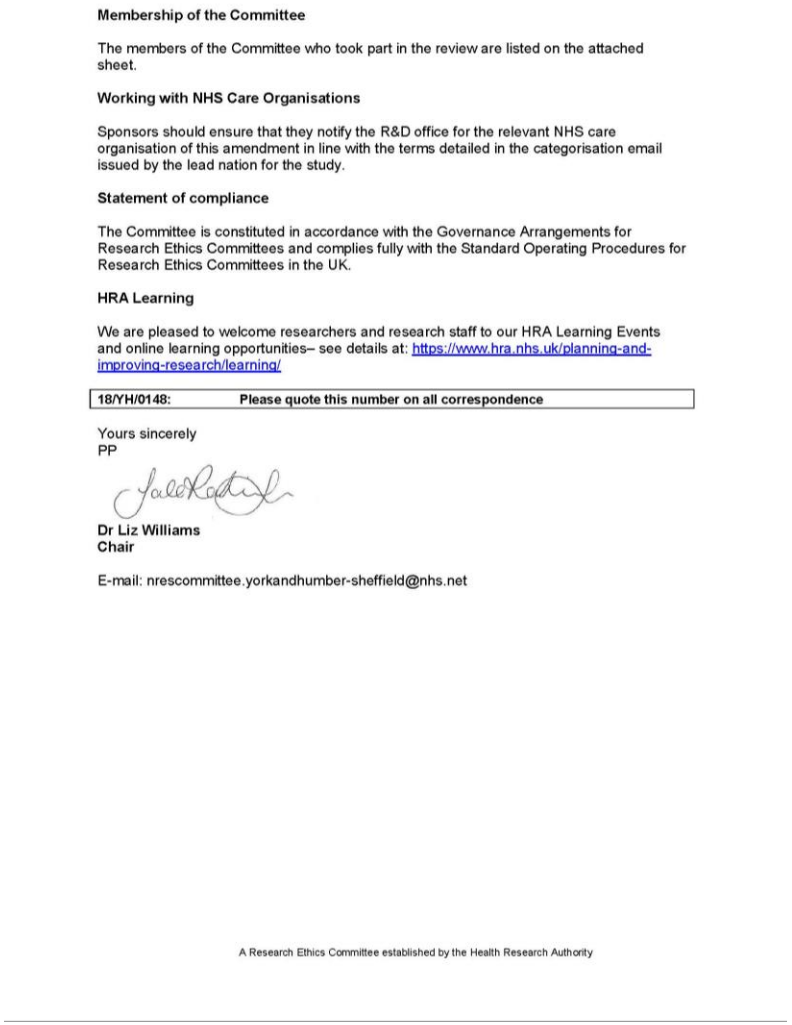
Ethical opinion

The members of the Committee taking part in the review gave a **favourable ethical opinion** of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMP)	Substantial amendment 1, 27-03-19	27 March 2019
Other [Draft Survey]	1.0	27 March 2019
Validated questionnaire [CARIES-QC Questionnaire]	1.0	27 March 2019



Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

Working with NHS Care Organisations

Sponsors should ensure that they notify the R&D office for the relevant NHS care organisation of this amendment in line with the terms detailed in the categorisation email issued by the lead nation for the study.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at: https://www.hra.nhs.uk/planning-and-improving-research/learning/

18/YH/0148:	**Please quote this number on all correspondence**

Yours sincerely
PP

Dr Liz Williams
Chair

E-mail: nrescommittee.yorkandhumber-sheffield@nhs.net

Yorkshire & The Humber - Sheffield Research Ethics Committee

Attendance at Sub-Committee of the REC meeting on 17 April 2019 by correspondence

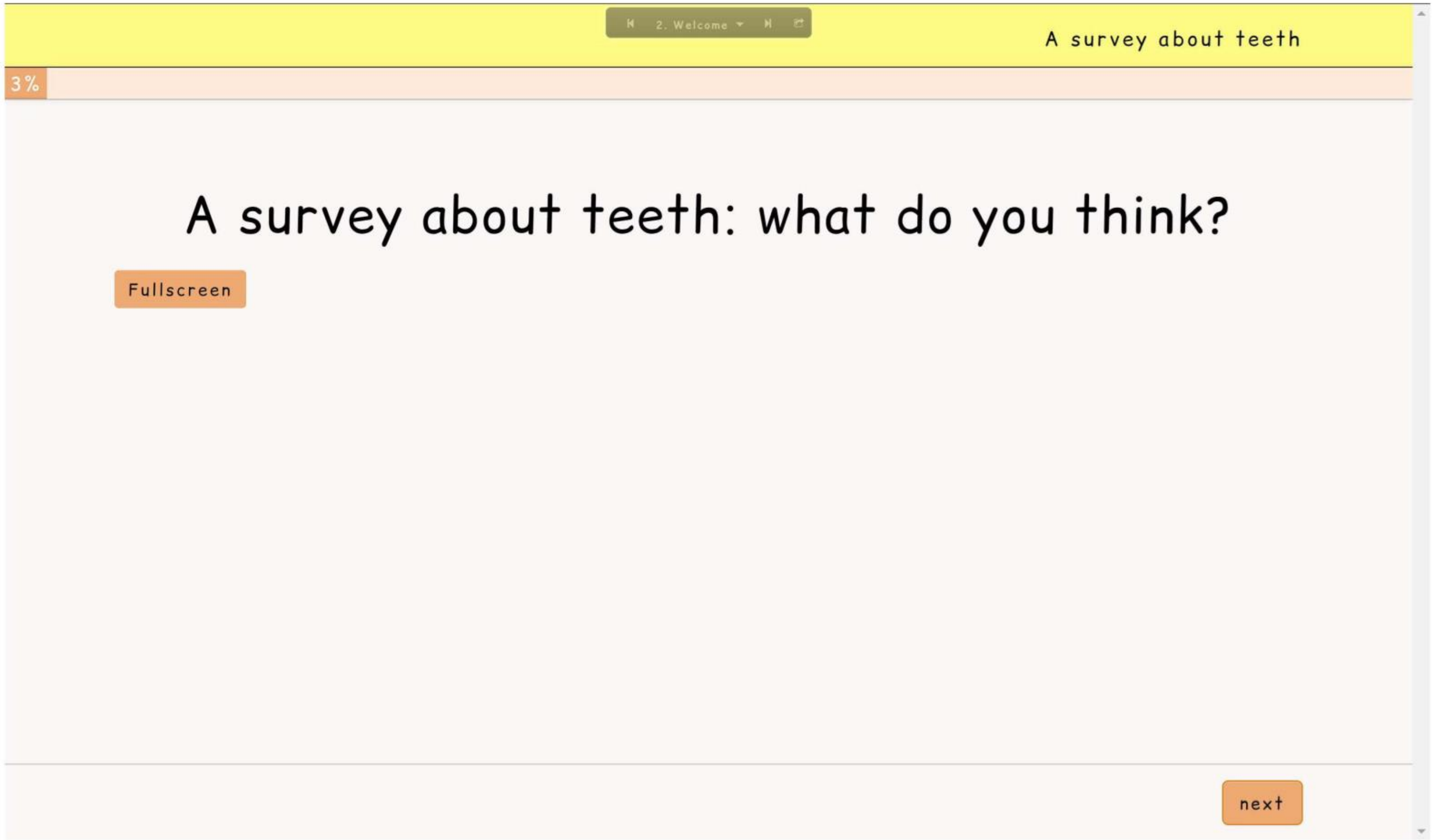
Committee Members:

Name	*Profession*	*Present*	*Notes*
Professor Basil Sharrack	Consultant Neurologist	Yes	
Dr Liz Williams	Senior Lecturer in Human Nutrition (Sheff)	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Miss Jade Robinson	Approvals Administrator

T. BWS valuation survey for adolescents



A survey about teeth

A survey about teeth: what do you think?

A survey about teeth: what do you think?

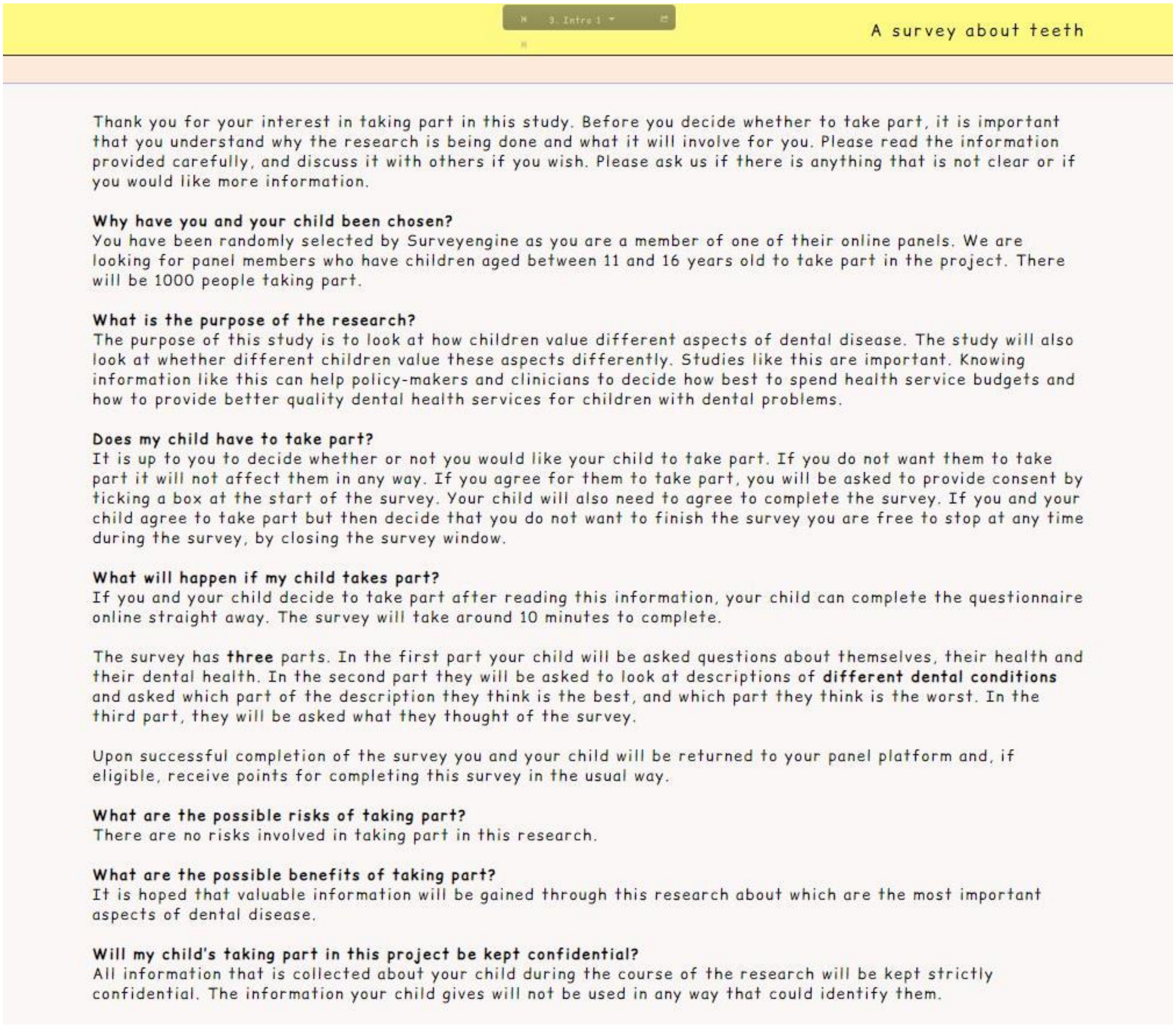
Today, we have a survey about teeth for which we are looking for children aged 11-16 years old to complete. The survey will involve some questions about different aspects of tooth decay. This survey will be approximately 10 minutes long. We would expect children to be able to complete this survey by themselves. Please read through the next pages to find out more about the study.

Are you the legal guardian of any children living in your household aged 11-16 years old?

Select only one answer

○ Yes

○ No



A survey about teeth

Thank you for your interest in taking part in this study. Before you decide whether to take part, it is important that you understand why the research is being done and what it will involve for you. Please read the information provided carefully, and discuss it with others if you wish. Please ask us if there is anything that is not clear or if you would like more information.

Why have you and your child been chosen?
You have been randomly selected by Surveyengine as you are a member of one of their online panels. We are looking for panel members who have children aged between 11 and 16 years old to take part in the project. There will be 1000 people taking part.

What is the purpose of the research?
The purpose of this study is to look at how children value different aspects of dental disease. The study will also look at whether different children value these aspects differently. Studies like this are important. Knowing information like this can help policy-makers and clinicians to decide how best to spend health service budgets and how to provide better quality dental health services for children with dental problems.

Does my child have to take part?
It is up to you to decide whether or not you would like your child to take part. If you do not want them to take part it will not affect them in any way. If you agree for them to take part, you will be asked to provide consent by ticking a box at the start of the survey. Your child will also need to agree to complete the survey. If you and your child agree to take part but then decide that you do not want to finish the survey you are free to stop at any time during the survey, by closing the survey window.

What will happen if my child takes part?
If you and your child decide to take part after reading this information, your child can complete the questionnaire online straight away. The survey will take around 10 minutes to complete.

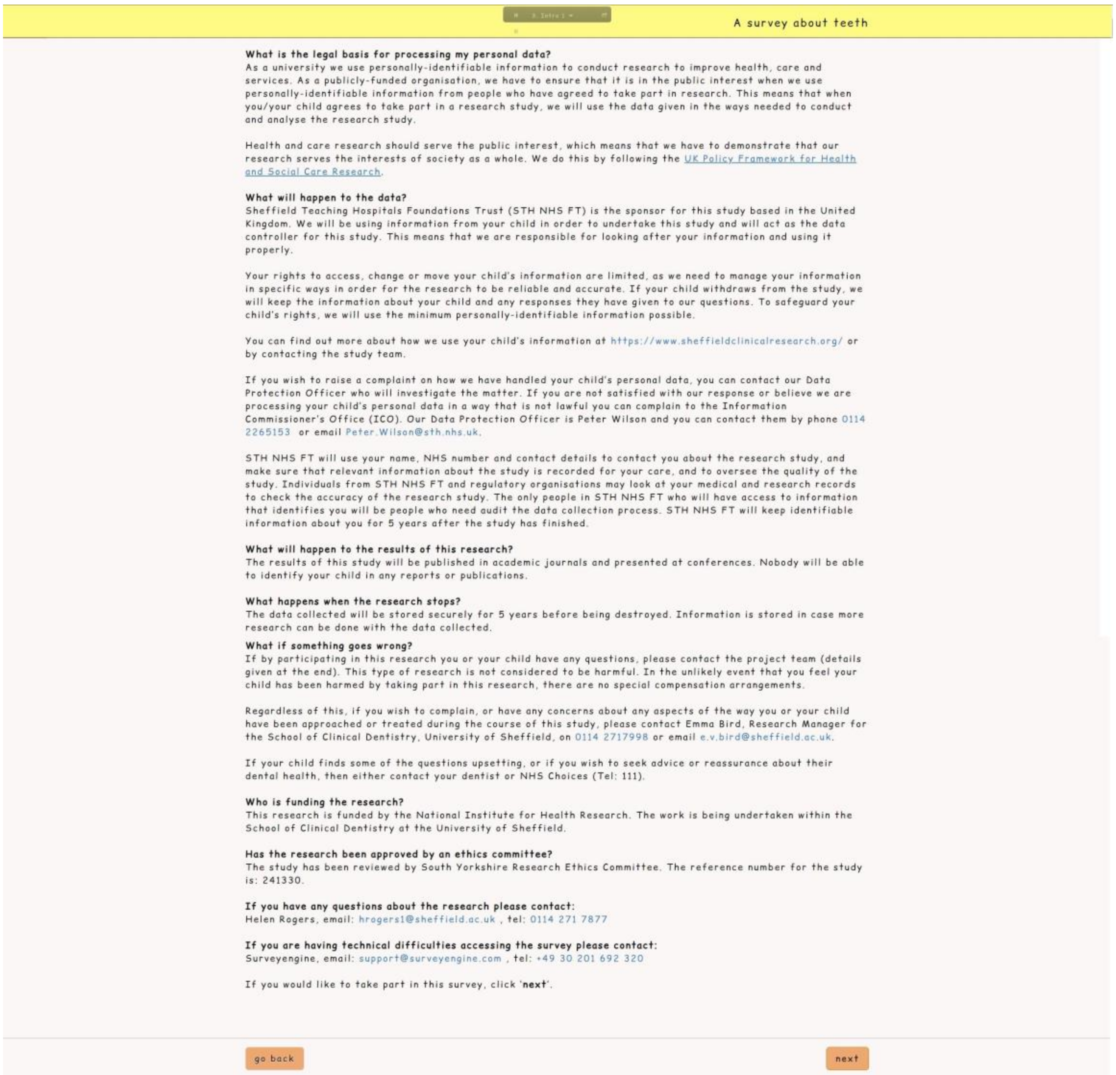
The survey has **three** parts. In the first part your child will be asked questions about themselves, their health and their dental health. In the second part they will be asked to look at descriptions of **different dental conditions** and asked which part of the description they think is the best, and which part they think is the worst. In the third part, they will be asked what they thought of the survey.

Upon successful completion of the survey you and your child will be returned to your panel platform and, if eligible, receive points for completing this survey in the usual way.

What are the possible risks of taking part?
There are no risks involved in taking part in this research.

What are the possible benefits of taking part?
It is hoped that valuable information will be gained through this research about which are the most important aspects of dental disease.

Will my child's taking part in this project be kept confidential?
All information that is collected about your child during the course of the research will be kept strictly confidential. The information your child gives will not be used in any way that could identify them.



What is the legal basis for processing my personal data?
As a university we use personally-identifiable information to conduct research to improve health, care and services. As a publicly-funded organisation, we have to ensure that it is in the public interest when we use personally-identifiable information from people who have agreed to take part in research. This means that when you/your child agrees to take part in a research study, we will use the data given in the ways needed to conduct and analyse the research study.

Health and care research should serve the public interest, which means that we have to demonstrate that our research serves the interests of society as a whole. We do this by following the UK Policy Framework for Health and Social Care Research.

What will happen to the data?
Sheffield Teaching Hospitals Foundations Trust (STH NHS FT) is the sponsor for this study based in the United Kingdom. We will be using information from your child in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly.

Your rights to access, change or move your child's information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If your child withdraws from the study, we will keep the information about your child and any responses they have given to our questions. To safeguard your child's rights, we will use the minimum personally-identifiable information possible.

You can find out more about how we use your child's information at https://www.sheffieldclinicalresearch.org/ or by contacting the study team.

If you wish to raise a complaint on how we have handled your child's personal data, you can contact our Data Protection Officer who will investigate the matter. If you are not satisfied with our response or believe we are processing your child's personal data in a way that is not lawful you can complain to the Information Commissioner's Office (ICO). Our Data Protection Officer is Peter Wilson and you can contact them by phone 0114 2265153 or email Peter.Wilson@sth.nhs.uk.

STH NHS FT will use your name, NHS number and contact details to contact you about the research study, and make sure that relevant information about the study is recorded for your care, and to oversee the quality of the study. Individuals from STH NHS FT and regulatory organisations may look at your medical and research records to check the accuracy of the research study. The only people in STH NHS FT who will have access to information that identifies you will be people who need audit the data collection process. STH NHS FT will keep identifiable information about you for 5 years after the study has finished.

What will happen to the results of this research?
The results of this study will be published in academic journals and presented at conferences. Nobody will be able to identify your child in any reports or publications.

What happens when the research stops?
The data collected will be stored securely for 5 years before being destroyed. Information is stored in case more research can be done with the data collected.

What if something goes wrong?
If by participating in this research you or your child have any questions, please contact the project team (details given at the end). This type of research is not considered to be harmful. In the unlikely event that you feel your child has been harmed by taking part in this research, there are no special compensation arrangements.

Regardless of this, if you wish to complain, or have any concerns about any aspects of the way you or your child have been approached or treated during the course of this study, please contact Emma Bird, Research Manager for the School of Clinical Dentistry, University of Sheffield, on 0114 2717998 or email e.v.bird@sheffield.ac.uk.

If your child finds some of the questions upsetting, or if you wish to seek advice or reassurance about their dental health, then either contact your dentist or NHS Choices (Tel: 111).

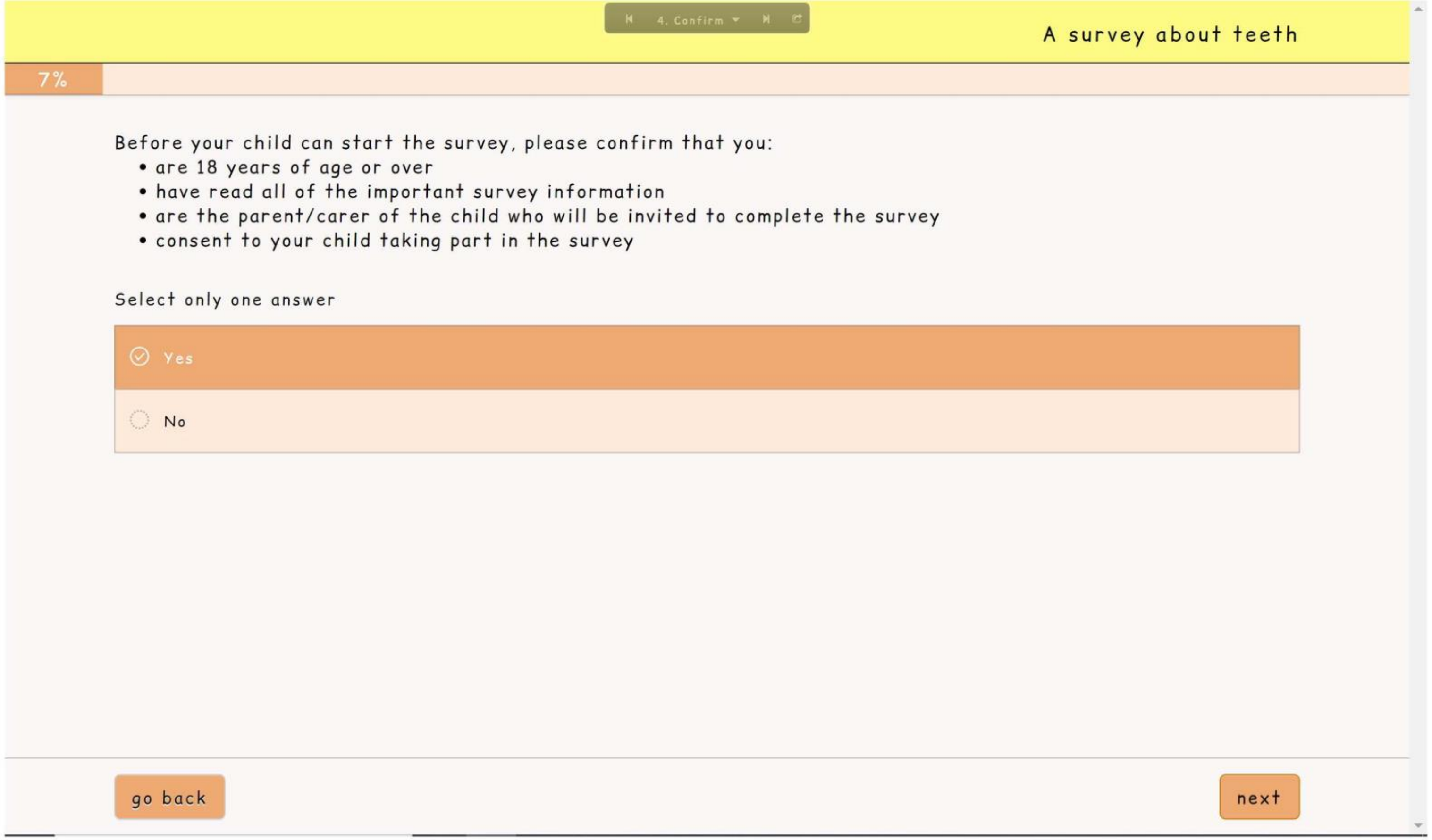
Who is funding the research?
This research is funded by the National Institute for Health Research. The work is being undertaken within the School of Clinical Dentistry at the University of Sheffield.

Has the research been approved by an ethics committee?
The study has been reviewed by South Yorkshire Research Ethics Committee. The reference number for the study is: 241330.

If you have any questions about the research please contact:
Helen Rogers, email: hrogers1@sheffield.ac.uk , tel: 0114 271 7877

If you are having technical difficulties accessing the survey please contact:
Surveyengine, email: support@surveyengine.com , tel: +49 30 201 692 320

If you would like to take part in this survey, click '**next**'.



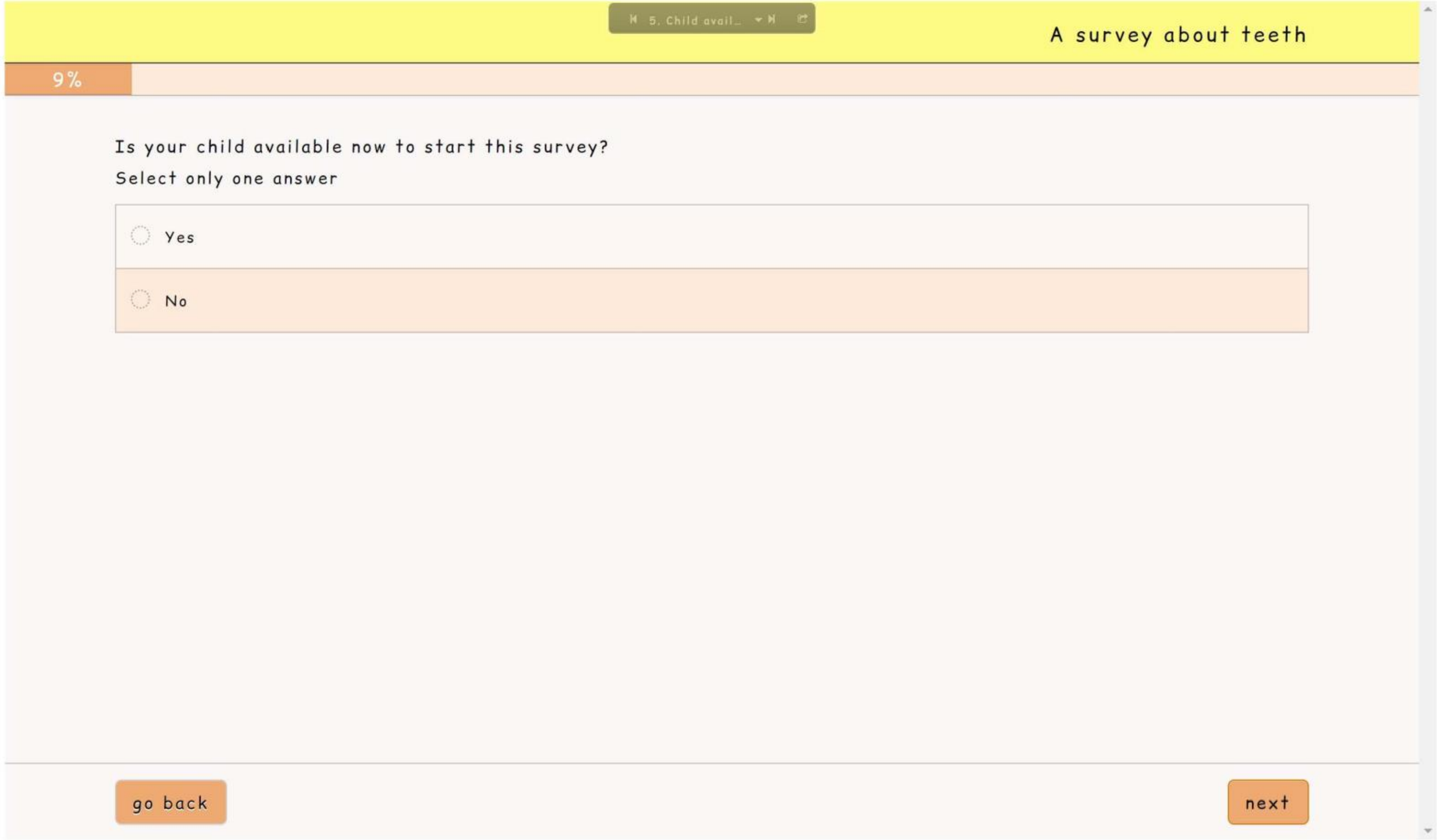
Before your child can start the survey, please confirm that you:

- are 18 years of age or over
- have read all of the important survey information
- are the parent/carer of the child who will be invited to complete the survey
- consent to your child taking part in the survey

Select only one answer

● Yes

○ No



Is your child available now to start this survey?

Select only one answer

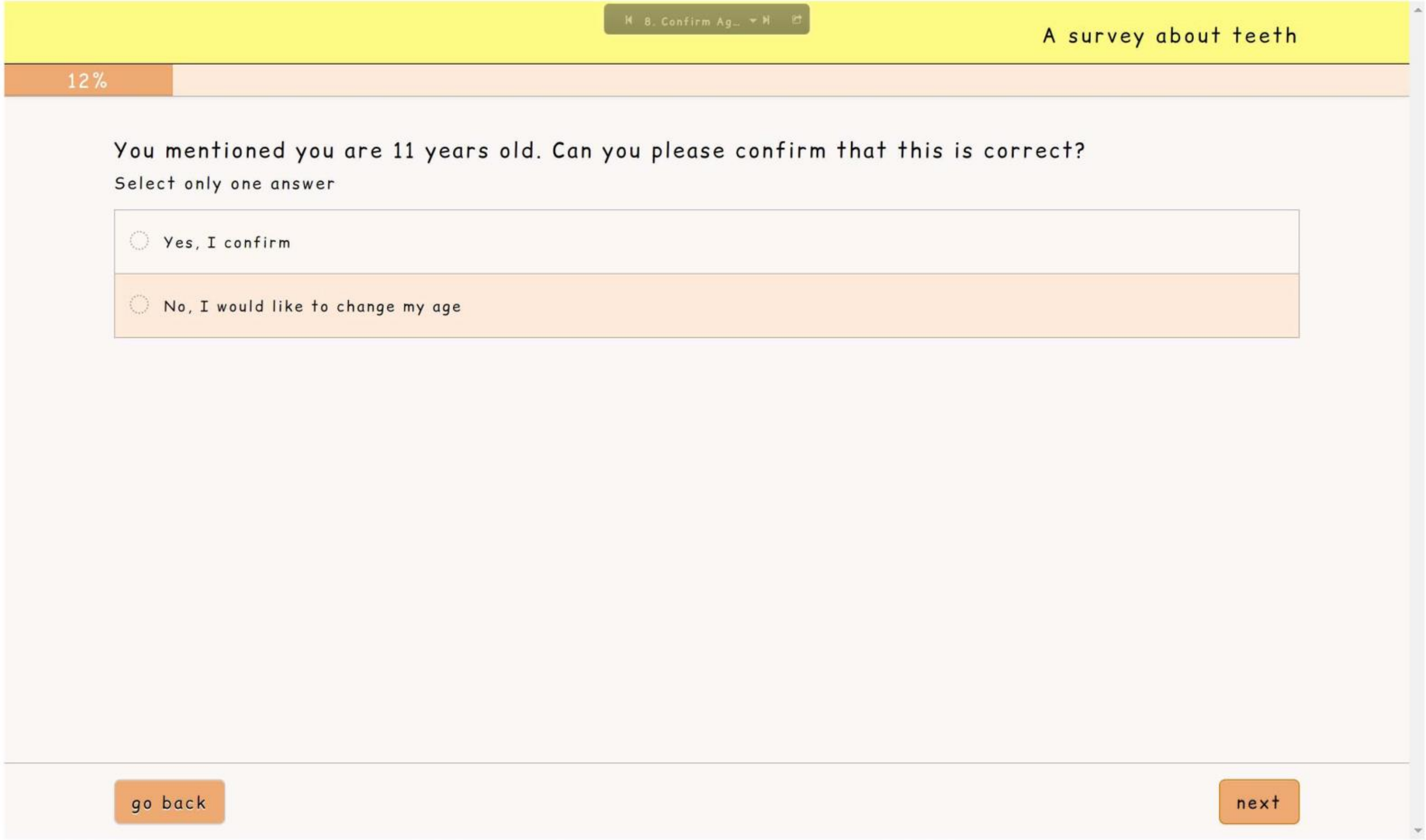
○ Yes

○ No

How old are you (in years)?

Select only one answer

- ○ 11
- ○ 12
- ○ 13
- ○ 14
- ○ 15
- ○ 16
- ○ Prefer not to say

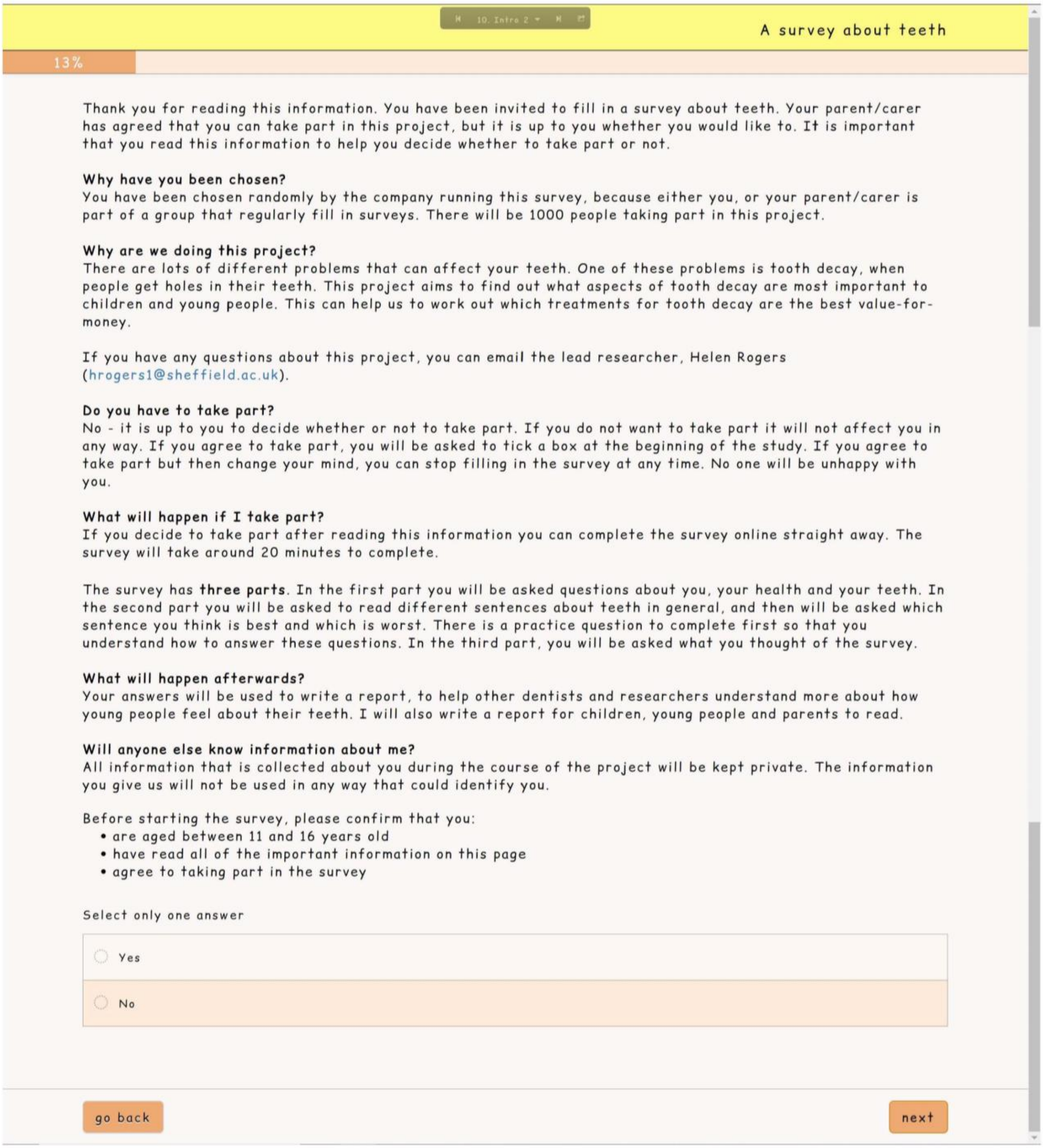


You mentioned you are 11 years old. Can you please confirm that this is correct?

Select only one answer

○ Yes, I confirm

○ No, I would like to change my age



A survey about teeth

Thank you for reading this information. You have been invited to fill in a survey about teeth. Your parent/carer has agreed that you can take part in this project, but it is up to you whether you would like to. It is important that you read this information to help you decide whether to take part or not.

Why have you been chosen?
You have been chosen randomly by the company running this survey, because either you, or your parent/carer is part of a group that regularly fill in surveys. There will be 1000 people taking part in this project.

Why are we doing this project?
There are lots of different problems that can affect your teeth. One of these problems is tooth decay, when people get holes in their teeth. This project aims to find out what aspects of tooth decay are most important to children and young people. This can help us to work out which treatments for tooth decay are the best value-for-money.

If you have any questions about this project, you can email the lead researcher, Helen Rogers (hrogers1@sheffield.ac.uk).

Do you have to take part?
No - it is up to you to decide whether or not to take part. If you do not want to take part it will not affect you in any way. If you agree to take part, you will be asked to tick a box at the beginning of the study. If you agree to take part but then change your mind, you can stop filling in the survey at any time. No one will be unhappy with you.

What will happen if I take part?
If you decide to take part after reading this information you can complete the survey online straight away. The survey will take around 20 minutes to complete.

The survey has **three parts**. In the first part you will be asked questions about you, your health and your teeth. In the second part you will be asked to read different sentences about teeth in general, and then will be asked which sentence you think is best and which is worst. There is a practice question to complete first so that you understand how to answer these questions. In the third part, you will be asked what you thought of the survey.

What will happen afterwards?
Your answers will be used to write a report, to help other dentists and researchers understand more about how young people feel about their teeth. I will also write a report for children, young people and parents to read.

Will anyone else know information about me?
All information that is collected about you during the course of the project will be kept private. The information you give us will not be used in any way that could identify you.

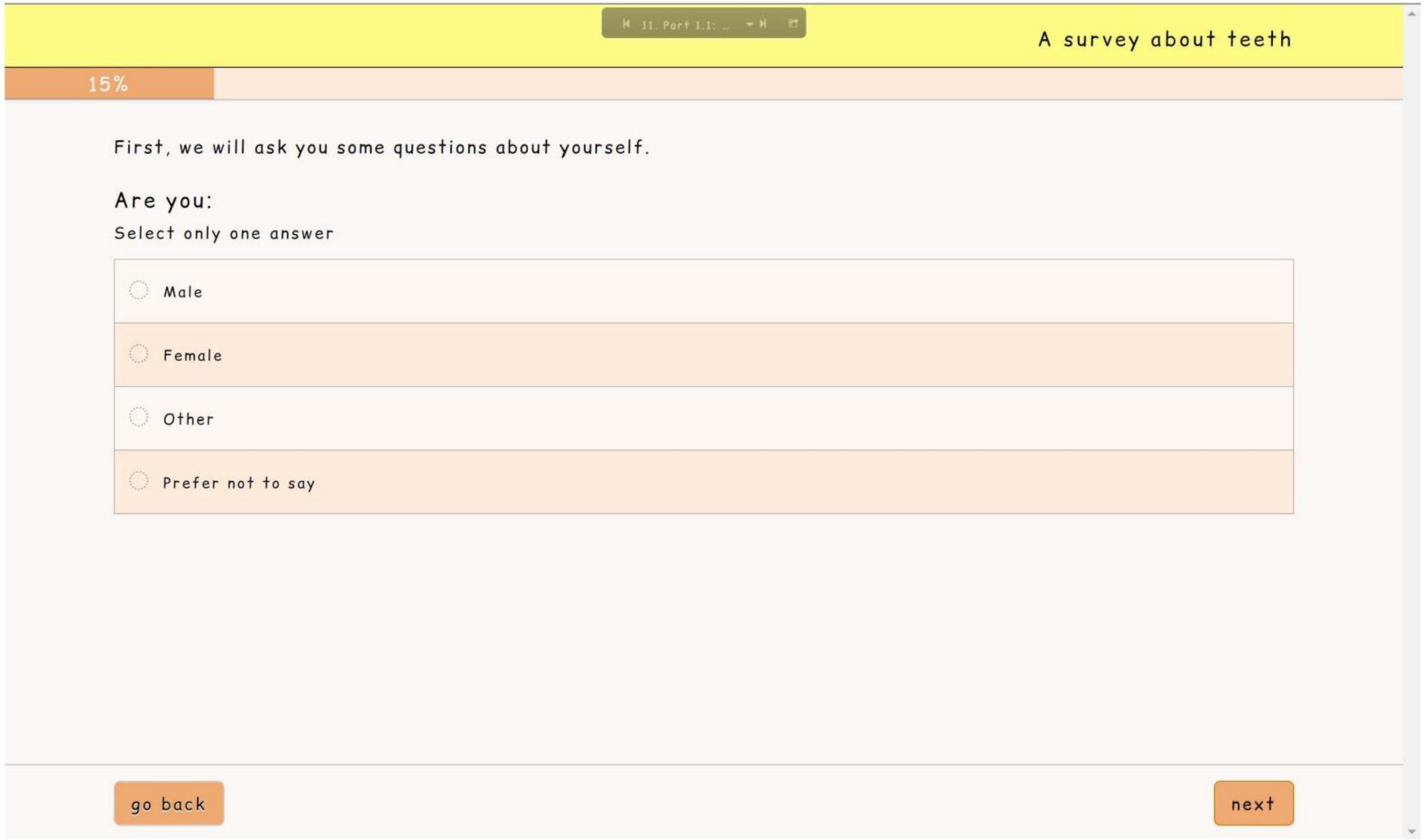
Before starting the survey, please confirm that you:

- are aged between 11 and 16 years old
- have read all of the important information on this page
- agree to taking part in the survey

Select only one answer

○ Yes

○ No



First, we will ask you some questions about yourself.

Are you:

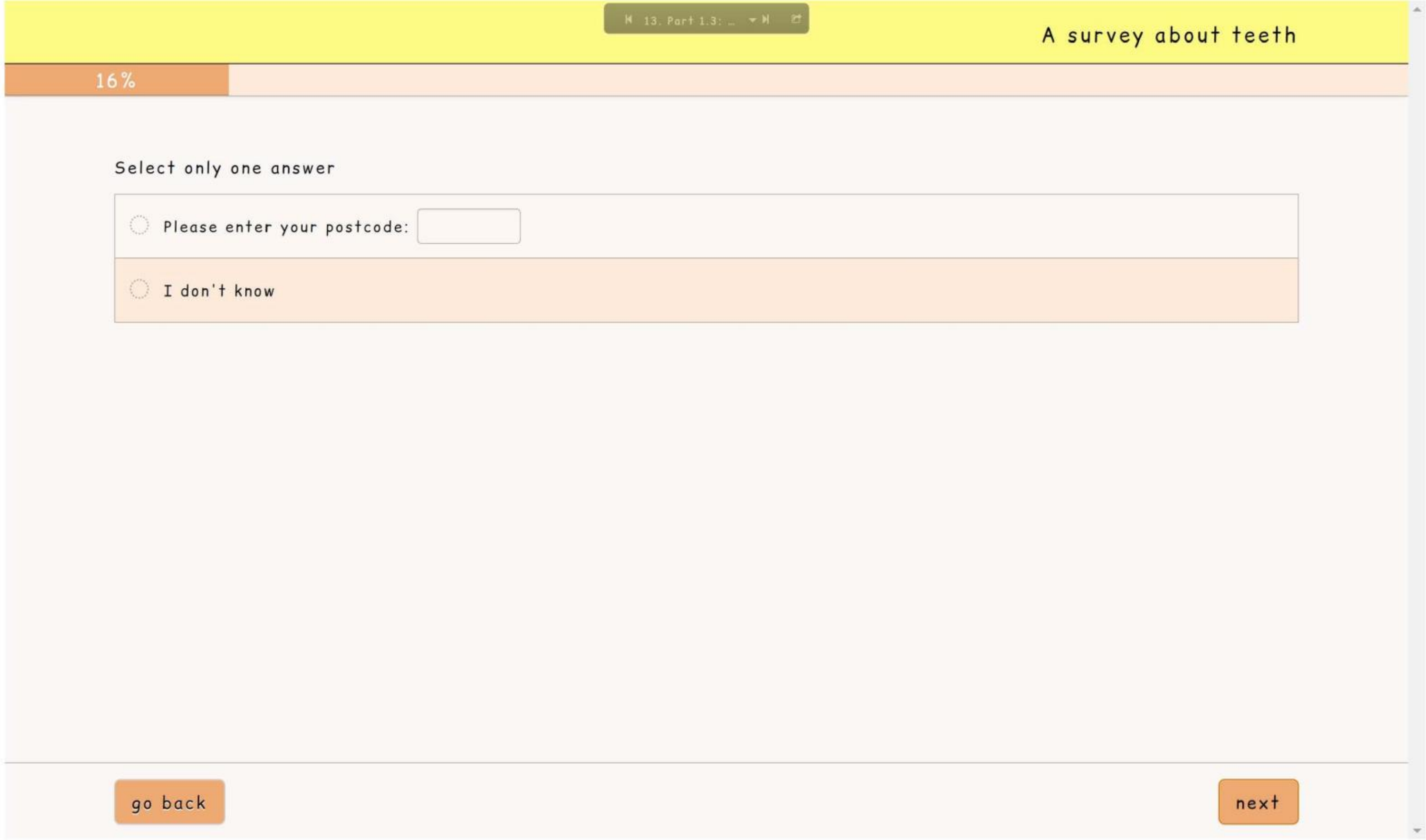
Select only one answer

○ Male

○ Female

○ Other

○ Prefer not to say

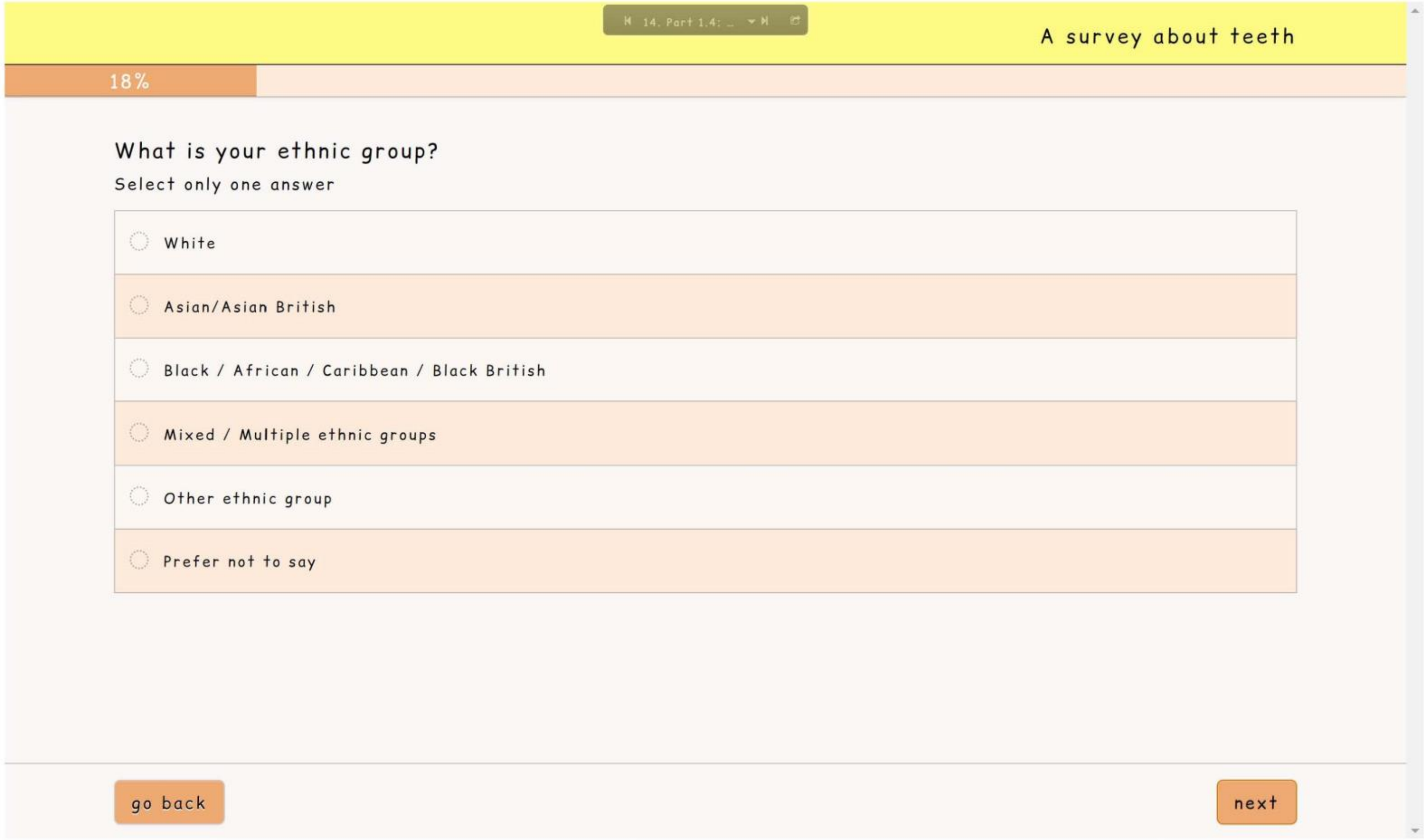


A survey about teeth

Select only one answer

○ Please enter your postcode:

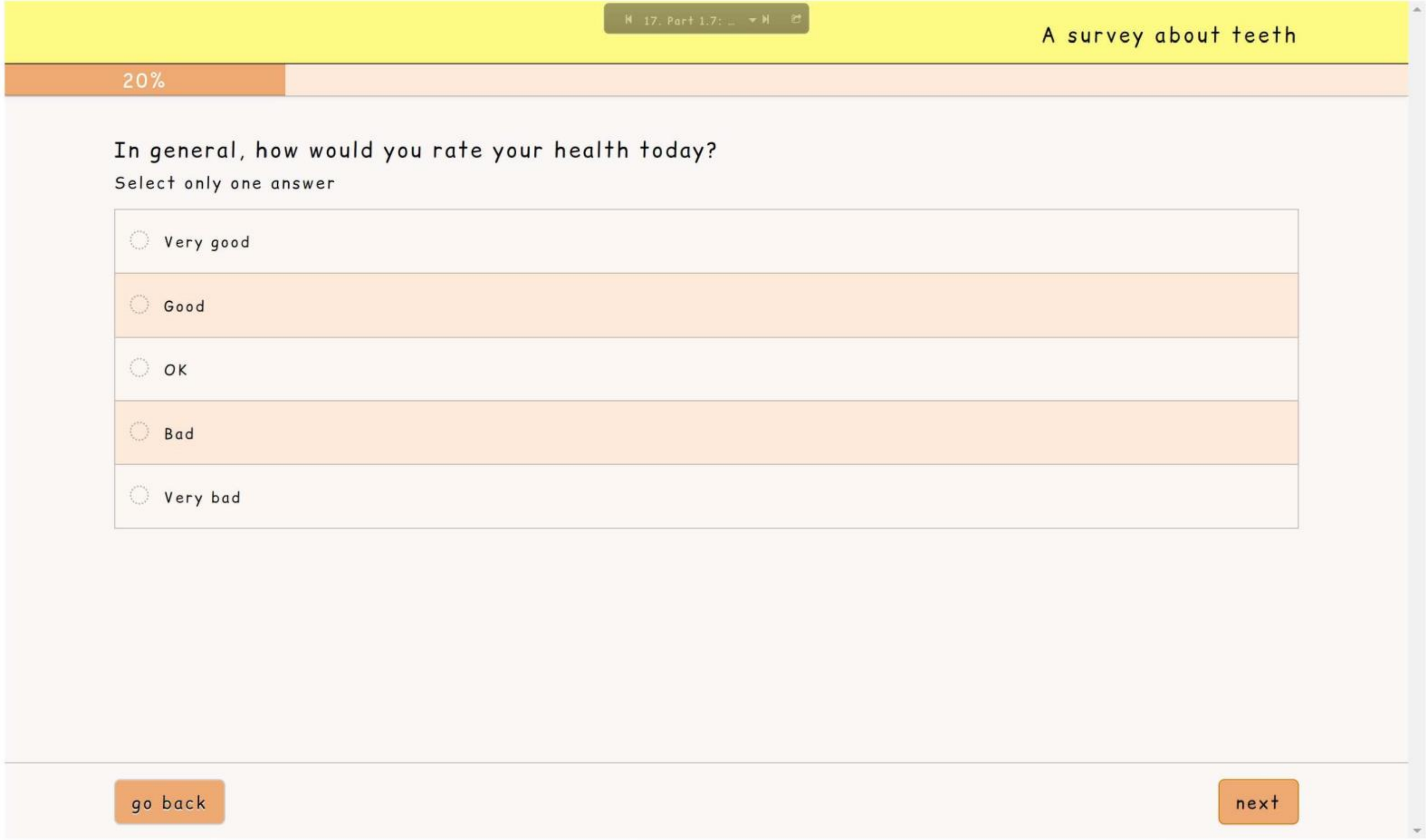
○ I don't know



What is your ethnic group?

Select only one answer

- ○ White
- ○ Asian/Asian British
- ○ Black / African / Caribbean / Black British
- ○ Mixed / Multiple ethnic groups
- ○ Other ethnic group
- ○ Prefer not to say



In general, how would you rate your health today?

Select only one answer

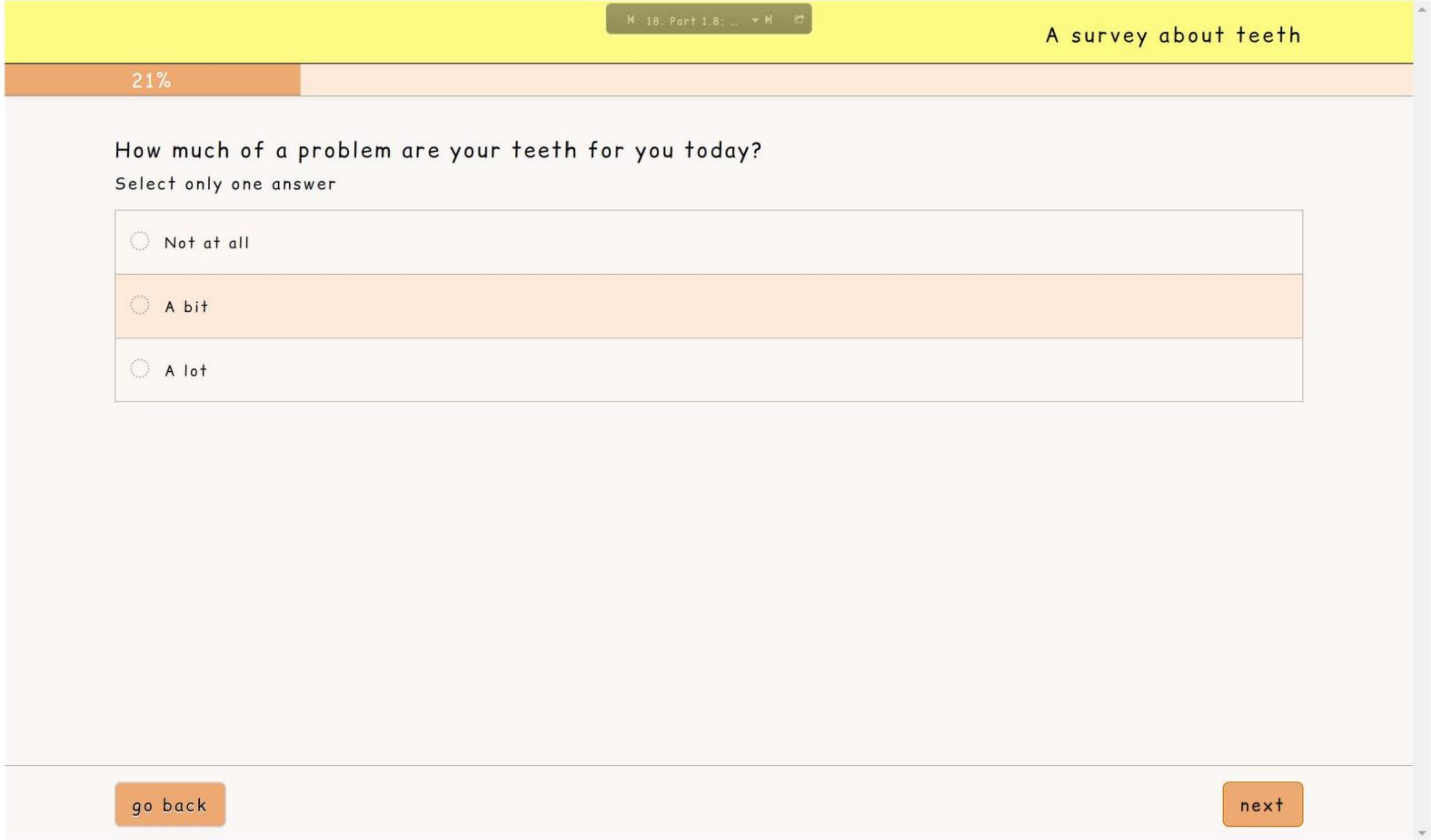
○ Very good

○ Good

○ OK

○ Bad

○ Very bad



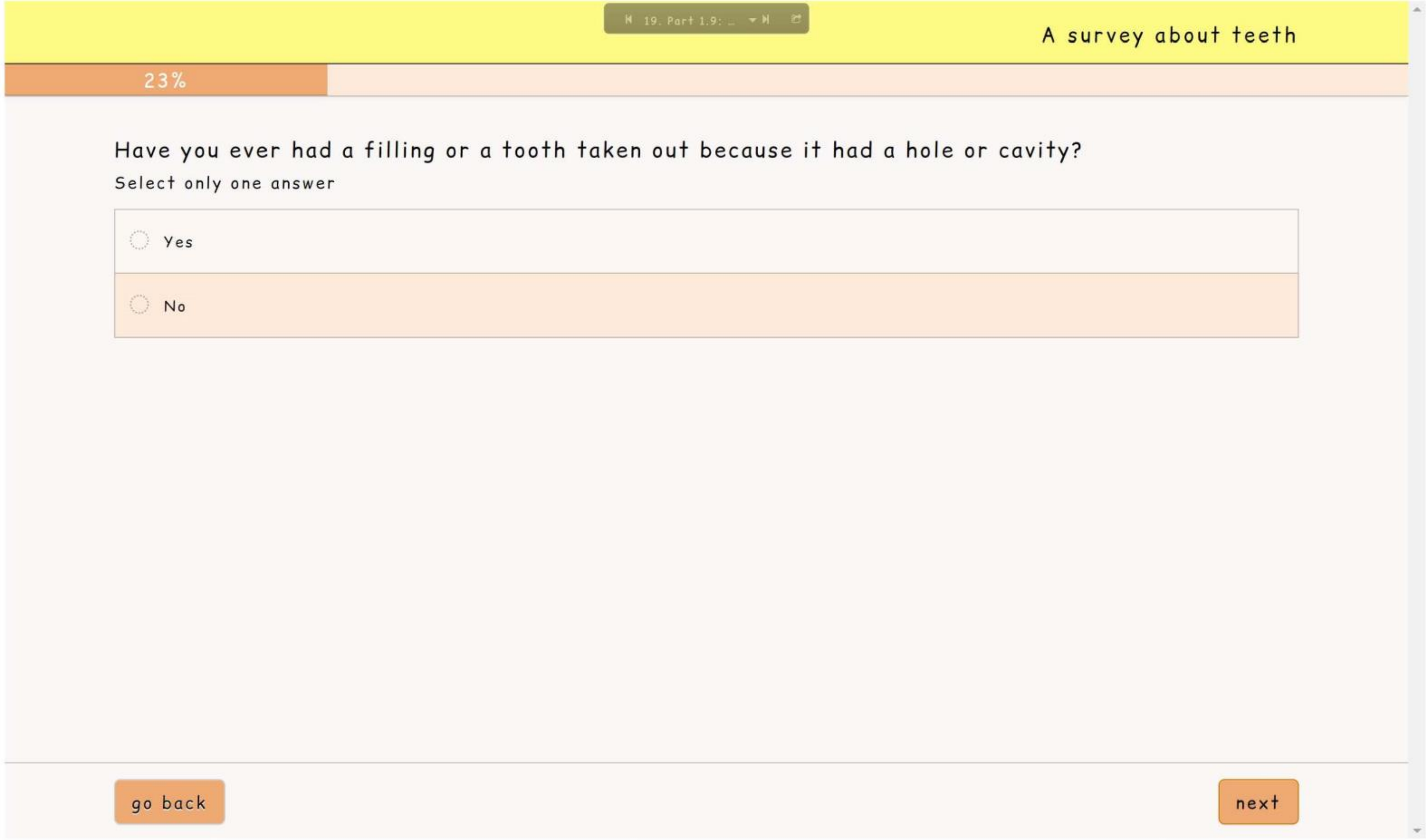
How much of a problem are your teeth for you today?

Select only one answer

○ Not at all

○ A bit

○ A lot

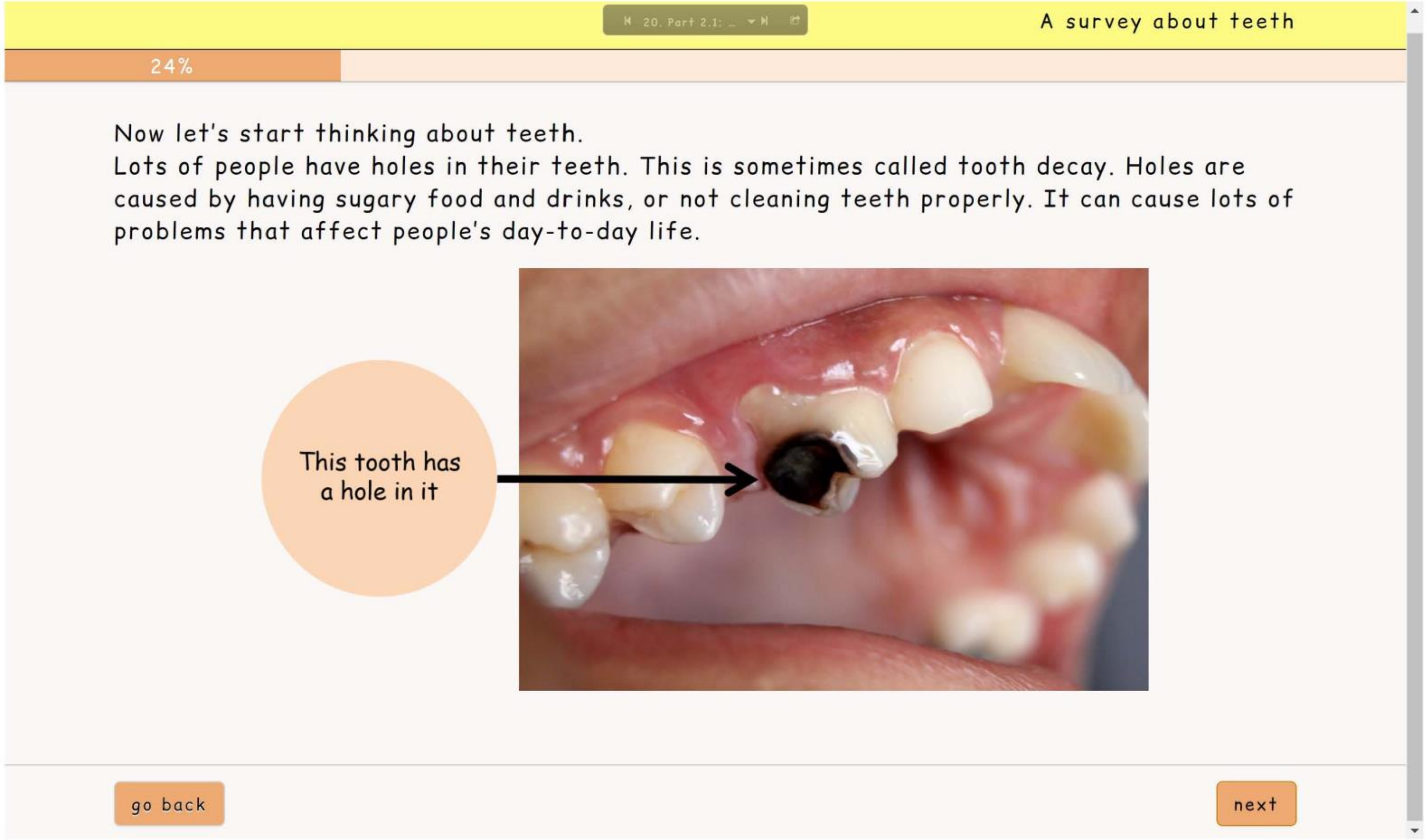


Have you ever had a filling or a tooth taken out because it had a hole or cavity?

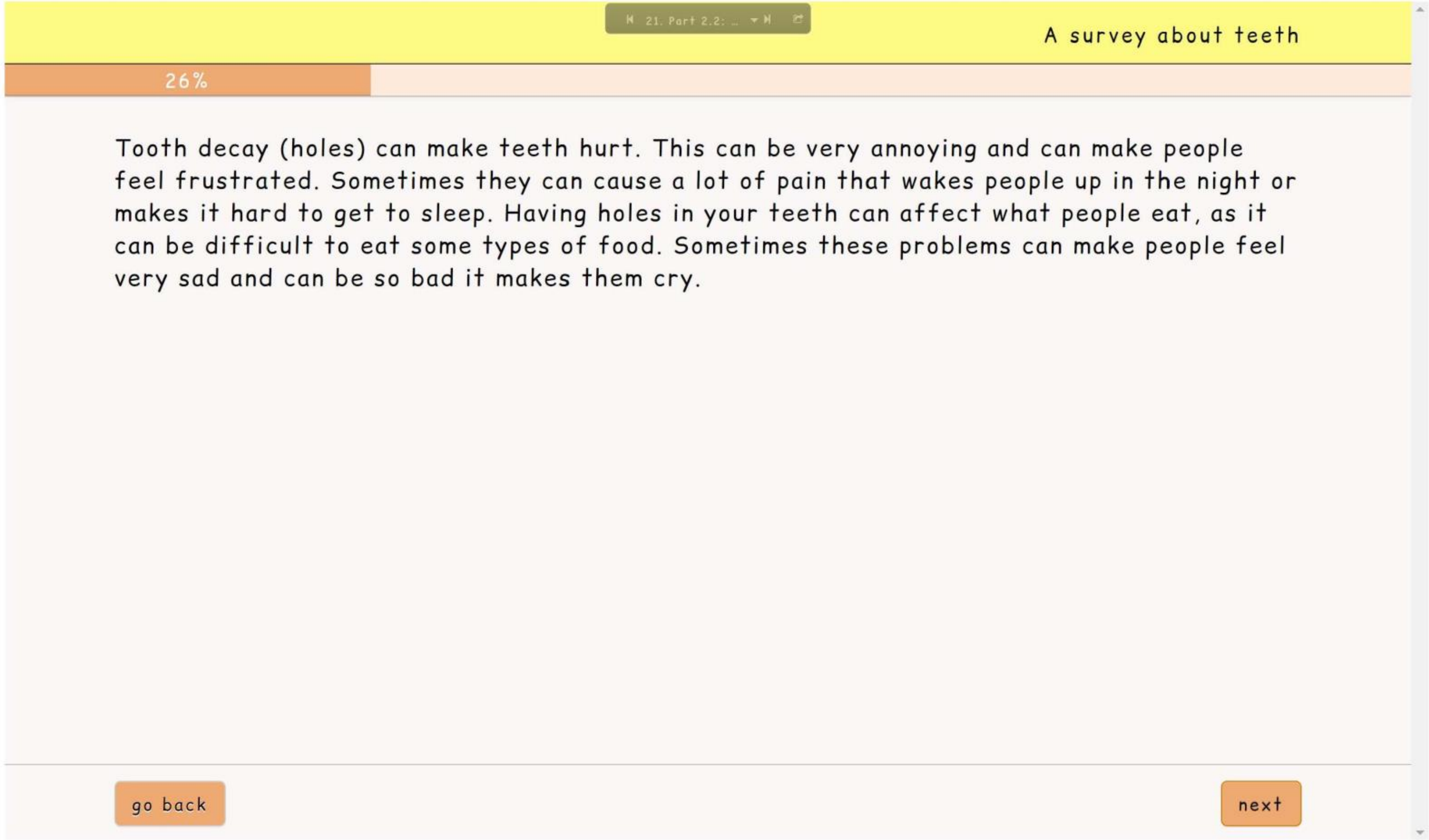
Select only one answer

○ Yes

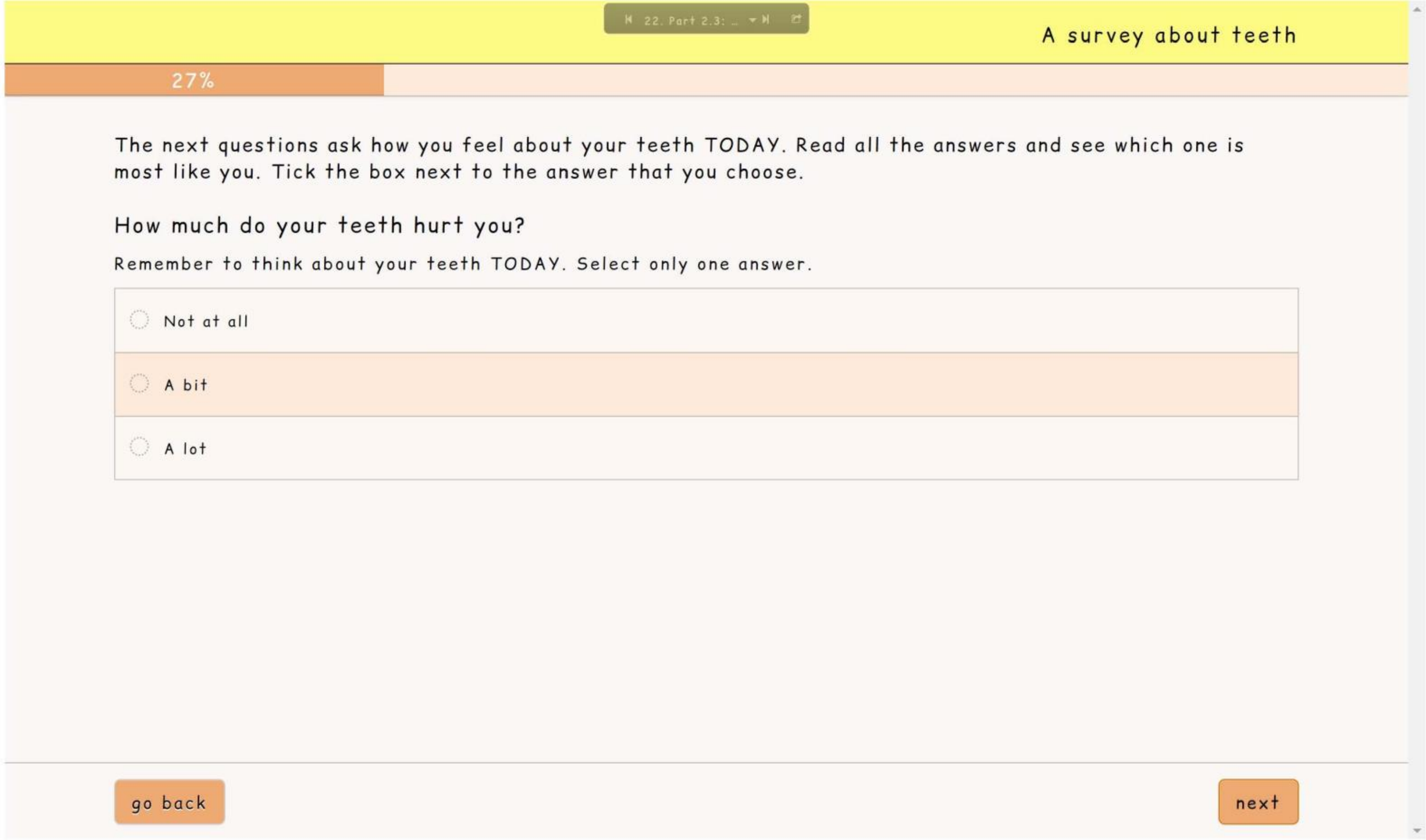
○ No



Now let's start thinking about teeth.
Lots of people have holes in their teeth. This is sometimes called tooth decay. Holes are caused by having sugary food and drinks, or not cleaning teeth properly. It can cause lots of problems that affect people's day-to-day life.



Tooth decay (holes) can make teeth hurt. This can be very annoying and can make people feel frustrated. Sometimes they can cause a lot of pain that wakes people up in the night or makes it hard to get to sleep. Having holes in your teeth can affect what people eat, as it can be difficult to eat some types of food. Sometimes these problems can make people feel very sad and can be so bad it makes them cry.



The next questions ask how you feel about your teeth TODAY. Read all the answers and see which one is most like you. Tick the box next to the answer that you choose.

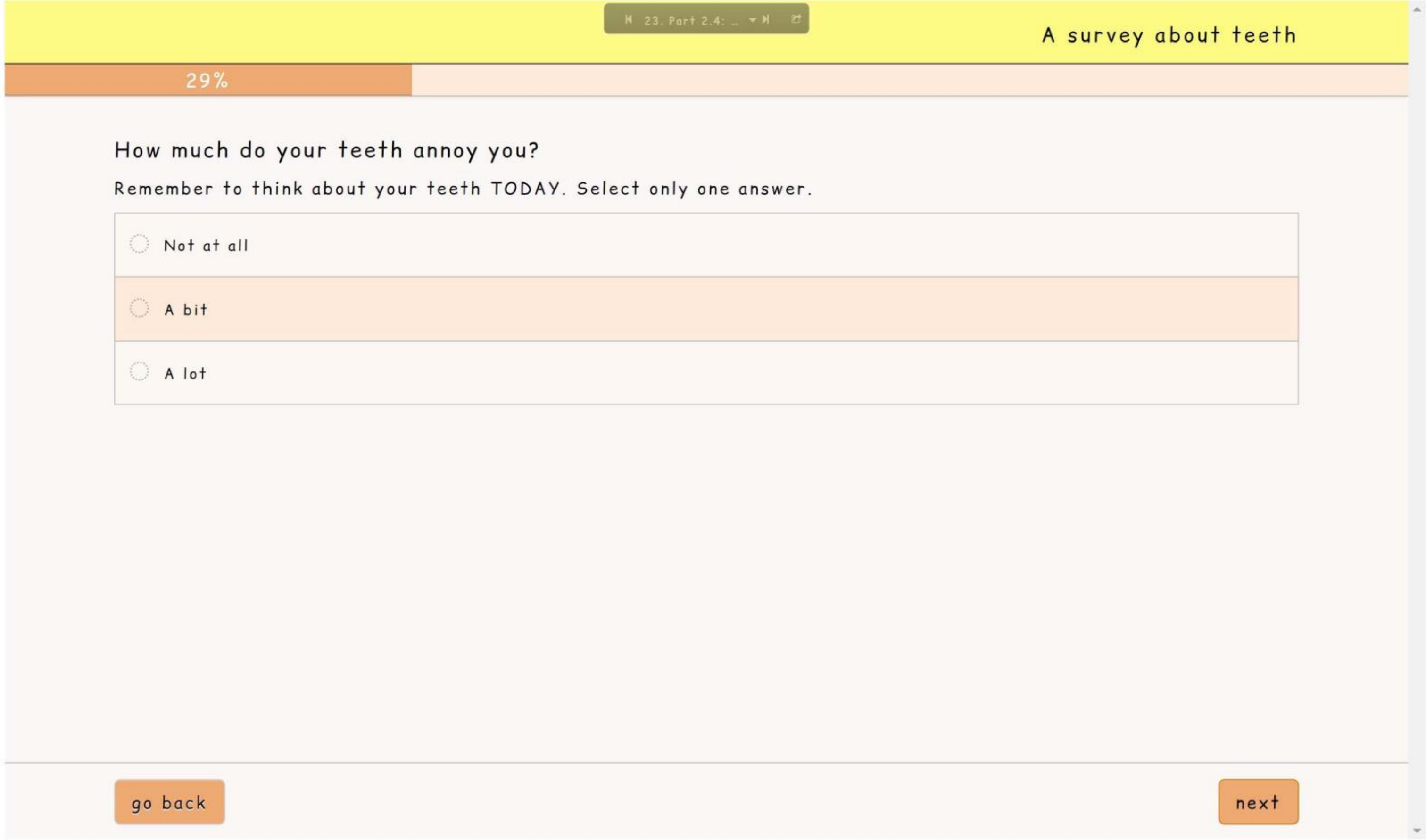
How much do your teeth hurt you?

Remember to think about your teeth TODAY. Select only one answer.

○ Not at all

○ A bit

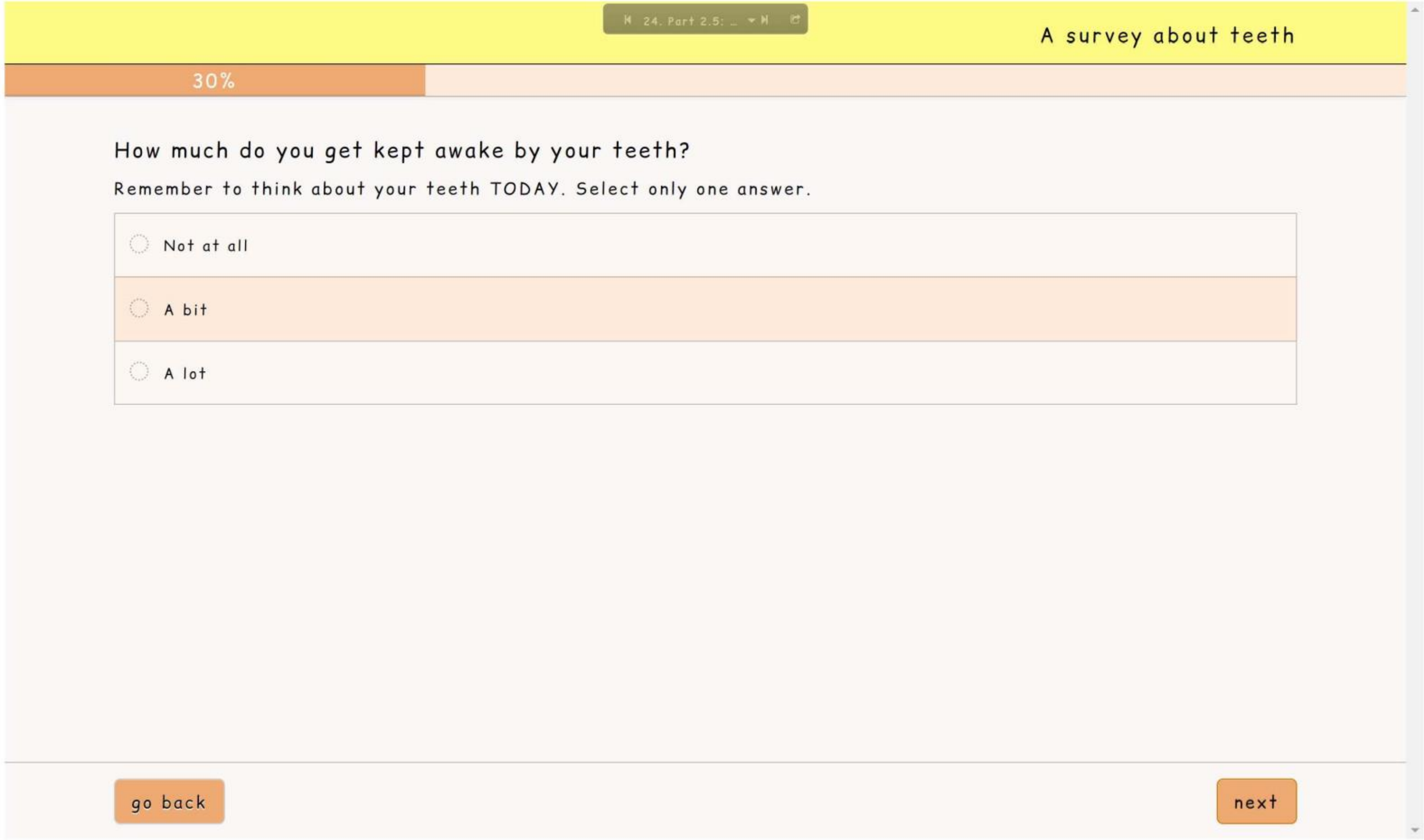
○ A lot



How much do your teeth annoy you?

Remember to think about your teeth TODAY. Select only one answer.

- ○ Not at all
- ○ A bit
- ○ A lot



How much do you get kept awake by your teeth?

Remember to think about your teeth TODAY. Select only one answer.

○ Not at all

○ A bit

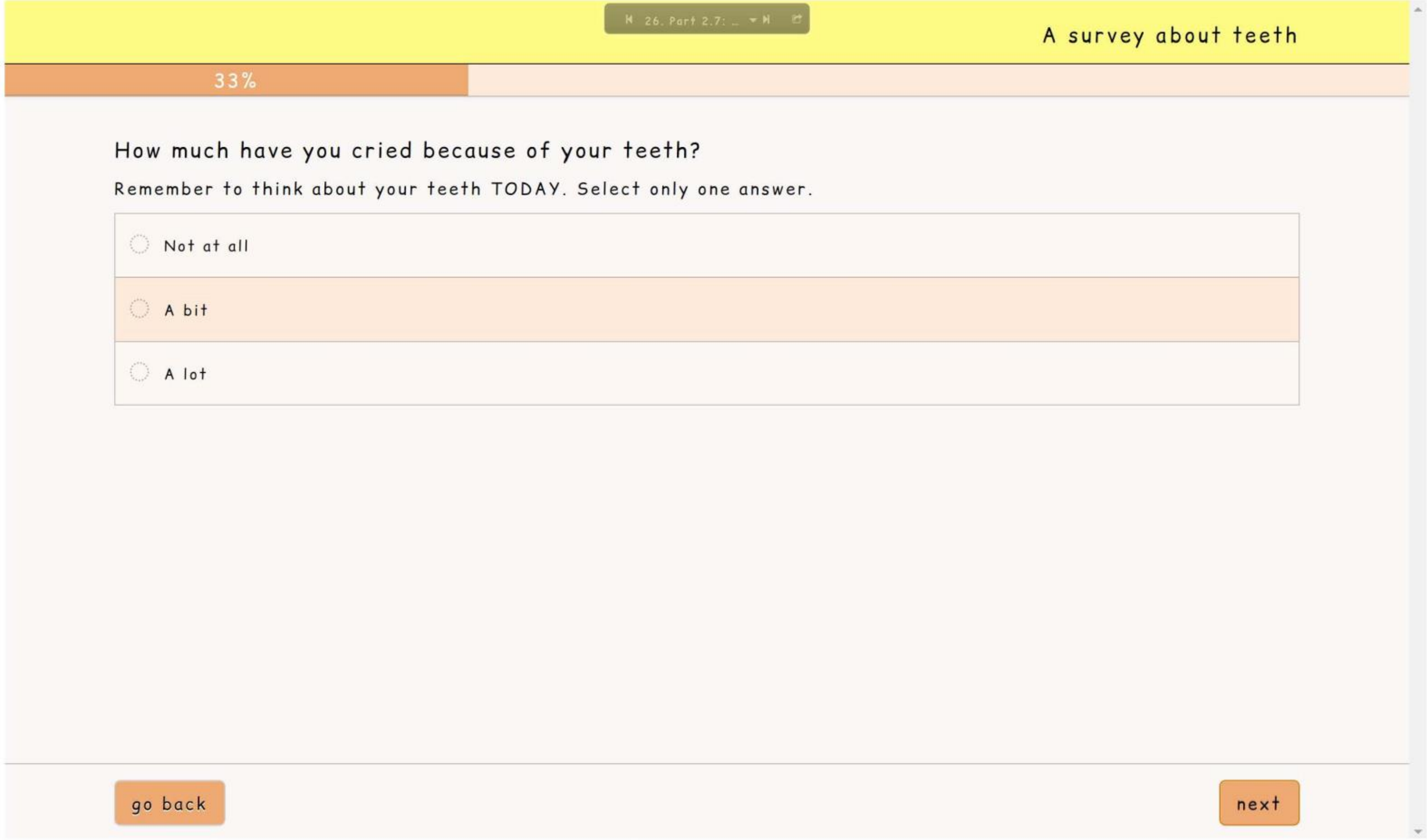
○ A lot

A survey about teeth

Do your teeth make it hard to eat some foods?

Remember to think about your teeth TODAY. Select only one answer.

○ Not at all
○ A bit
○ A lot



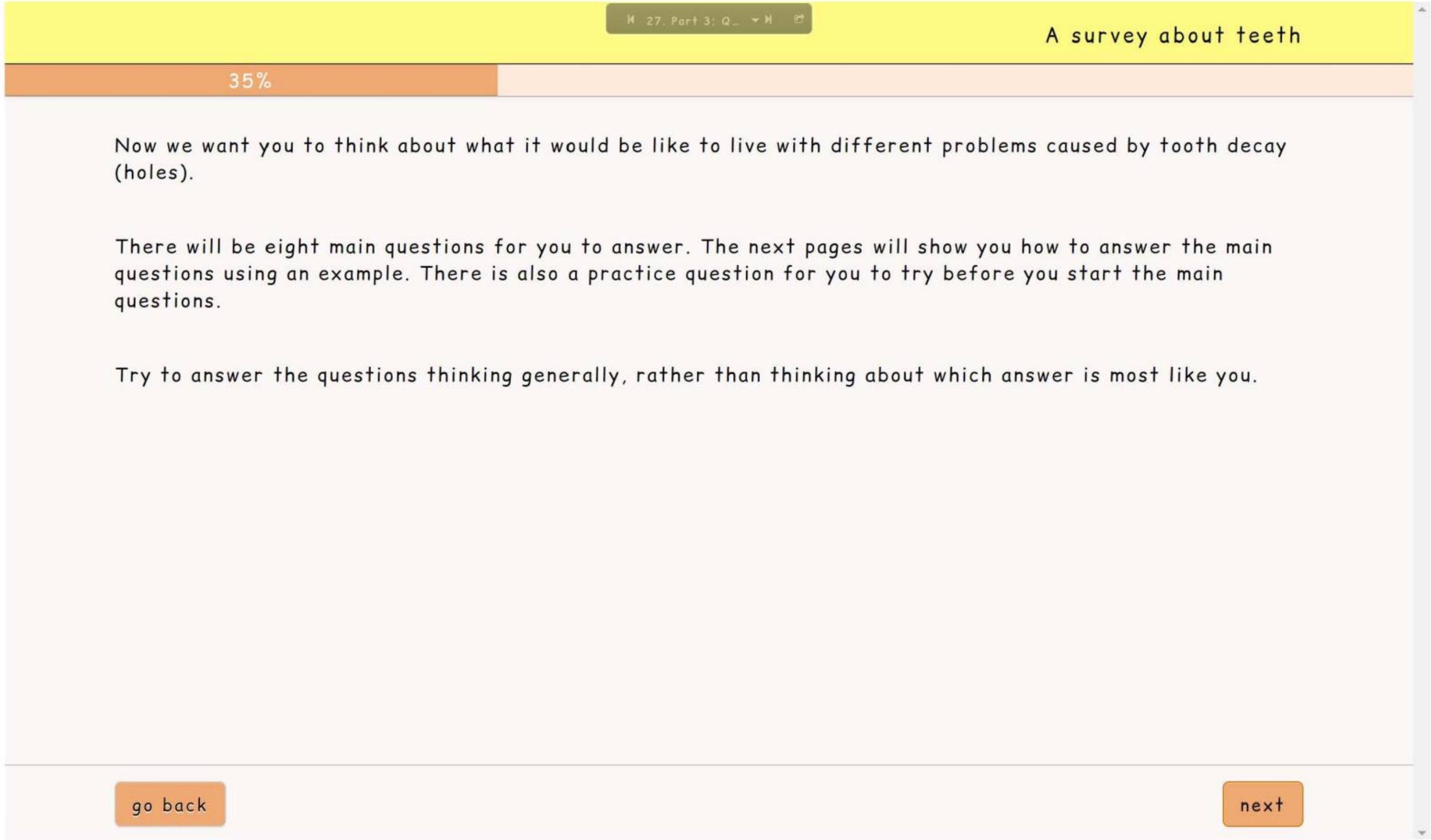
How much have you cried because of your teeth?

Remember to think about your teeth TODAY. Select only one answer.

○ Not at all

○ A bit

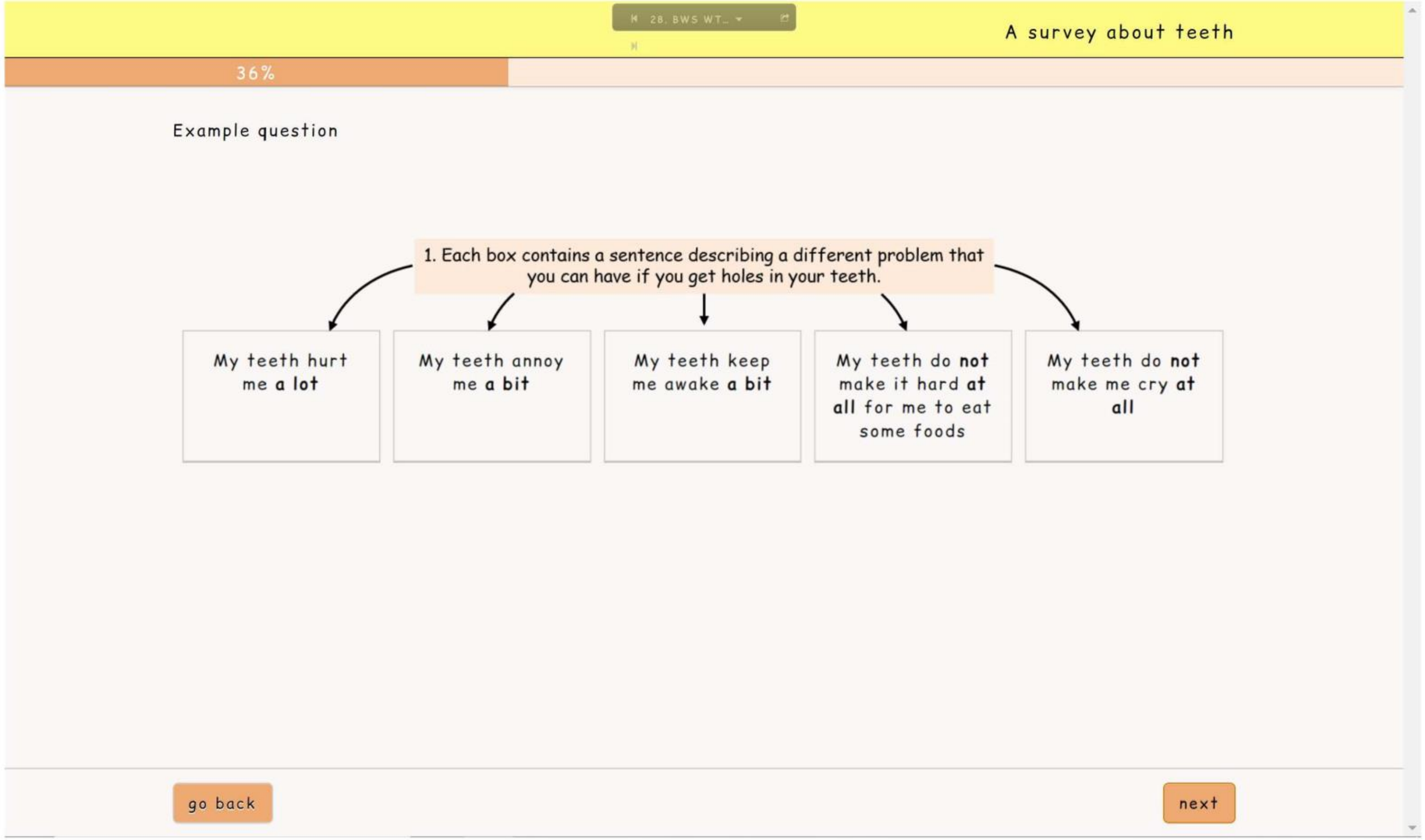
○ A lot



Now we want you to think about what it would be like to live with different problems caused by tooth decay (holes).

There will be eight main questions for you to answer. The next pages will show you how to answer the main questions using an example. There is also a practice question for you to try before you start the main questions.

Try to answer the questions thinking generally, rather than thinking about which answer is most like you.



Example question

1. Each box contains a sentence describing a different problem that you can have if you get holes in your teeth.

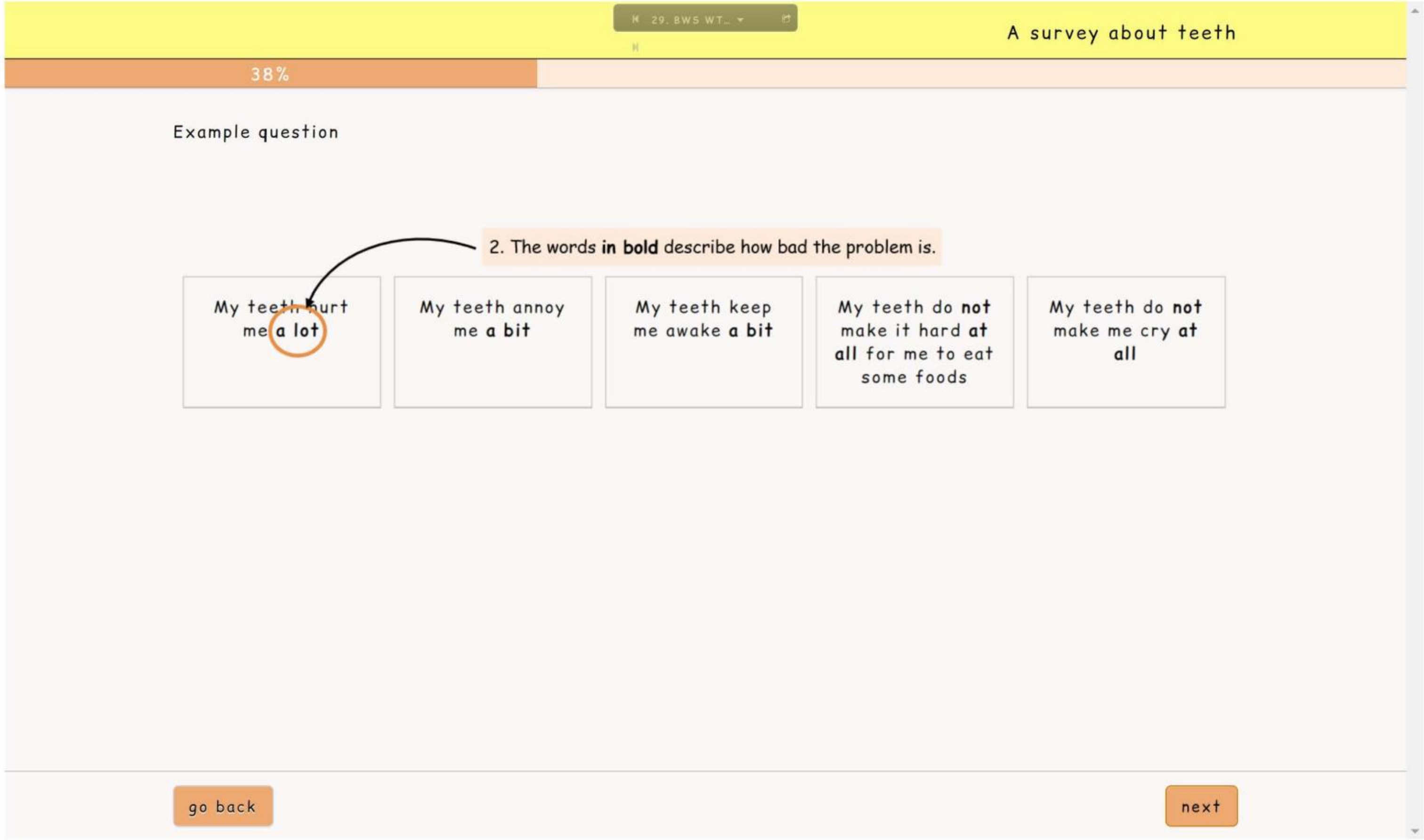
My teeth hurt me **a lot**

My teeth annoy me **a bit**

My teeth keep me awake **a bit**

My teeth do **not** make it hard **at all** for me to eat some foods

My teeth do **not** make me cry **at all**



A survey about teeth

Example question

2. The words **in bold** describe how bad the problem is.

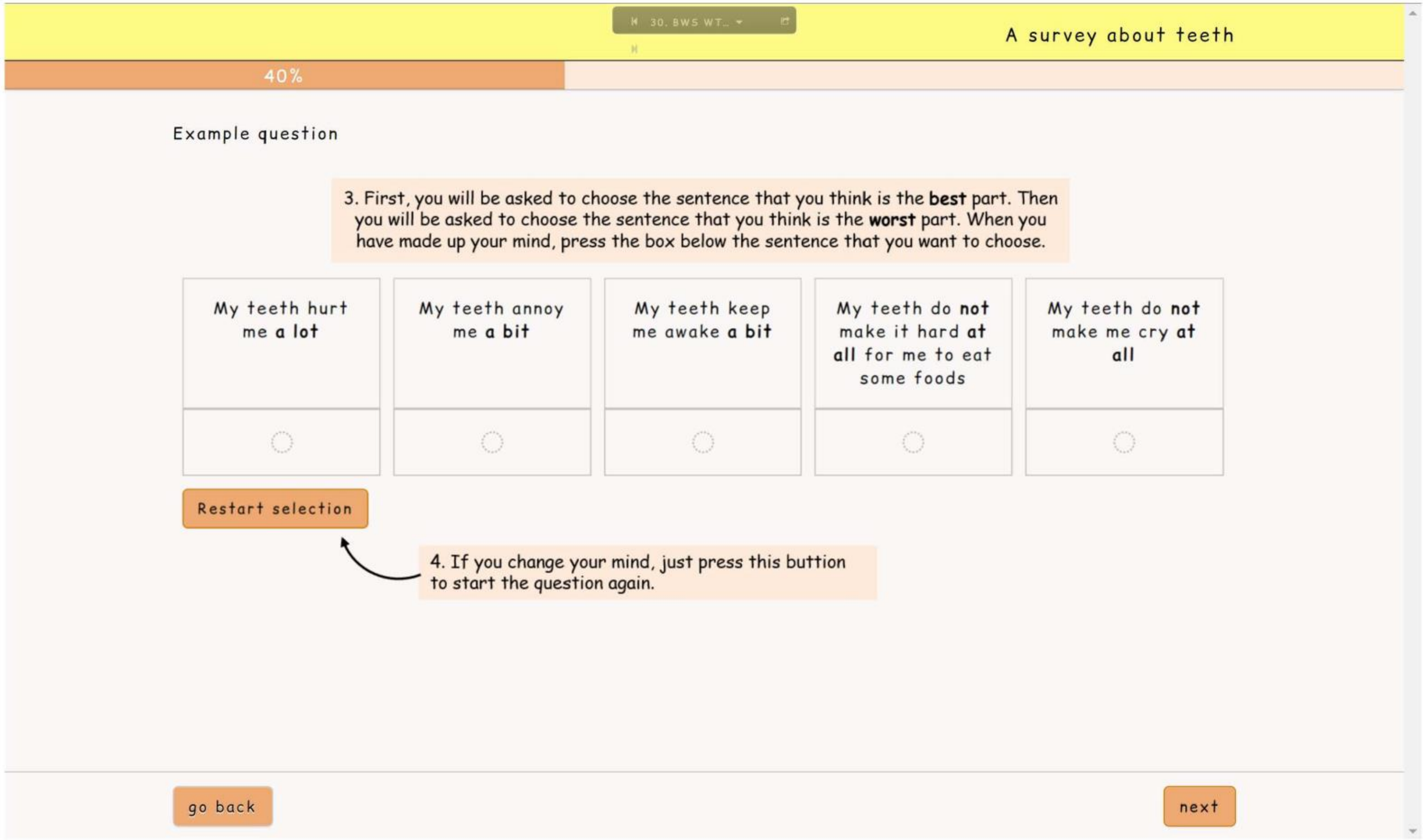
My teeth hurt me **a lot**

My teeth annoy me **a bit**

My teeth keep me awake **a bit**

My teeth do **not** make it hard **at all** for me to eat some foods

My teeth do **not** make me cry **at all**

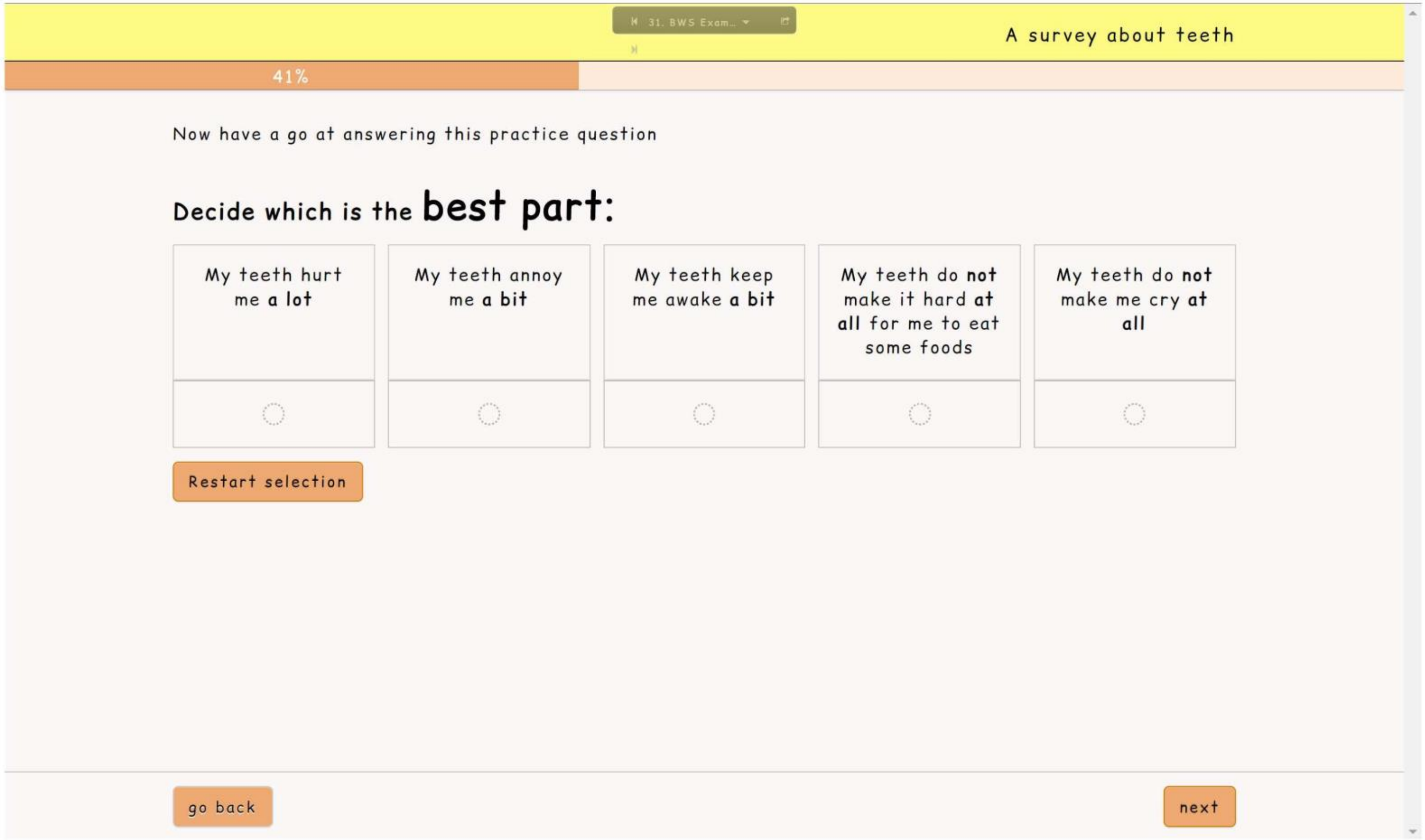


Example question

3. First, you will be asked to choose the sentence that you think is the **best** part. Then you will be asked to choose the sentence that you think is the **worst** part. When you have made up your mind, press the box below the sentence that you want to choose.

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○

4. If you change your mind, just press this buttion to start the question again.

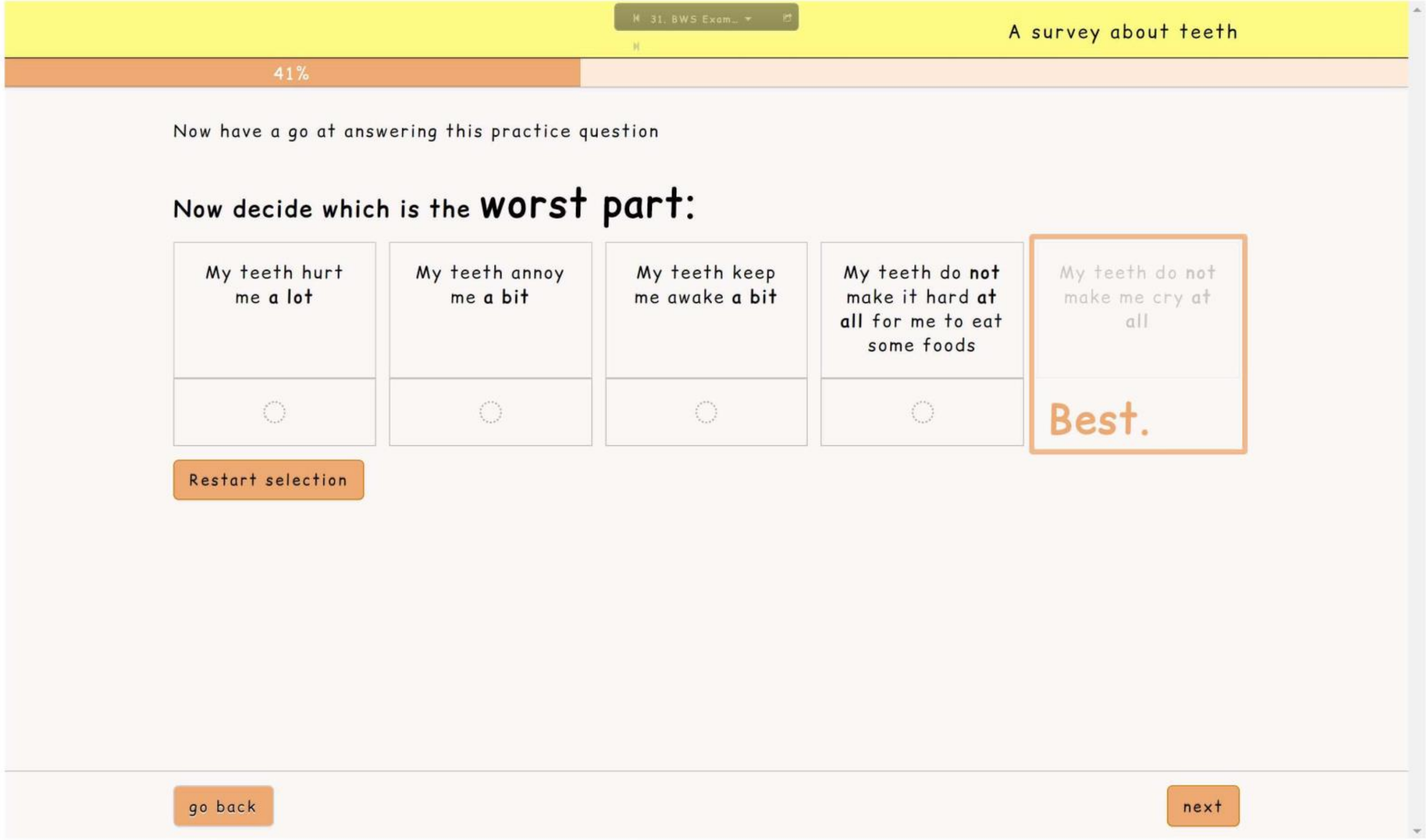


A survey about teeth

Now have a go at answering this practice question

Decide which is the **best part:**

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○



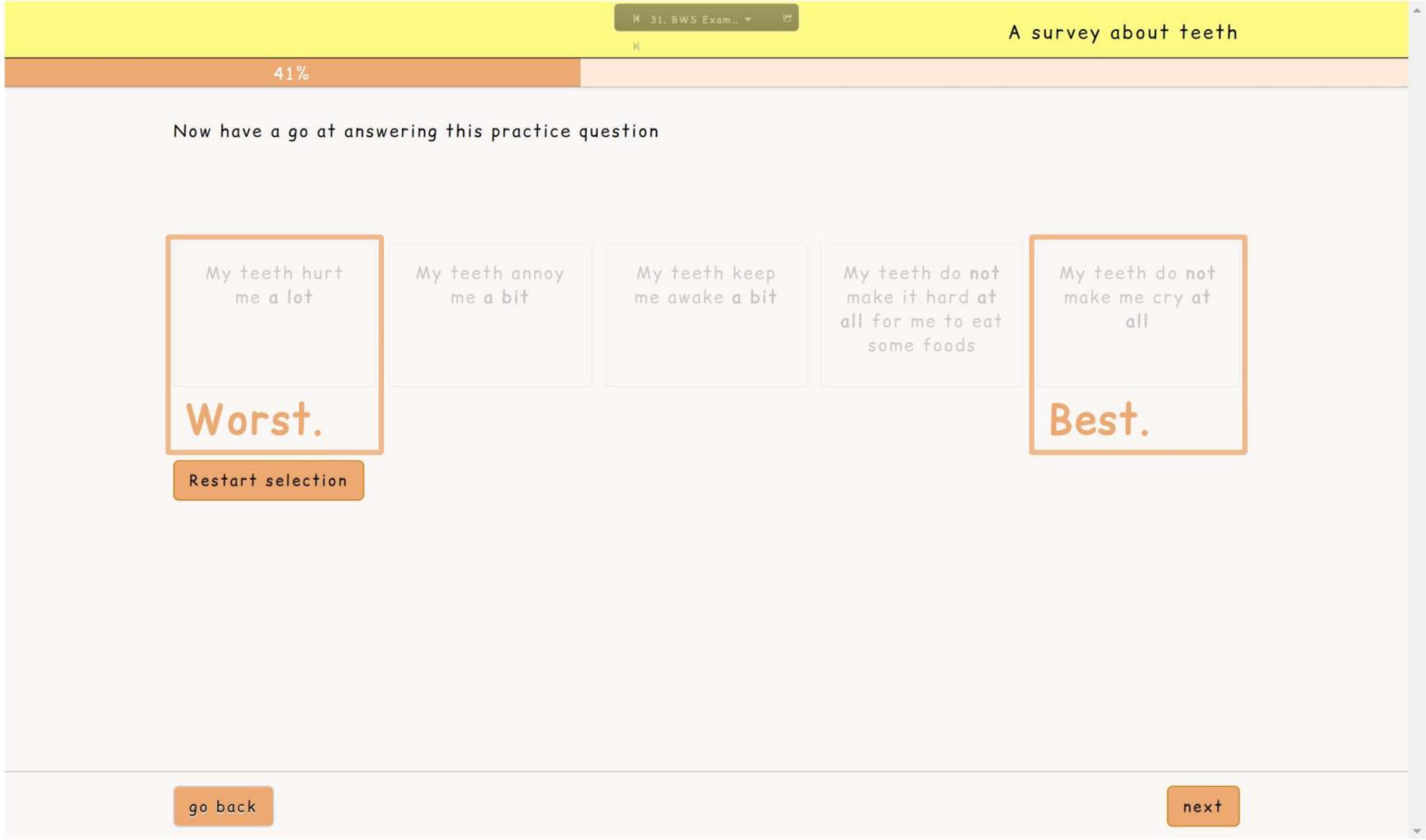
A survey about teeth

41%

Now have a go at answering this practice question

Now decide which is the worst part:

My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth keep me awake **a bit**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	**Best.**



Now have a go at answering this practice question

My teeth hurt me **a lot**

My teeth annoy me **a bit**

My teeth keep me awake **a bit**

My teeth do **not** make it hard **at all** for me to eat some foods

My teeth do **not** make me cry **at all**

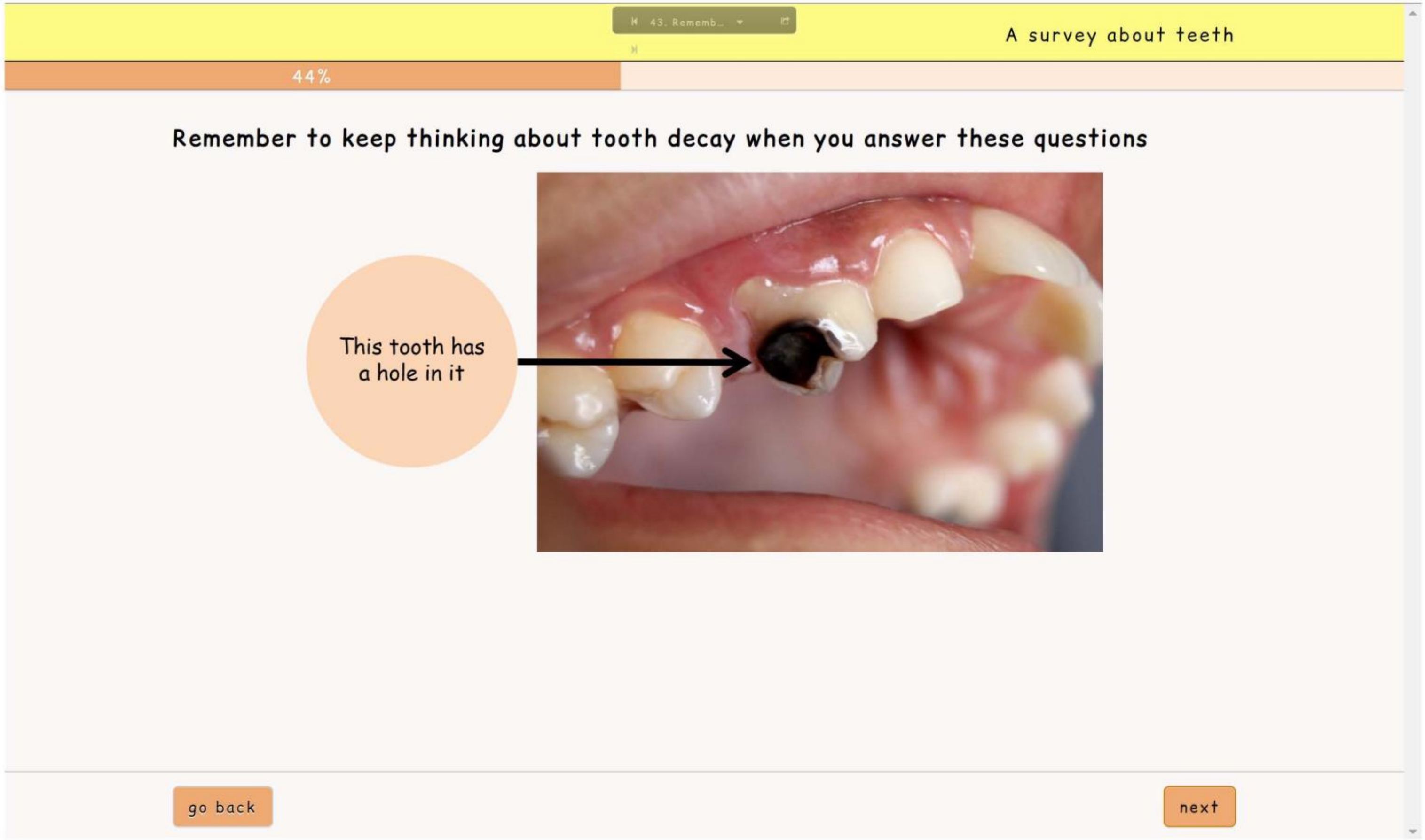
You have chosen **My teeth do not make me cry at all** as the best part of the description, and **My teeth hurt me a lot** as the worst part of the description.

Do you feel ready to answer some of these questions?

Select only one answer

○ Yes, start the main questions.

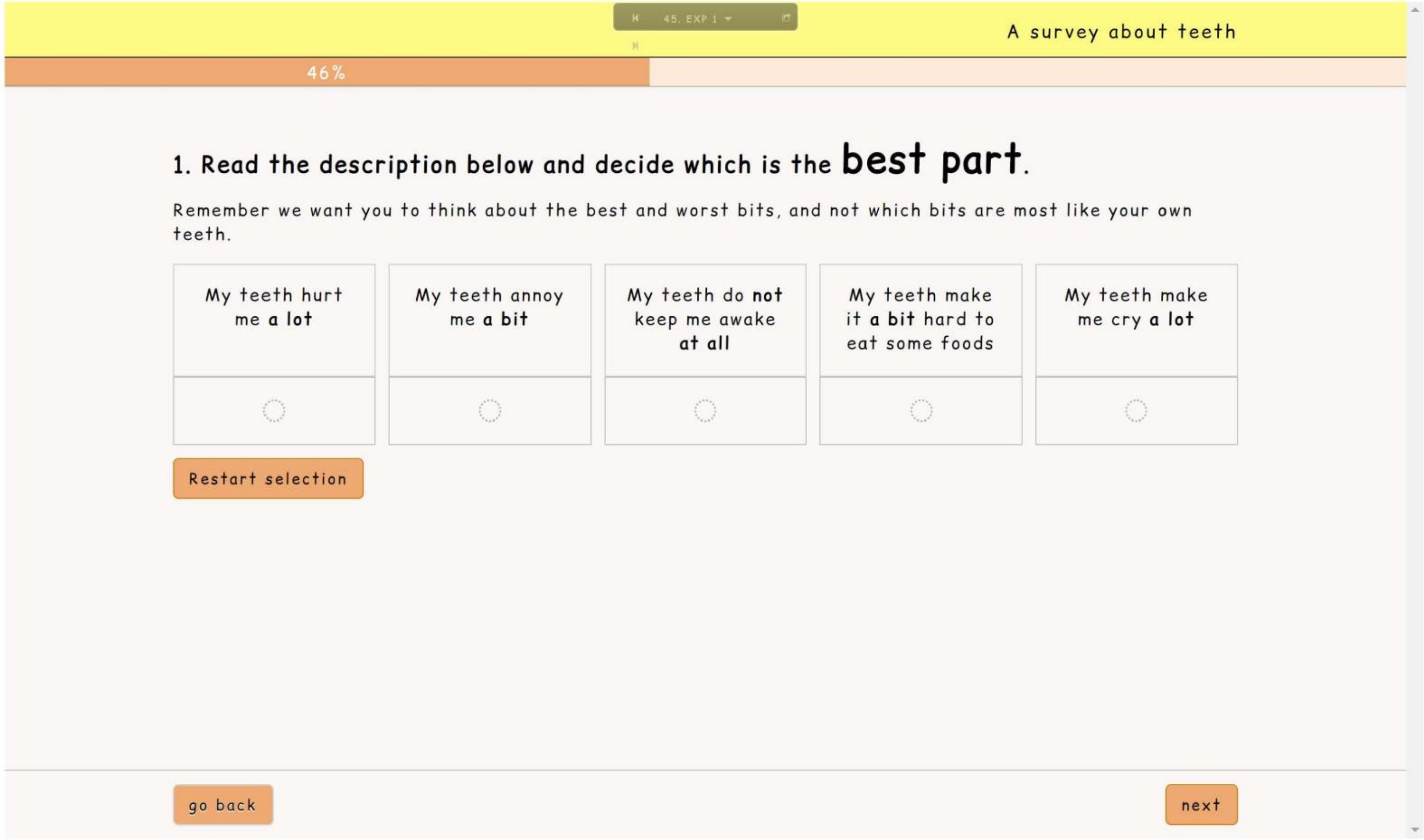
○ No, do the practice question again.



A survey about teeth

Remember to keep thinking about tooth decay when you answer these questions

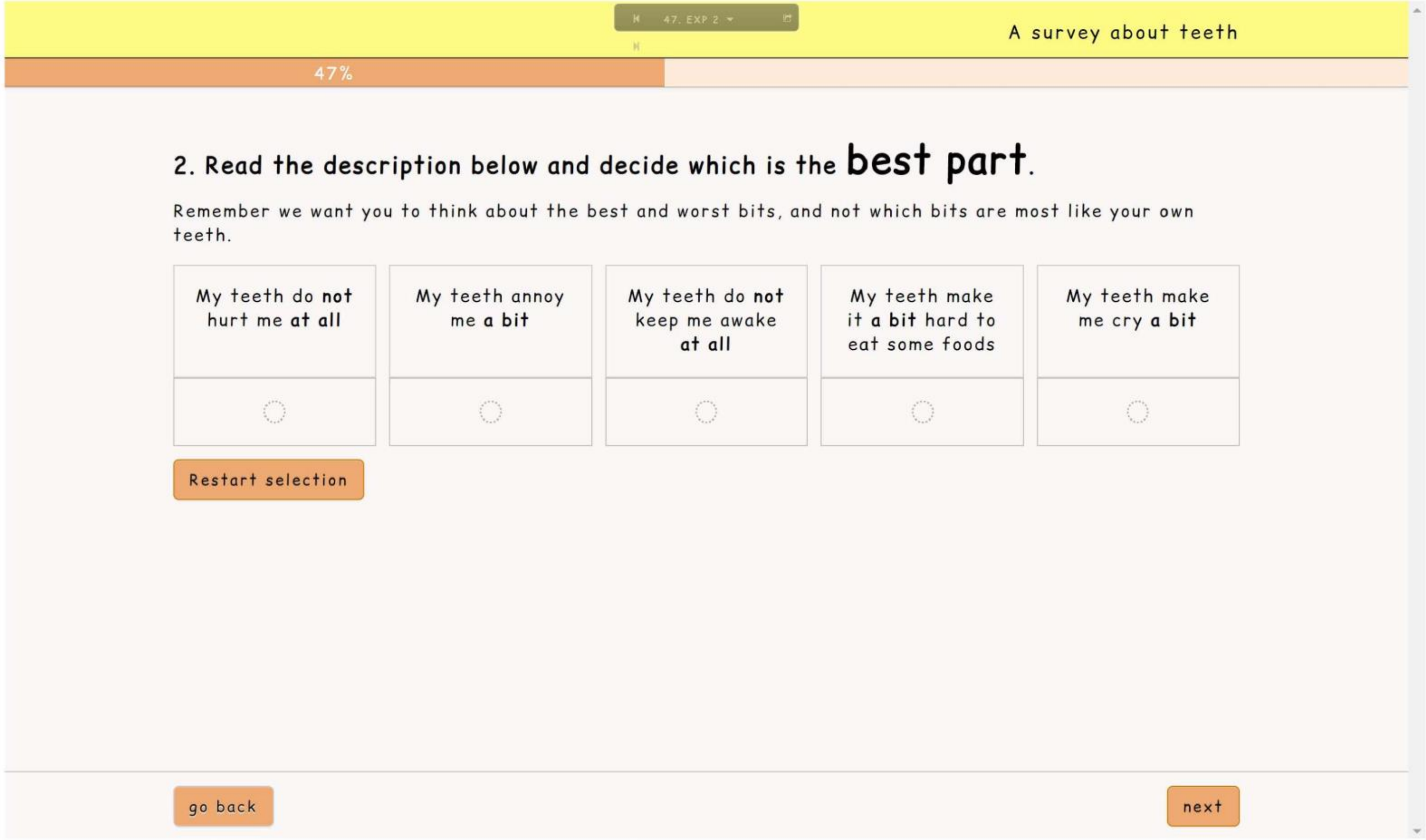
This tooth has a hole in it



1. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

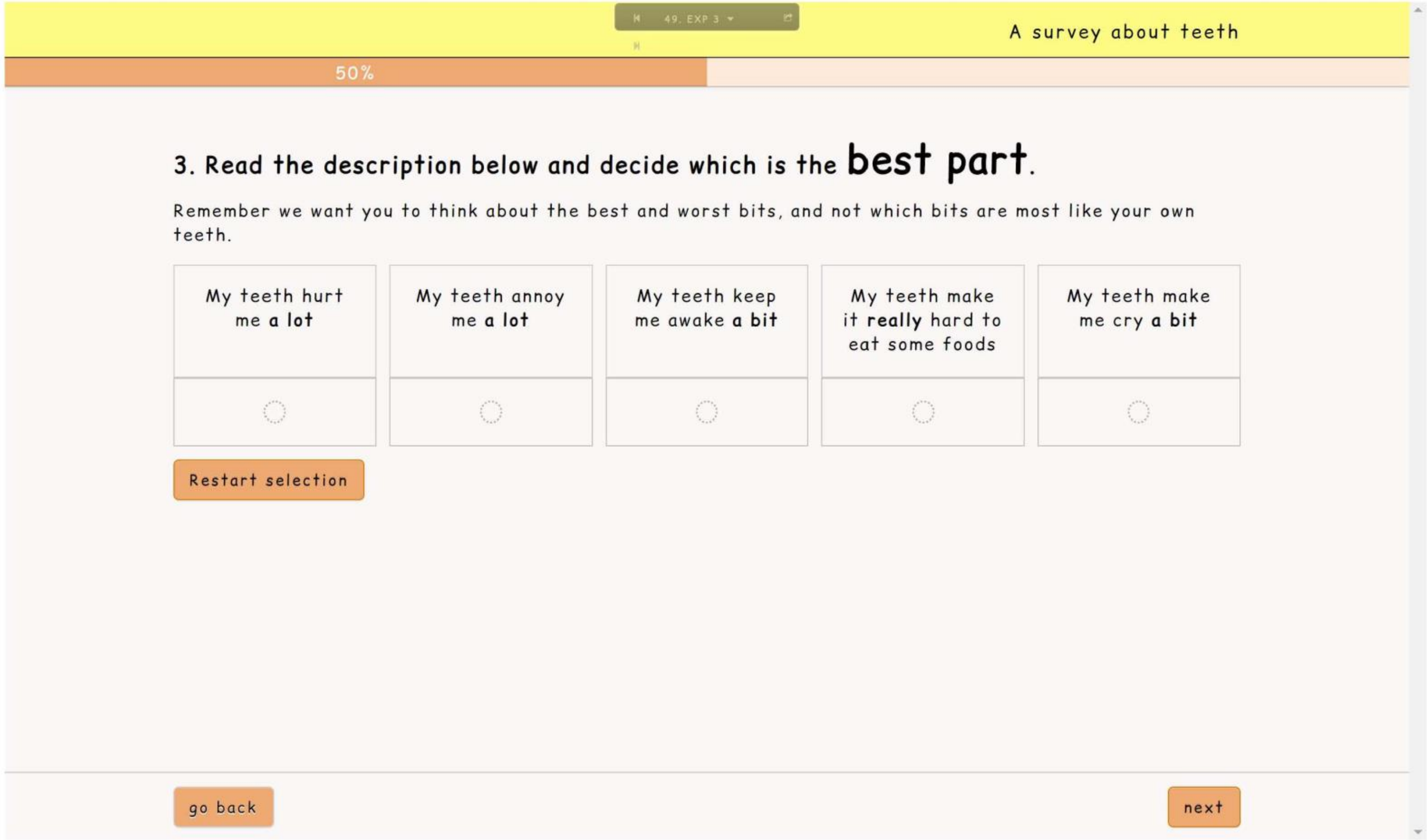
My teeth hurt me **a lot**	My teeth annoy me **a bit**	My teeth do **not** keep me awake **at all**	My teeth make it **a bit** hard to eat some foods	My teeth make me cry **a lot**
○	○	○	○	○



2. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

My teeth do **not** hurt me **at all**	My teeth annoy me **a bit**	My teeth do **not** keep me awake **at all**	My teeth make it **a bit** hard to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○



3. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

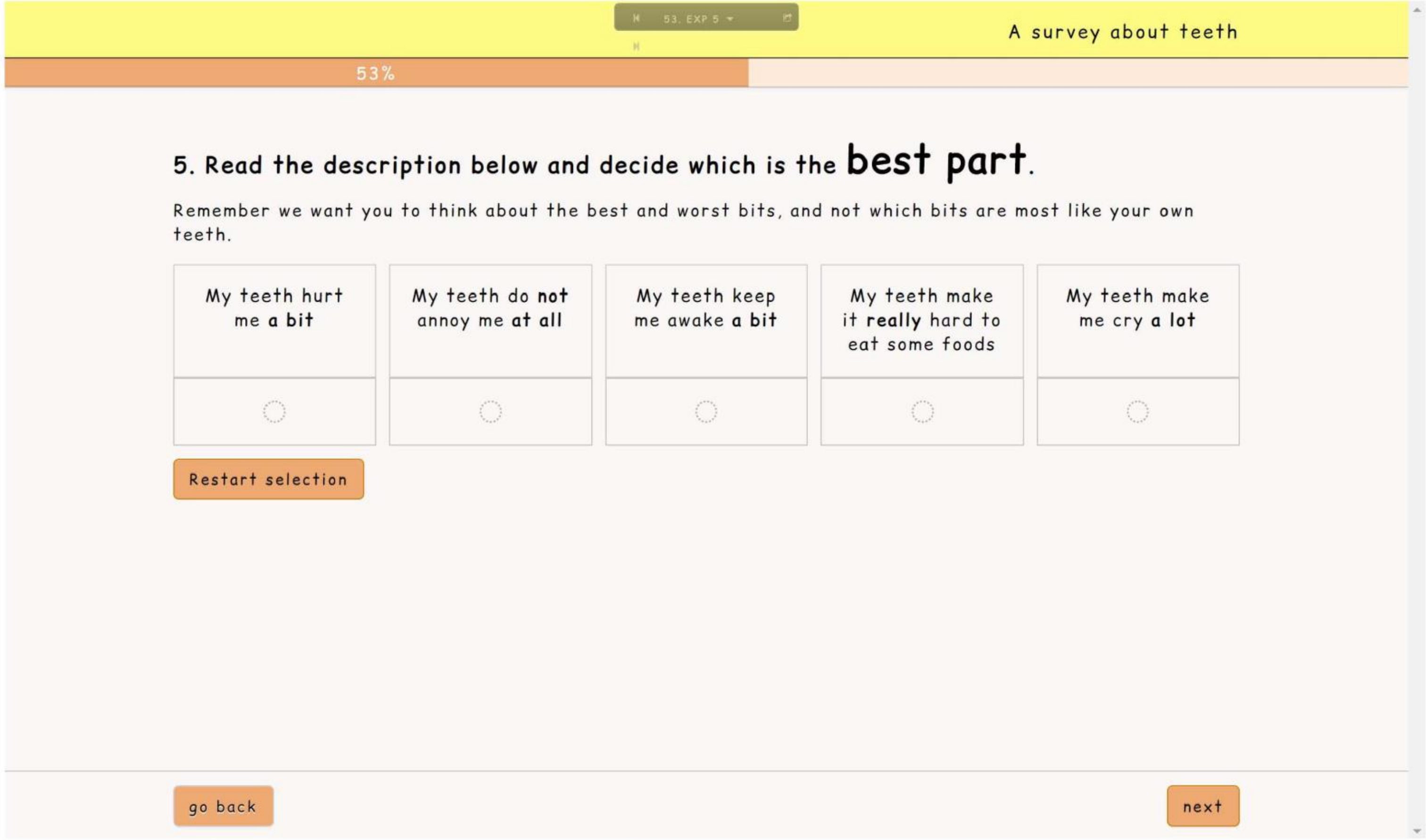
My teeth hurt me **a lot**	My teeth annoy me **a lot**	My teeth keep me awake **a bit**	My teeth make it **really** hard to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○

A survey about teeth

4. Read the description below and decide which is the best part.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

My teeth hurt me **a bit**	My teeth annoy me **a bit**	My teeth do **not** keep me awake **at all**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○

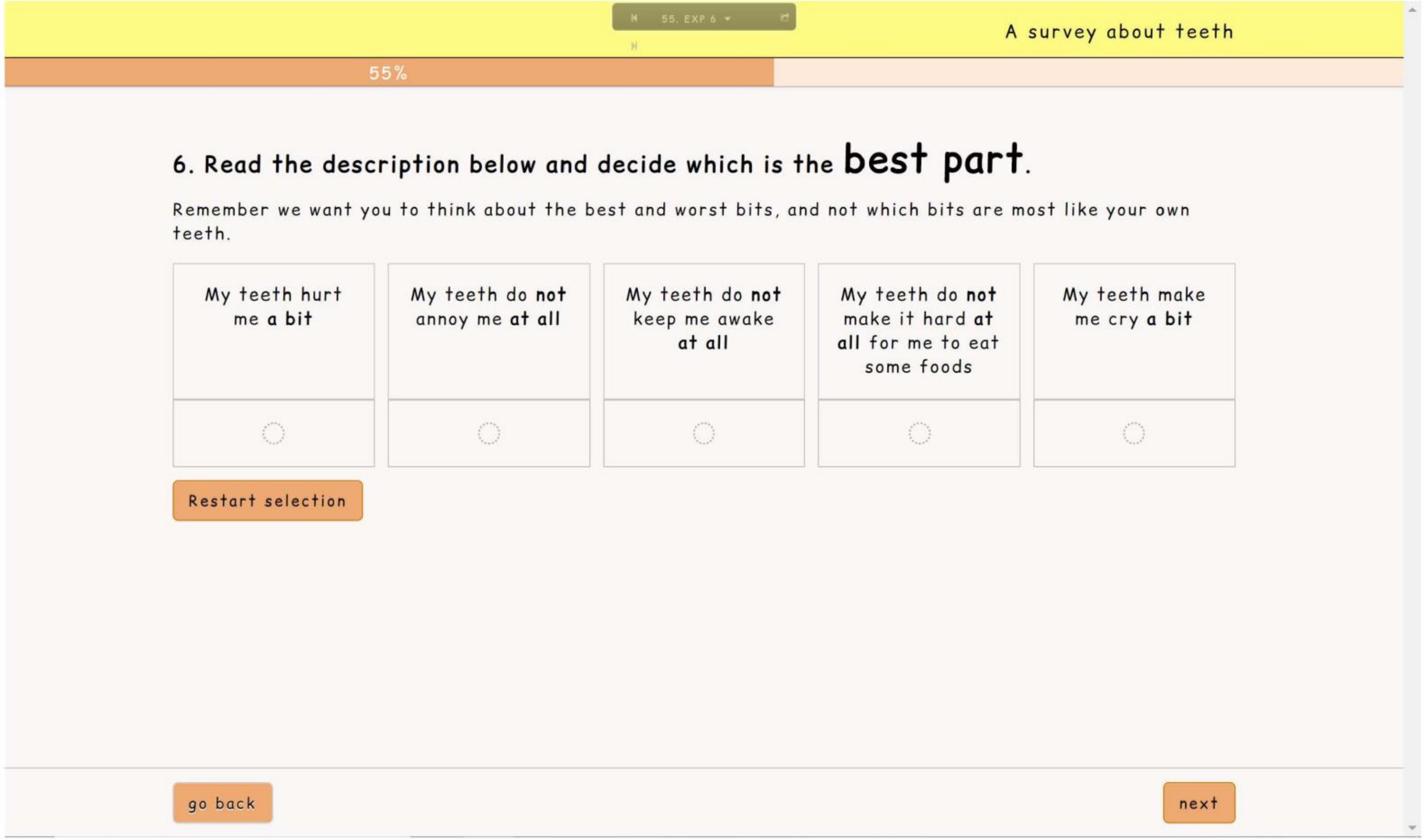


A survey about teeth

5. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

My teeth hurt me **a bit**	My teeth do **not** annoy me **at all**	My teeth keep me awake **a bit**	My teeth make it **really** hard to eat some foods	My teeth make me cry **a lot**
○	○	○	○	○



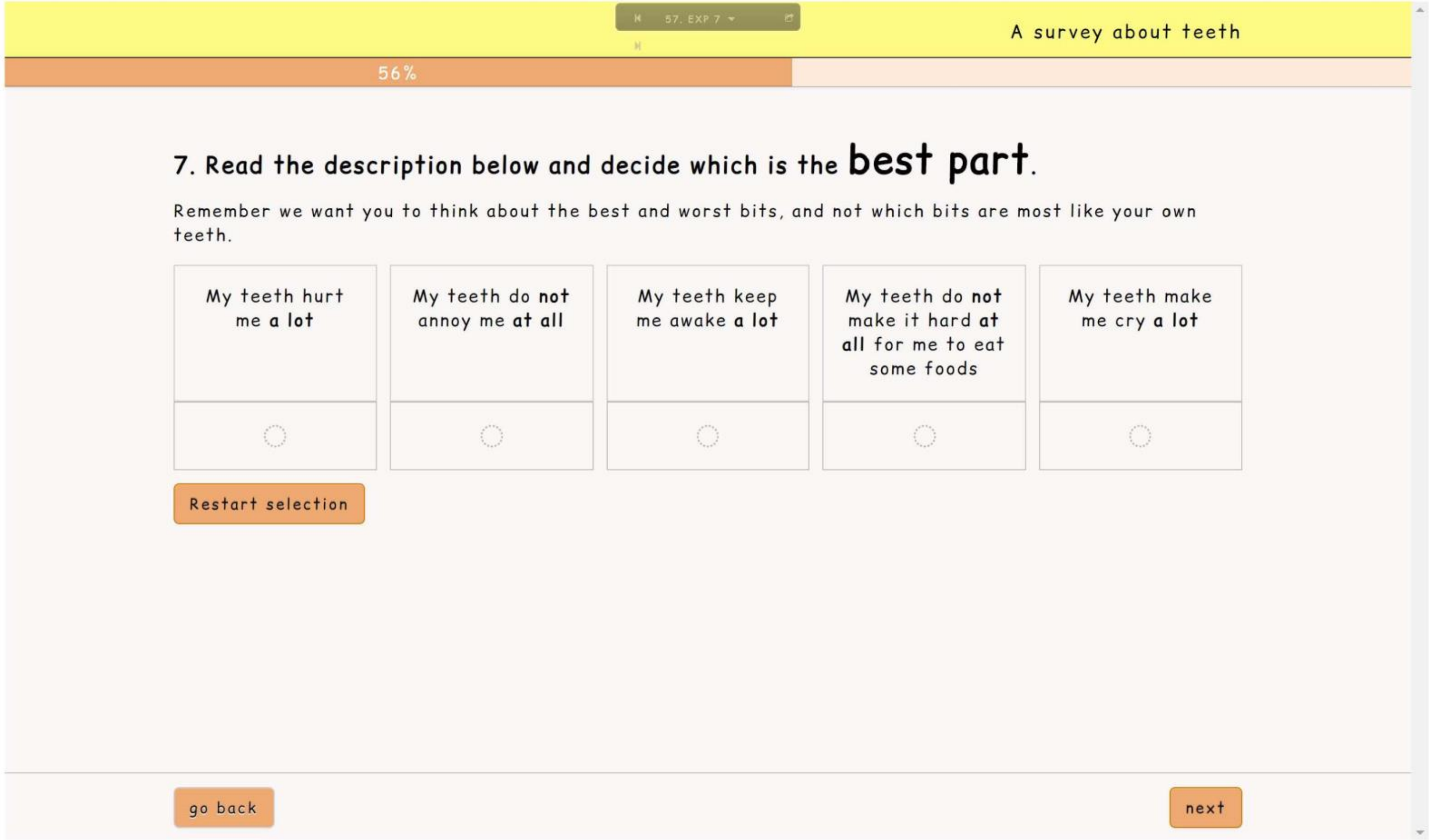
A survey about teeth

55%

6. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

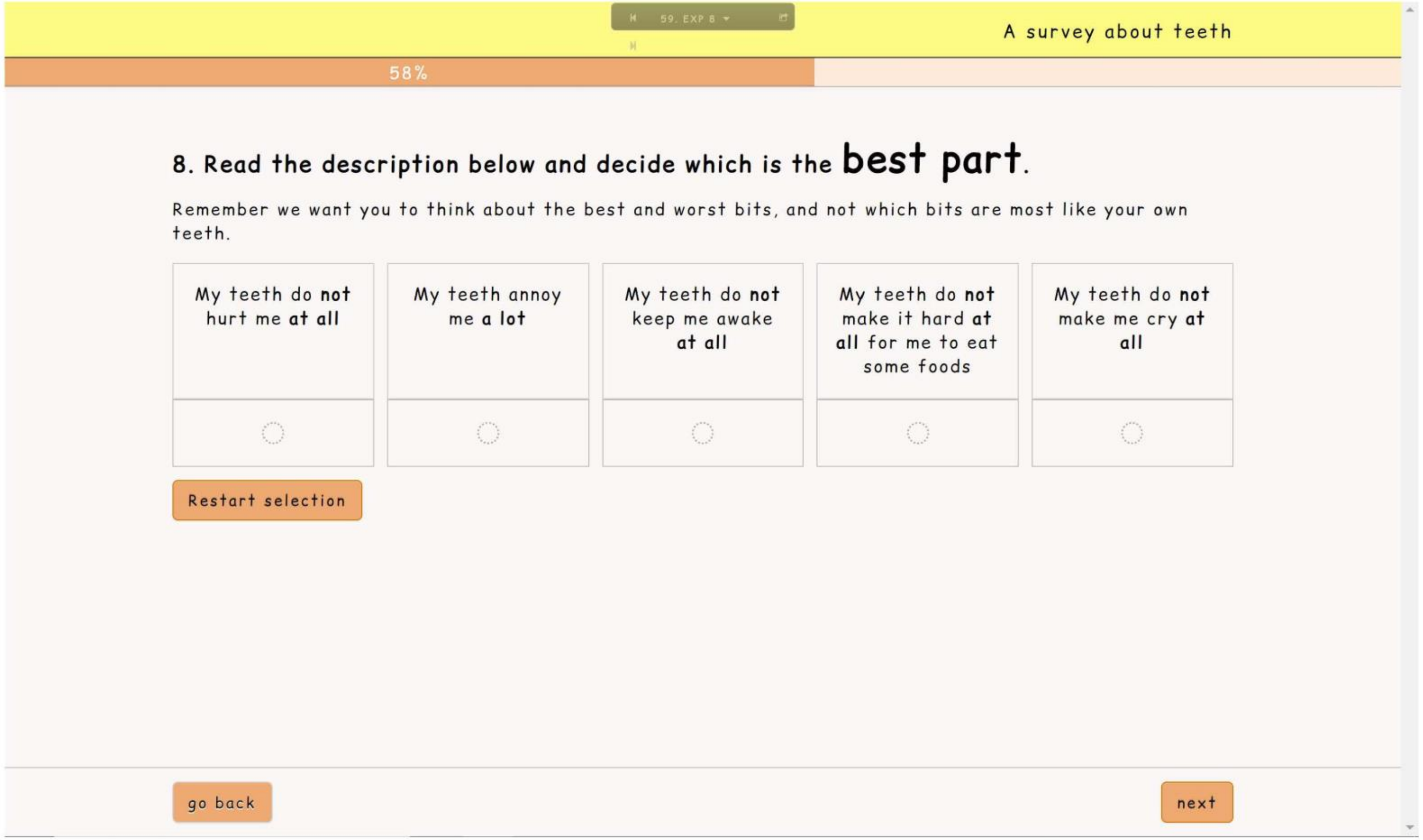
My teeth hurt me **a bit**	My teeth do **not** annoy me **at all**	My teeth do **not** keep me awake **at all**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make me cry **a bit**
○	○	○	○	○



7. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

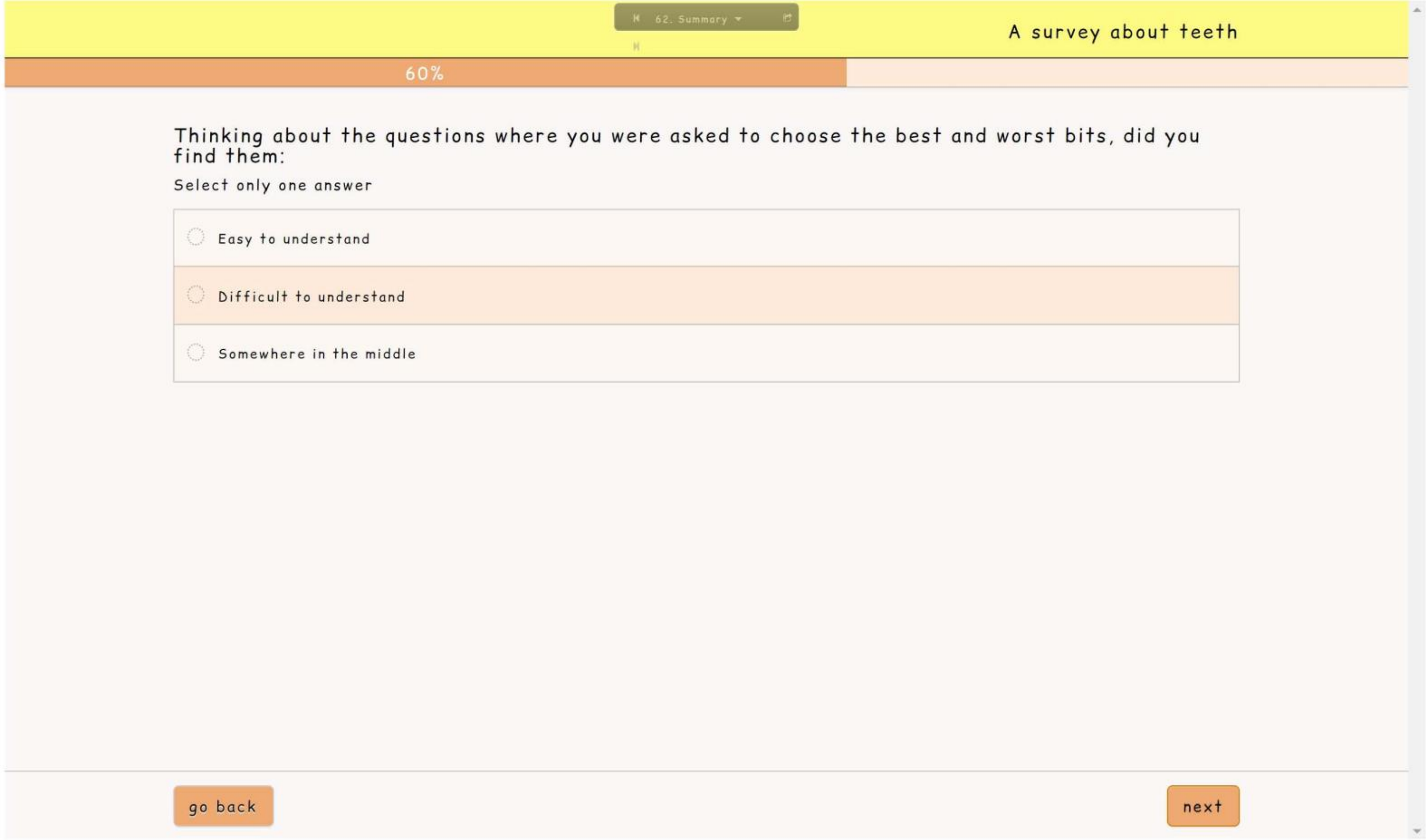
My teeth hurt me **a lot**	My teeth do **not** annoy me **at all**	My teeth keep me awake **a lot**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make me cry **a lot**
○	○	○	○	○



8. Read the description below and decide which is the **best part**.

Remember we want you to think about the best and worst bits, and not which bits are most like your own teeth.

My teeth do **not** hurt me **at all**	My teeth annoy me **a lot**	My teeth do **not** keep me awake **at all**	My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make me cry **at all**
○	○	○	○	○



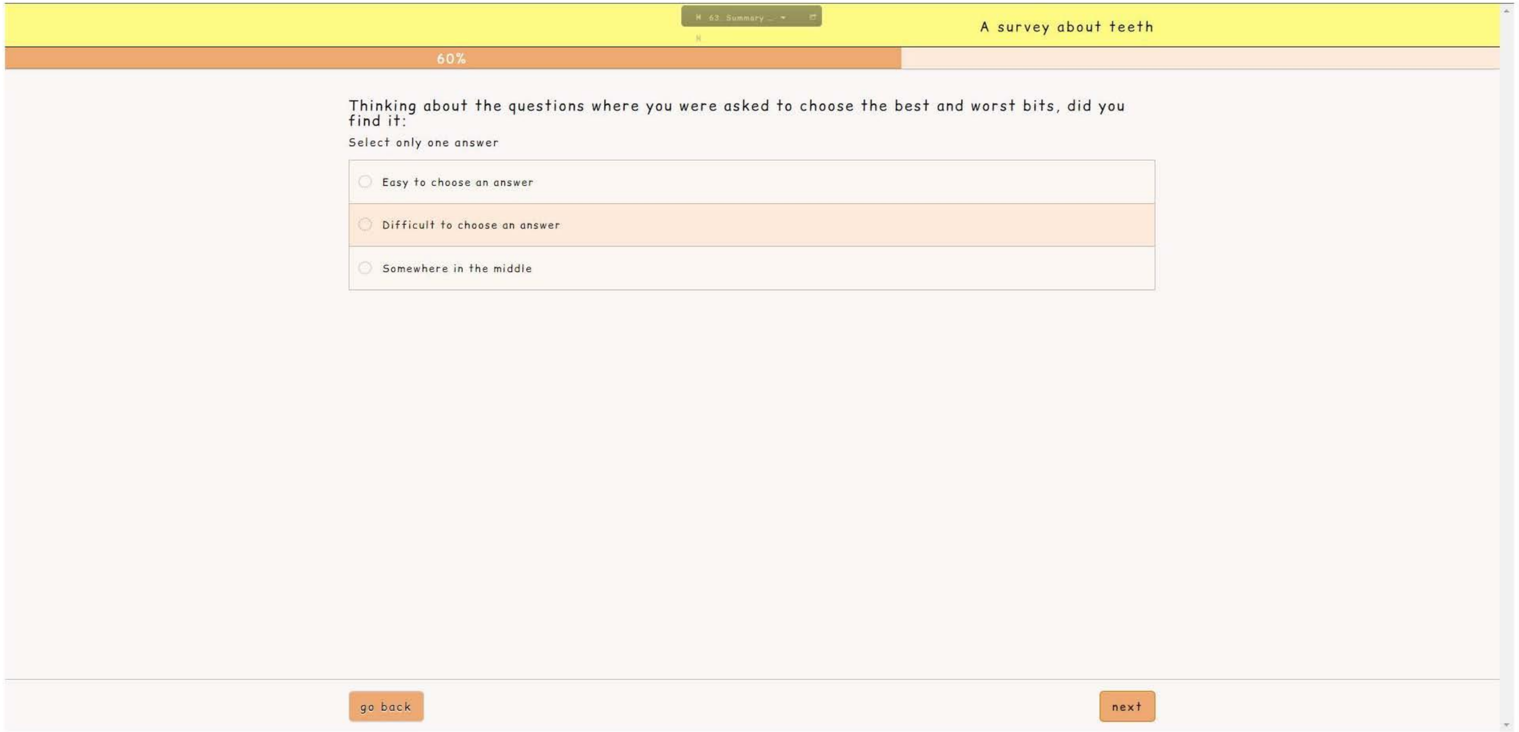
Thinking about the questions where you were asked to choose the best and worst bits, did you find them:

Select only one answer

○ Easy to understand

○ Difficult to understand

○ Somewhere in the middle



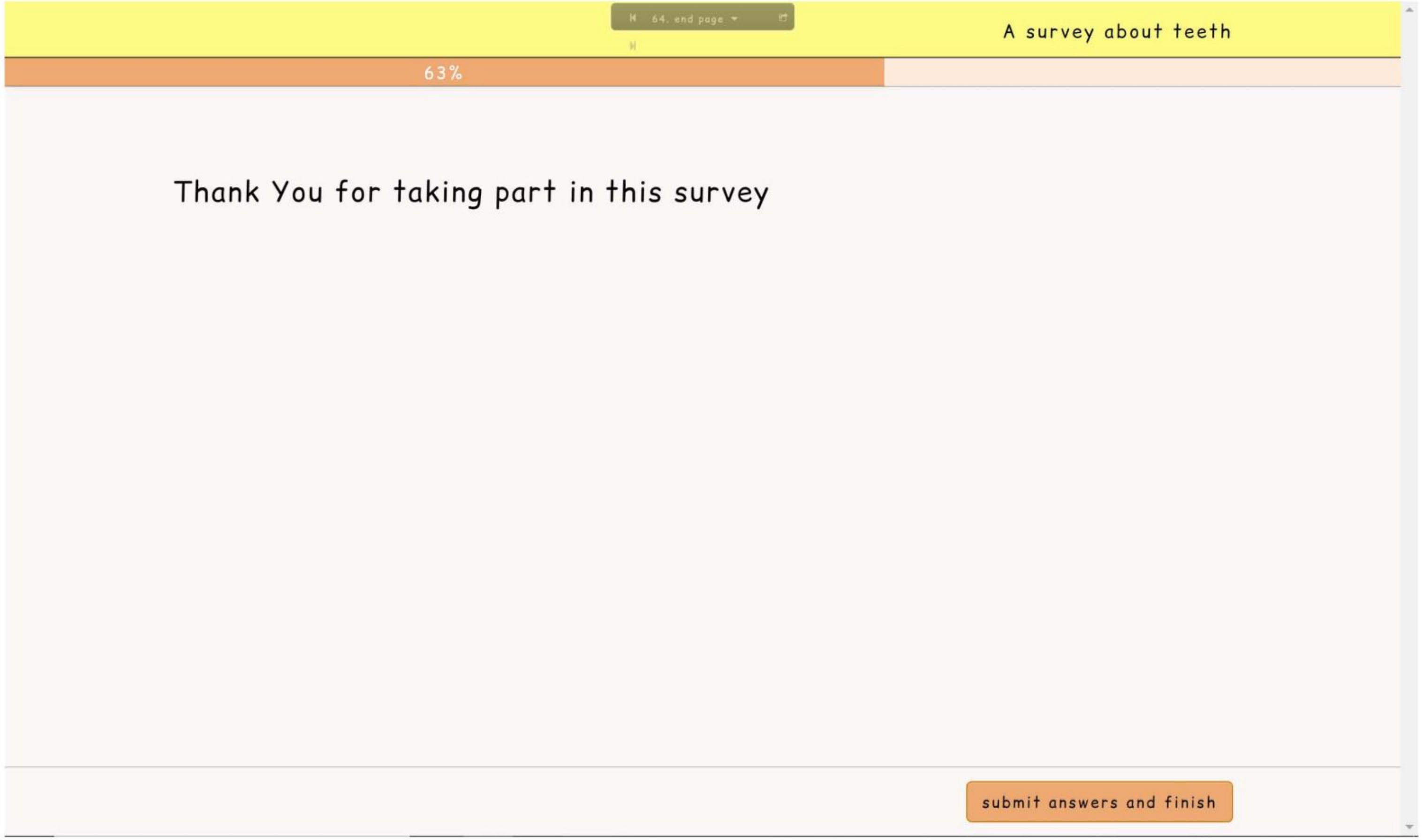
Thinking about the questions where you were asked to choose the best and worst bits, did you find it:

Select only one answer

○ Easy to choose an answer

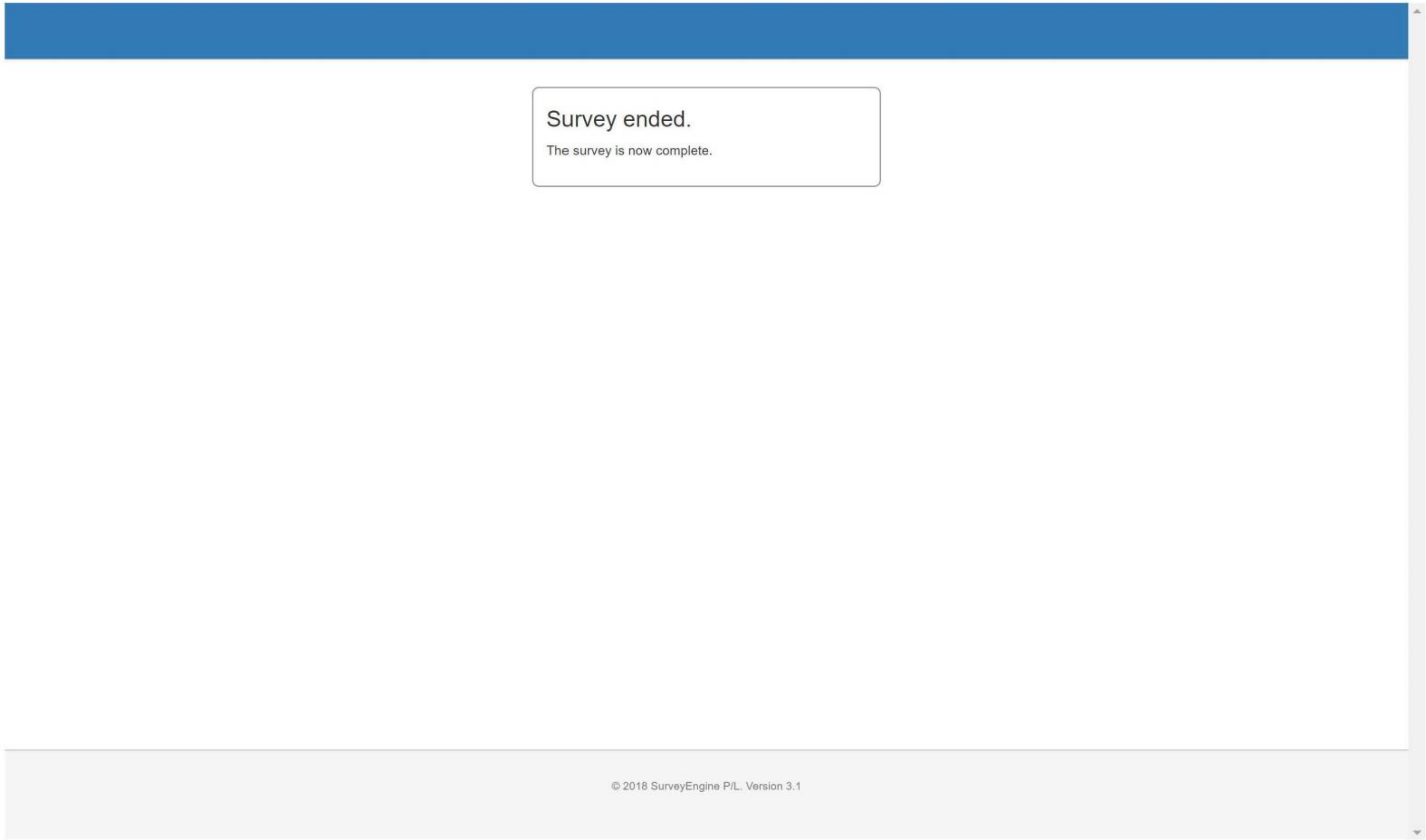
○ Difficult to choose an answer

○ Somewhere in the middle



A survey about teeth

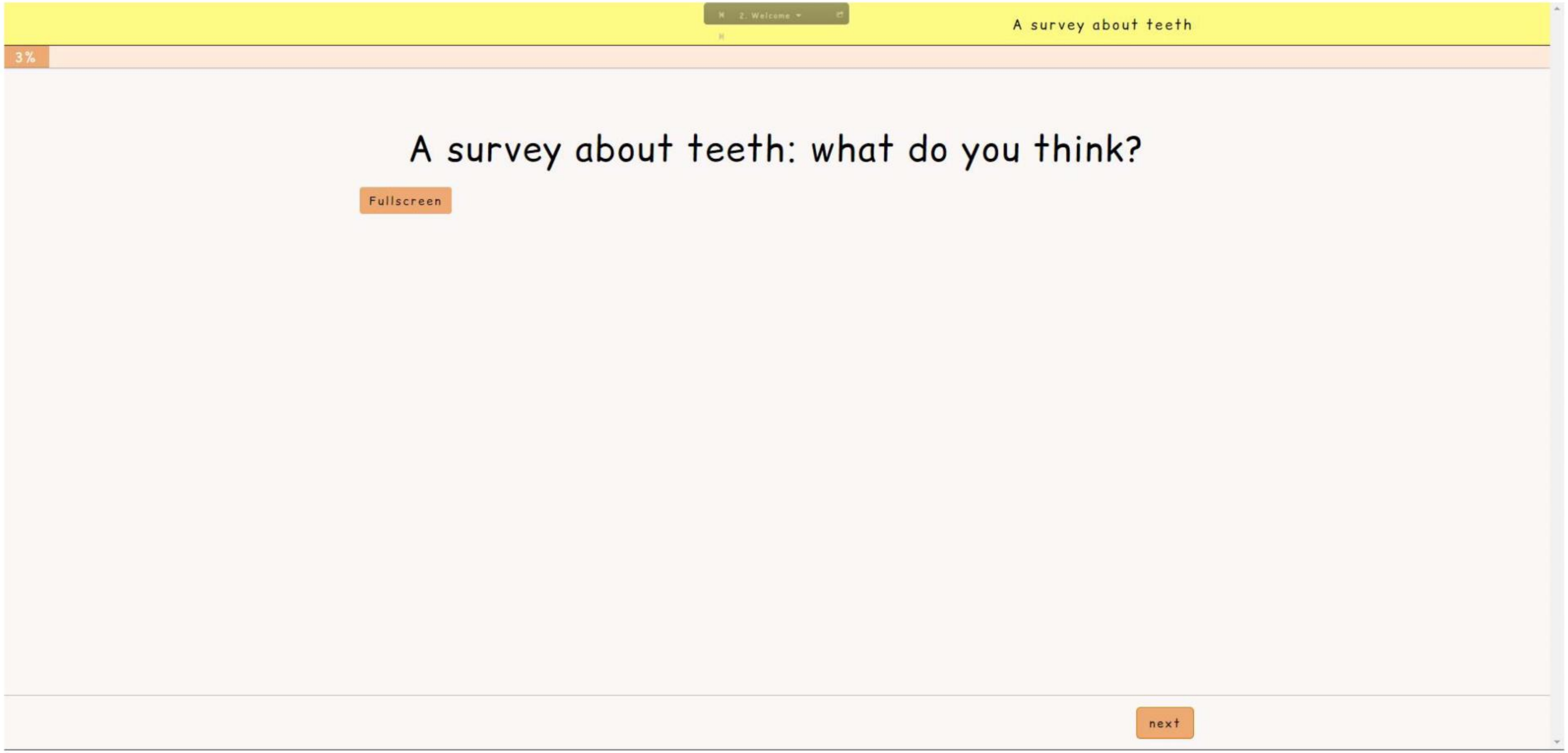
Thank You for taking part in this survey



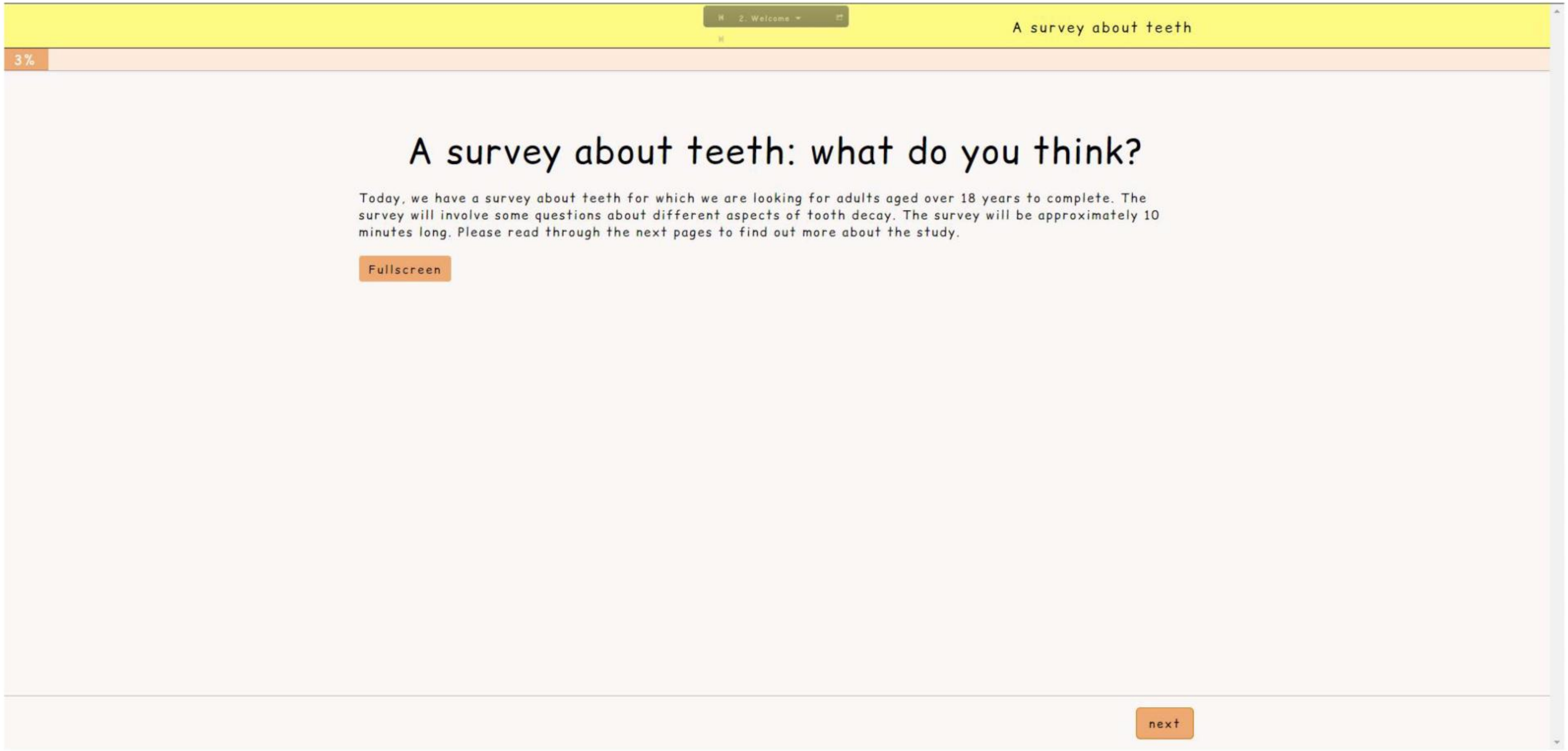
Survey ended.

The survey is now complete.

U. DCE$_{TTO}$ valuation survey for adults

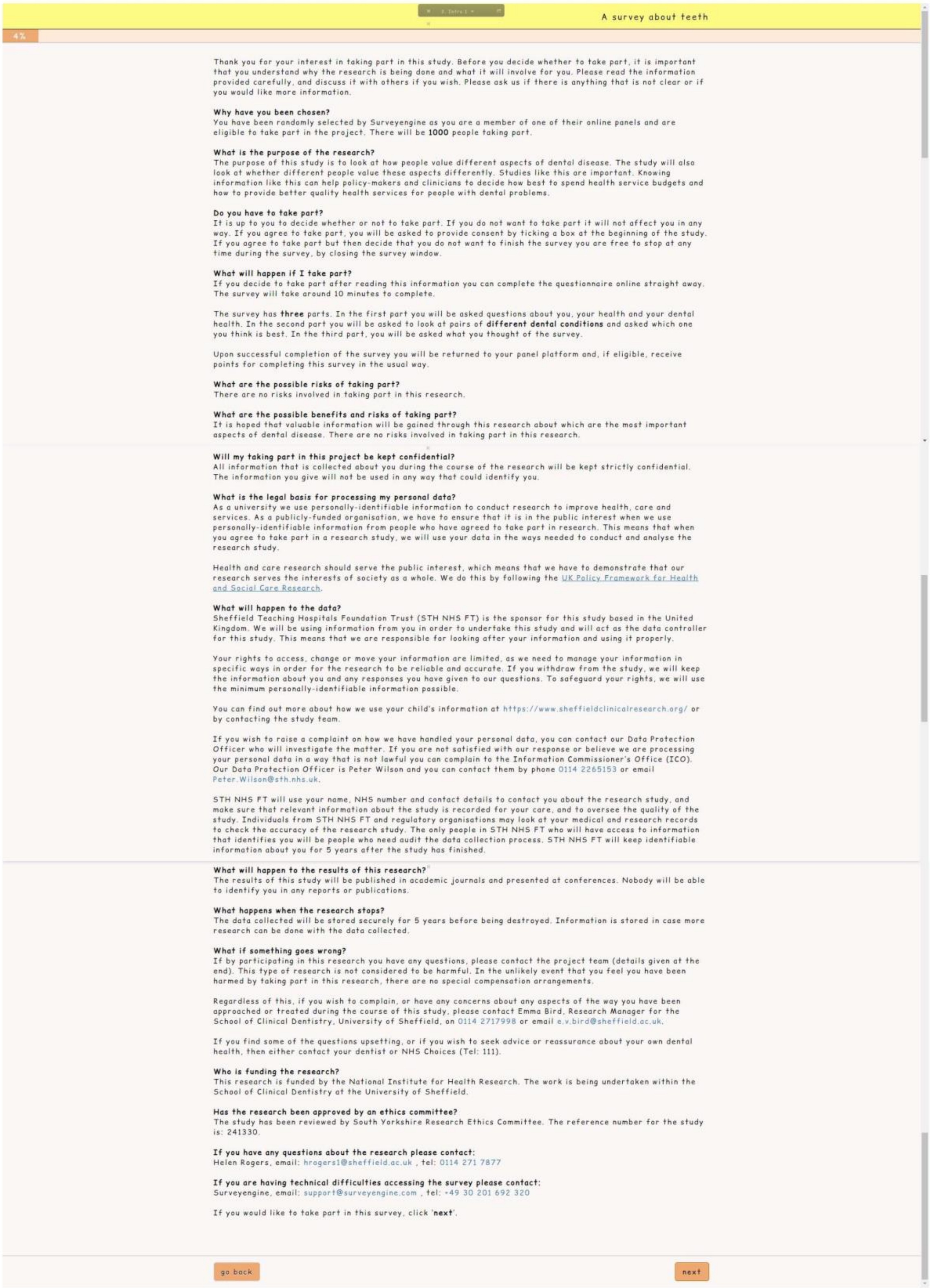


A survey about teeth: what do you think?



A survey about teeth: what do you think?

Today, we have a survey about teeth for which we are looking for adults aged over 18 years to complete. The survey will involve some questions about different aspects of tooth decay. The survey will be approximately 10 minutes long. Please read through the next pages to find out more about the study.



Thank you for your interest in taking part in this study. Before you decide whether to take part, it is important that you understand why the research is being done and what it will involve for you. Please read the information provided carefully, and discuss it with others if you wish. Please ask us if there is anything that is not clear or if you would like more information.

Why have you been chosen?
You have been randomly selected by Surveyengine as you are a member of one of their online panels and are eligible to take part in the project. There will be **1000** people taking part.

What is the purpose of the research?
The purpose of this study is to look at how people value different aspects of dental disease. The study will also look at whether different people value these aspects differently. Studies like this are important. Knowing information like this can help policy-makers and clinicians to decide how best to spend health service budgets and how to provide better quality health services for people with dental problems.

Do you have to take part?
It is up to you to decide whether or not to take part. If you do not want to take part it will not affect you in any way. If you agree to take part, you will be asked to provide consent by ticking a box at the beginning of the study. If you agree to take part but then decide that you do not want to finish the survey you are free to stop at any time during the survey, by closing the survey window.

What will happen if I take part?
If you decide to take part after reading this information you can complete the questionnaire online straight away. The survey will take around 10 minutes to complete.

The survey has **three** parts. In the first part you will be asked questions about you, your health and your dental health. In the second part you will be asked to look at pairs of **different dental conditions** and asked which one you think is best. In the third part, you will be asked what you thought of the survey.

Upon successful completion of the survey you will be returned to your panel platform and, if eligible, receive points for completing this survey in the usual way.

What are the possible risks of taking part?
There are no risks involved in taking part in this research.

What are the possible benefits and risks of taking part?
It is hoped that valuable information will be gained through this research about which are the most important aspects of dental disease. There are no risks involved in taking part in this research.

Will my taking part in this project be kept confidential?
All information that is collected about you during the course of the research will be kept strictly confidential. The information you give will not be used in any way that could identify you.

What is the legal basis for processing my personal data?
As a university we use personally-identifiable information to conduct research to improve health, care and services. As a publicly-funded organisation, we have to ensure that it is in the public interest when we use personally-identifiable information from people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use your data in the ways needed to conduct and analyse the research study.

Health and care research should serve the public interest, which means that we have to demonstrate that our research serves the interests of society as a whole. We do this by following the UK Policy Framework for Health and Social Care Research.

What will happen to the data?
Sheffield Teaching Hospitals Foundation Trust (STH NHS FT) is the sponsor for this study based in the United Kingdom. We will be using information from you in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly.

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you and any responses you have given to our questions. To safeguard your rights, we will use the minimum personally-identifiable information possible.

You can find out more about how we use your child's information at https://www.sheffieldclinicalresearch.org/ or by contacting the study team.

If you wish to raise a complaint on how we have handled your personal data, you can contact our Data Protection Officer who will investigate the matter. If you are not satisfied with our response or believe we are processing your personal data in a way that is not lawful you can complain to the Information Commissioner's Office (ICO). Our Data Protection Officer is Peter Wilson and you can contact them by phone 0114 2265153 or email Peter.Wilson@sth.nhs.uk.

STH NHS FT will use your name, NHS number and contact details to contact you about the research study, and make sure that relevant information about the study is recorded for your care, and to oversee the quality of the study. Individuals from STH NHS FT and regulatory organisations may look at your medical and research records to check the accuracy of the research study. The only people in STH NHS FT who will have access to information that identifies you will be people who need audit the data collection process. STH NHS FT will keep identifiable information about you for 5 years after the study has finished.

What will happen to the results of this research?
The results of this study will be published in academic journals and presented at conferences. Nobody will be able to identify you in any reports or publications.

What happens when the research stops?
The data collected will be stored securely for 5 years before being destroyed. Information is stored in case more research can be done with the data collected.

What if something goes wrong?
If by participating in this research you have any questions, please contact the project team (details given at the end). This type of research is not considered to be harmful. In the unlikely event that you feel you have been harmed by taking part in this research, there are no special compensation arrangements.

Regardless of this, if you wish to complain, or have any concerns about any aspects of the way you have been approached or treated during the course of this study, please contact Emma Bird, Research Manager for the School of Clinical Dentistry, University of Sheffield, on 0114 2717998 or email e.v.bird@sheffield.ac.uk.

If you find some of the questions upsetting, or if you wish to seek advice or reassurance about your own dental health, then either contact your dentist or NHS Choices (Tel: 111).

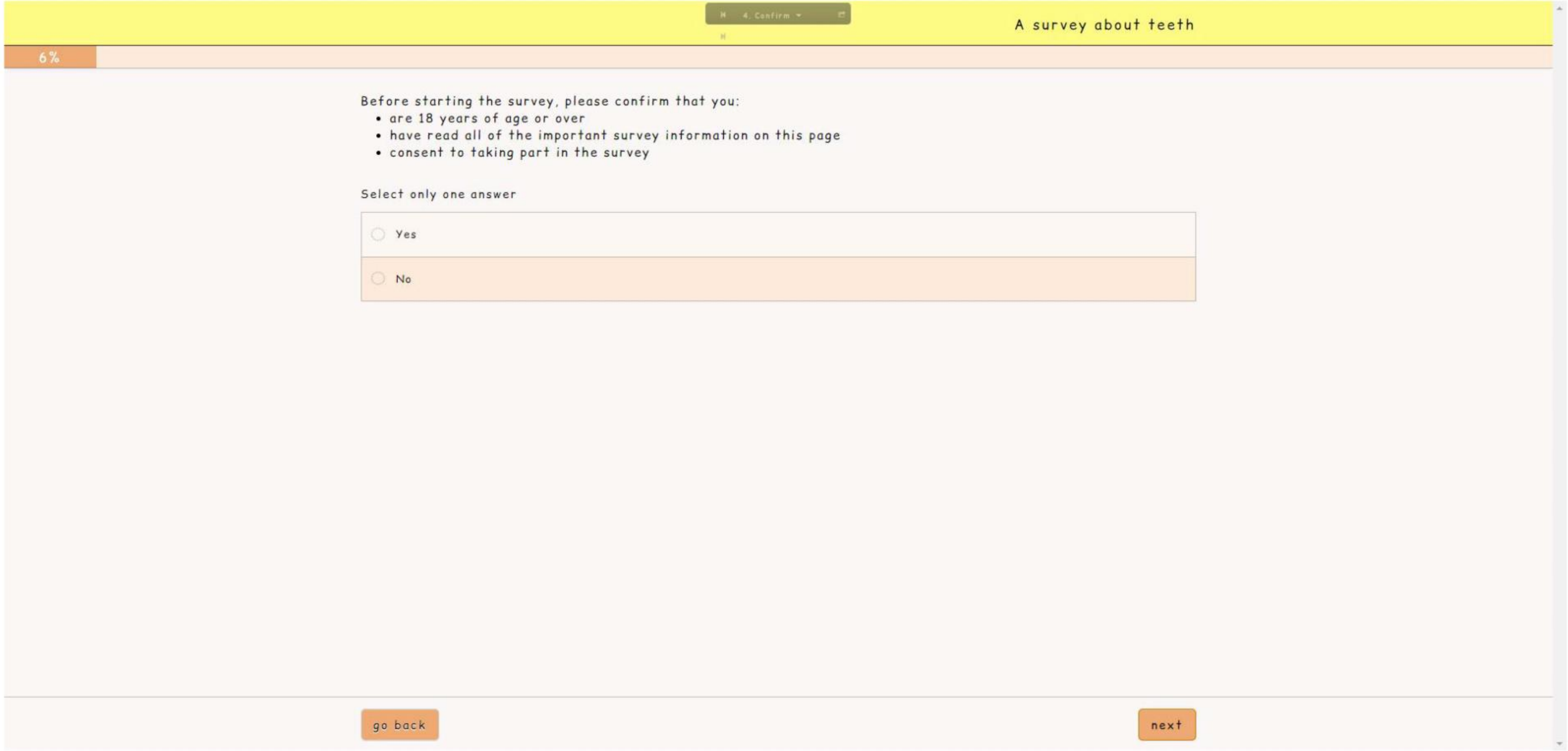
Who is funding the research?
This research is funded by the National Institute for Health Research. The work is being undertaken within the School of Clinical Dentistry at the University of Sheffield.

Has the research been approved by an ethics committee?
The study has been reviewed by South Yorkshire Research Ethics Committee. The reference number for the study is: 241330.

If you have any questions about the research please contact:
Helen Rogers, email: hrogers1@sheffield.ac.uk , tel: 0114 271 7877

If you are having technical difficulties accessing the survey please contact:
Surveyengine, email: support@surveyengine.com , tel: +49 30 201 692 320

If you would like to take part in this survey, click '**next**'.



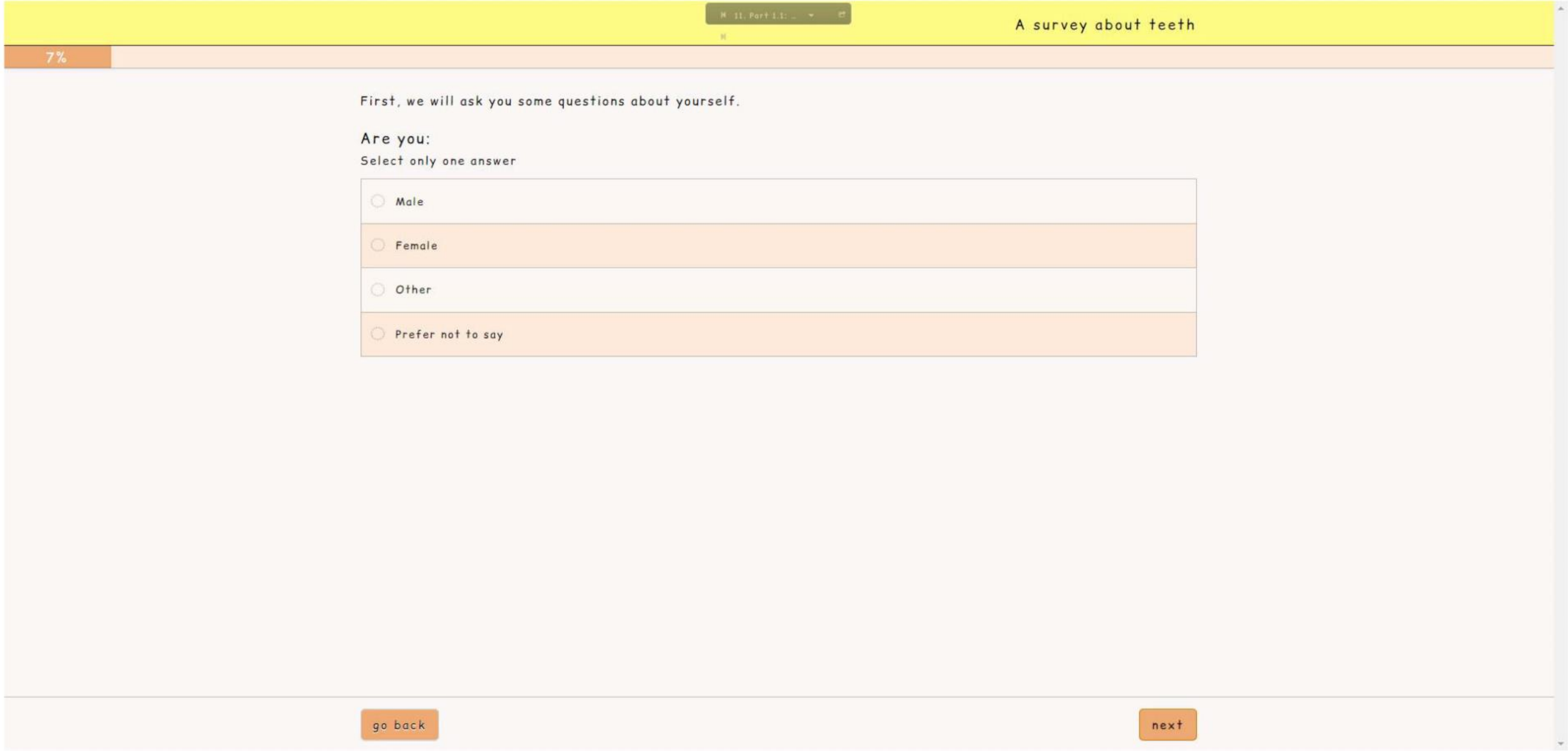
Before starting the survey, please confirm that you:

- are 18 years of age or over
- have read all of the important survey information on this page
- consent to taking part in the survey

Select only one answer

○ Yes

○ No



First, we will ask you some questions about yourself.

Are you:

Select only one answer

○ Male

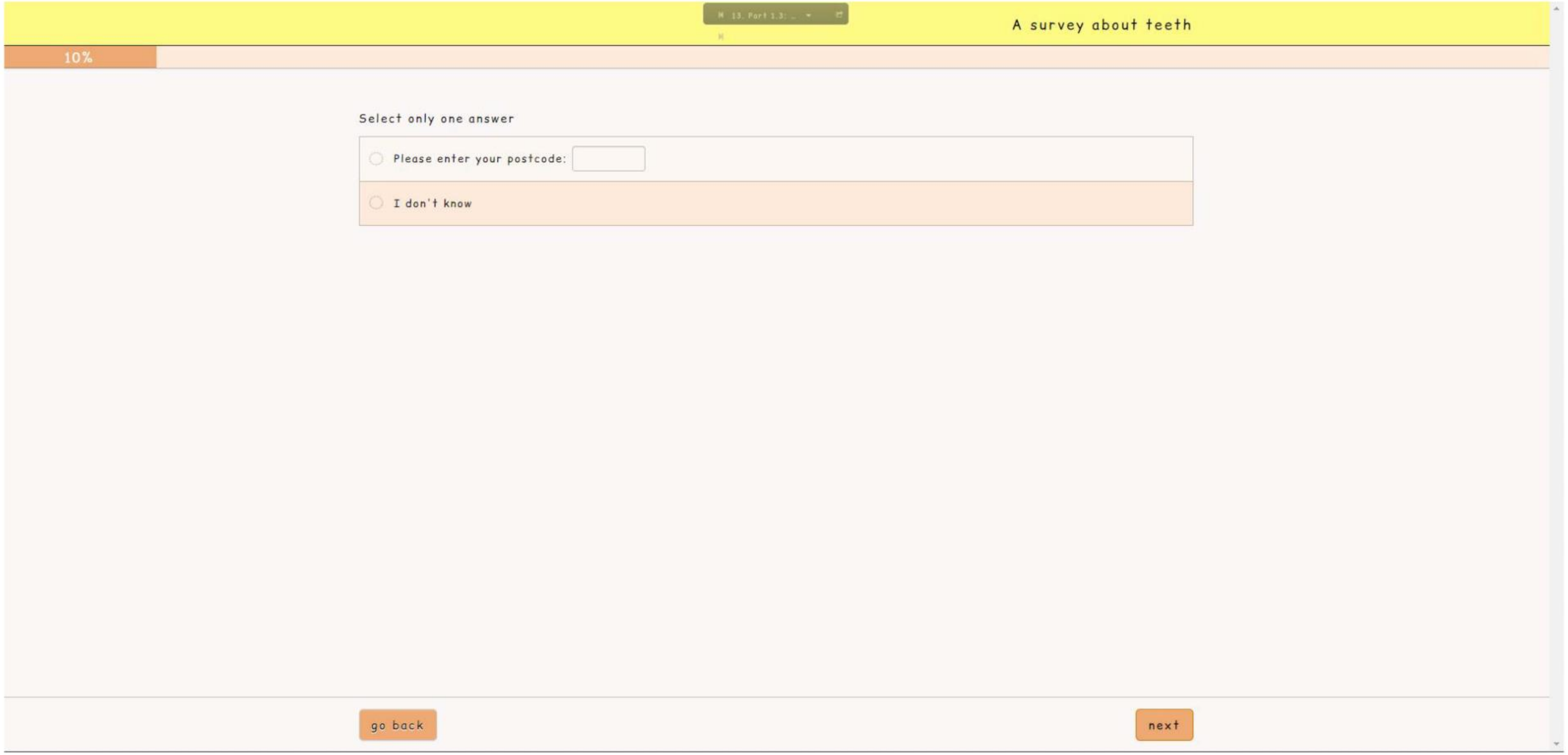
○ Female

○ Other

○ Prefer not to say

How old are you (in years)?

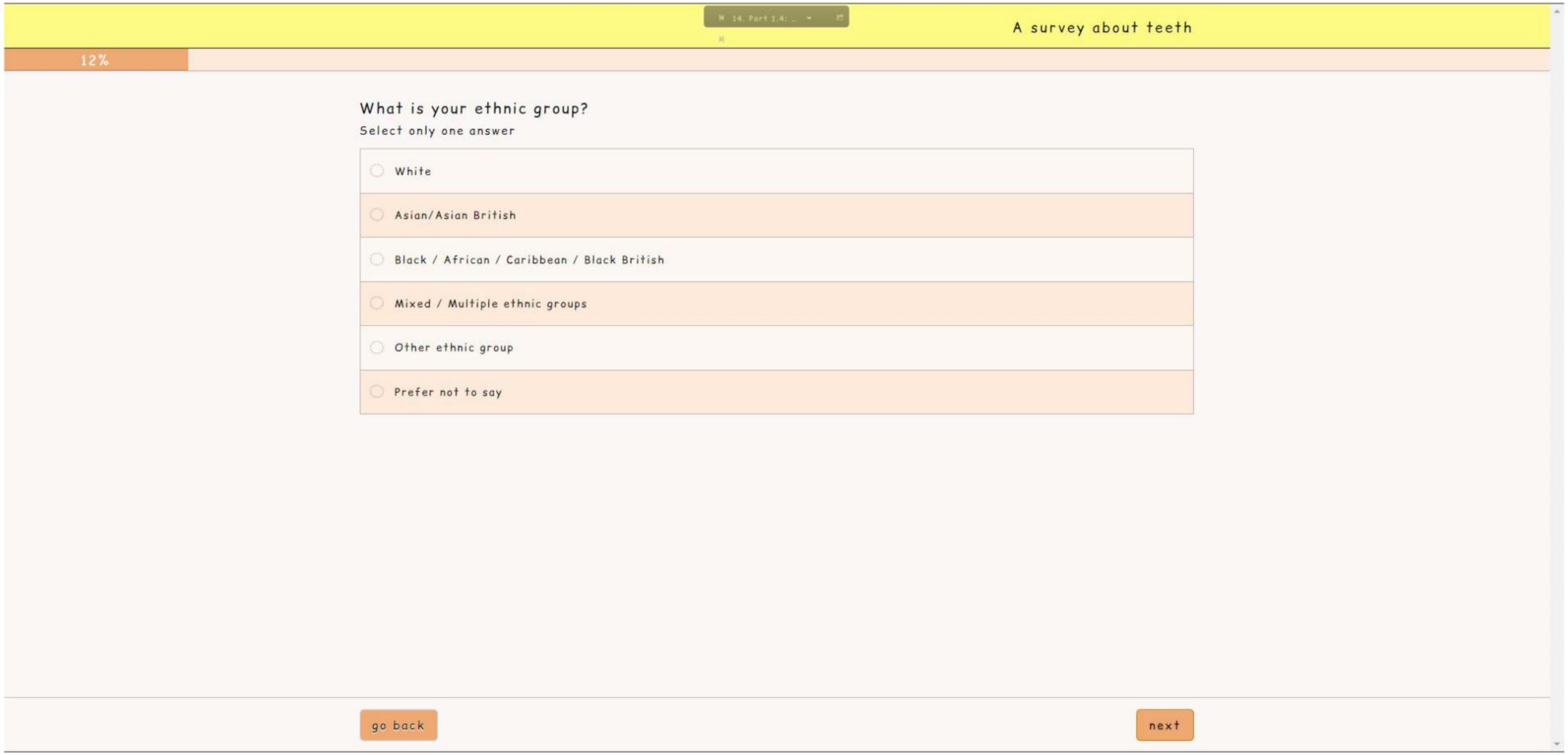
Enter text below



Select only one answer

○ Please enter your postcode:

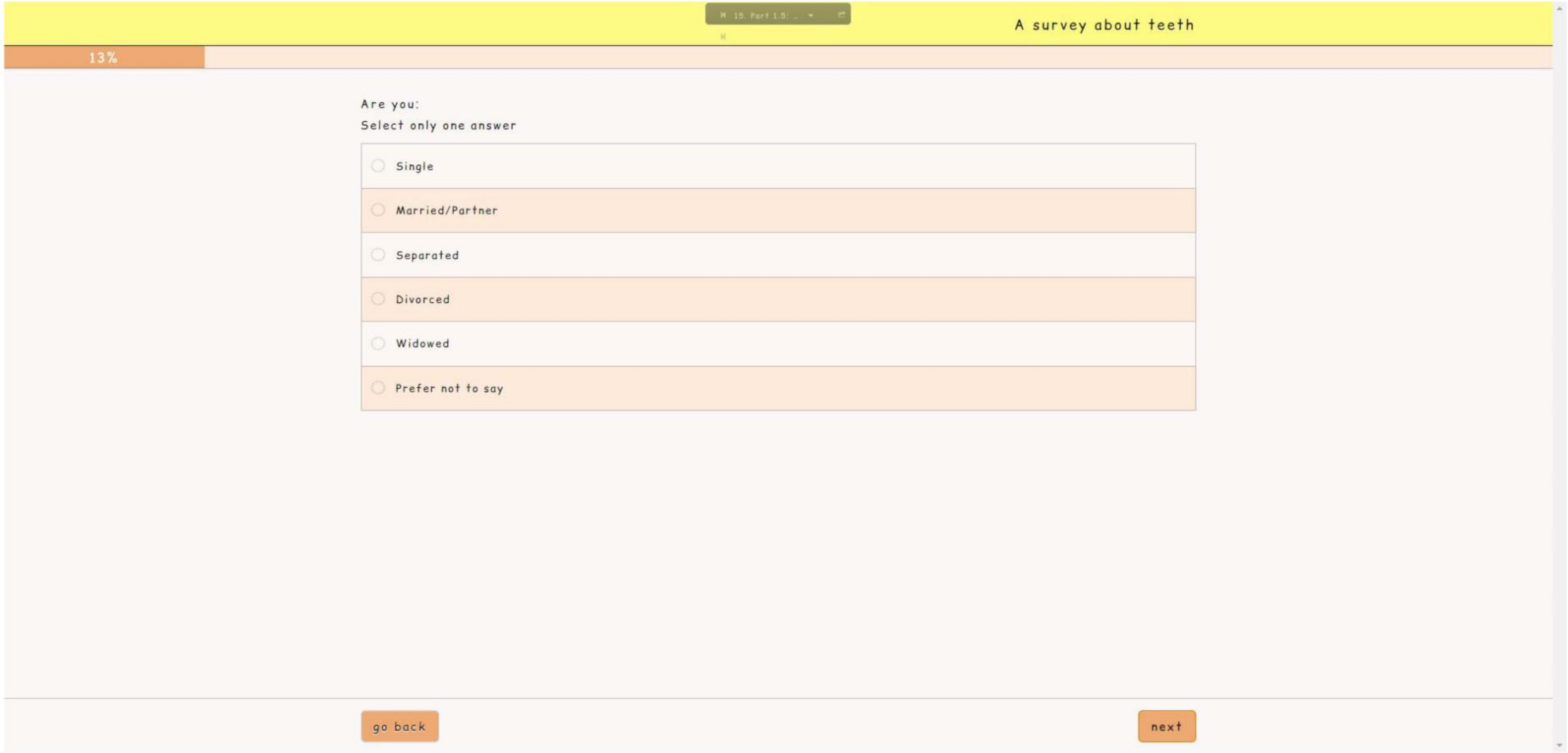
○ I don't know



What is your ethnic group?

Select only one answer

- ○ White
- ○ Asian/Asian British
- ○ Black / African / Caribbean / Black British
- ○ Mixed / Multiple ethnic groups
- ○ Other ethnic group
- ○ Prefer not to say



Are you:

Select only one answer

○ Single

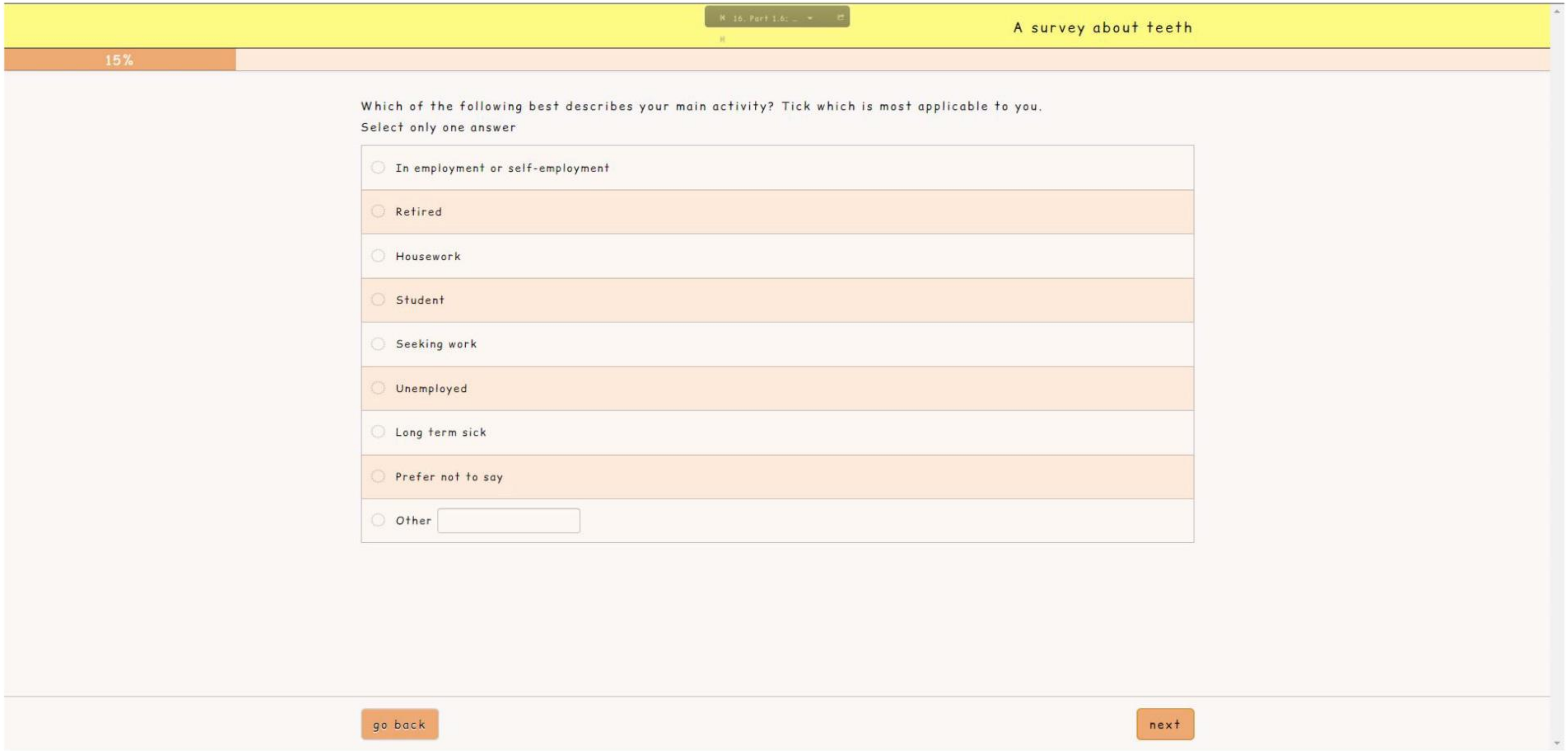
○ Married/Partner

○ Separated

○ Divorced

○ Widowed

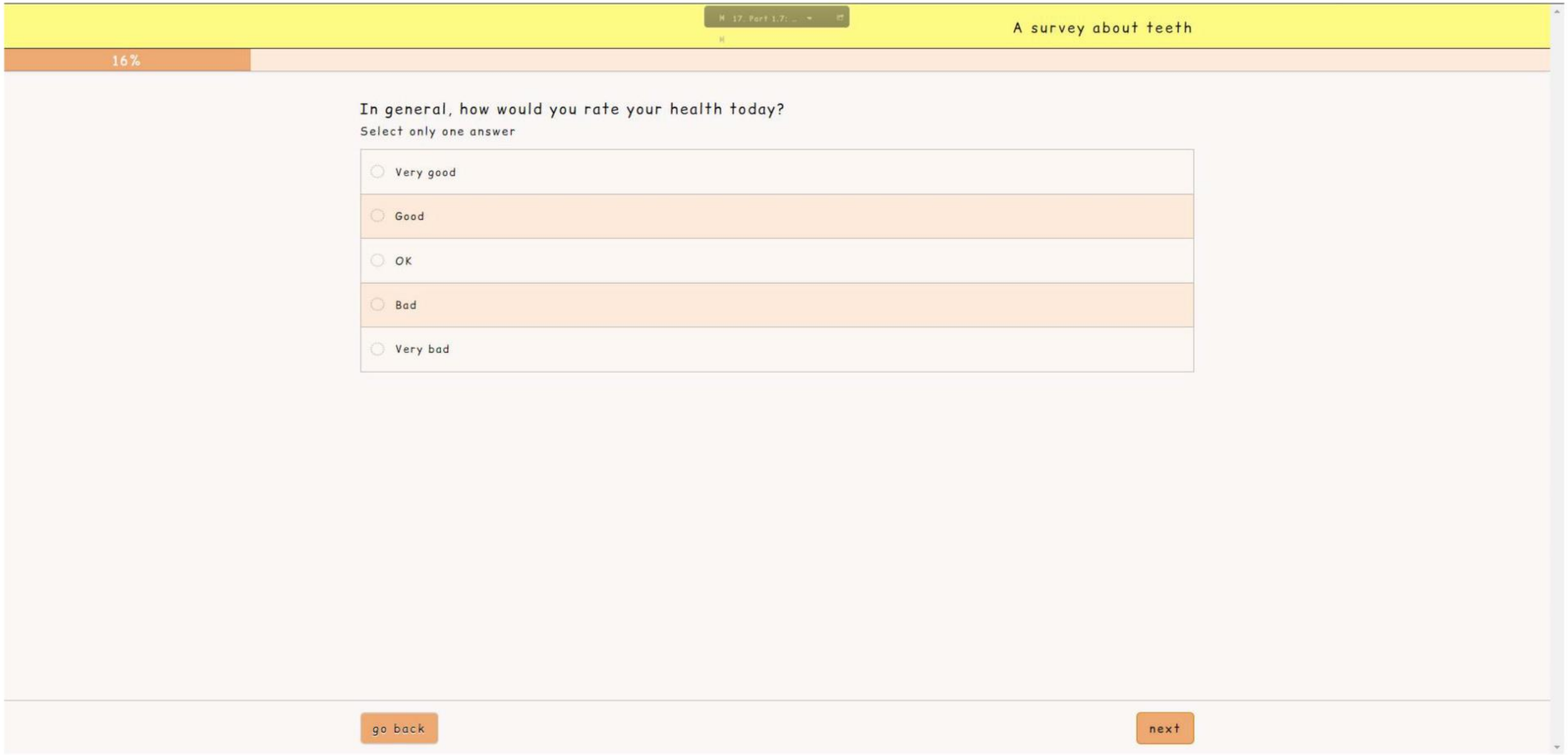
○ Prefer not to say



Which of the following best describes your main activity? Tick which is most applicable to you.

Select only one answer

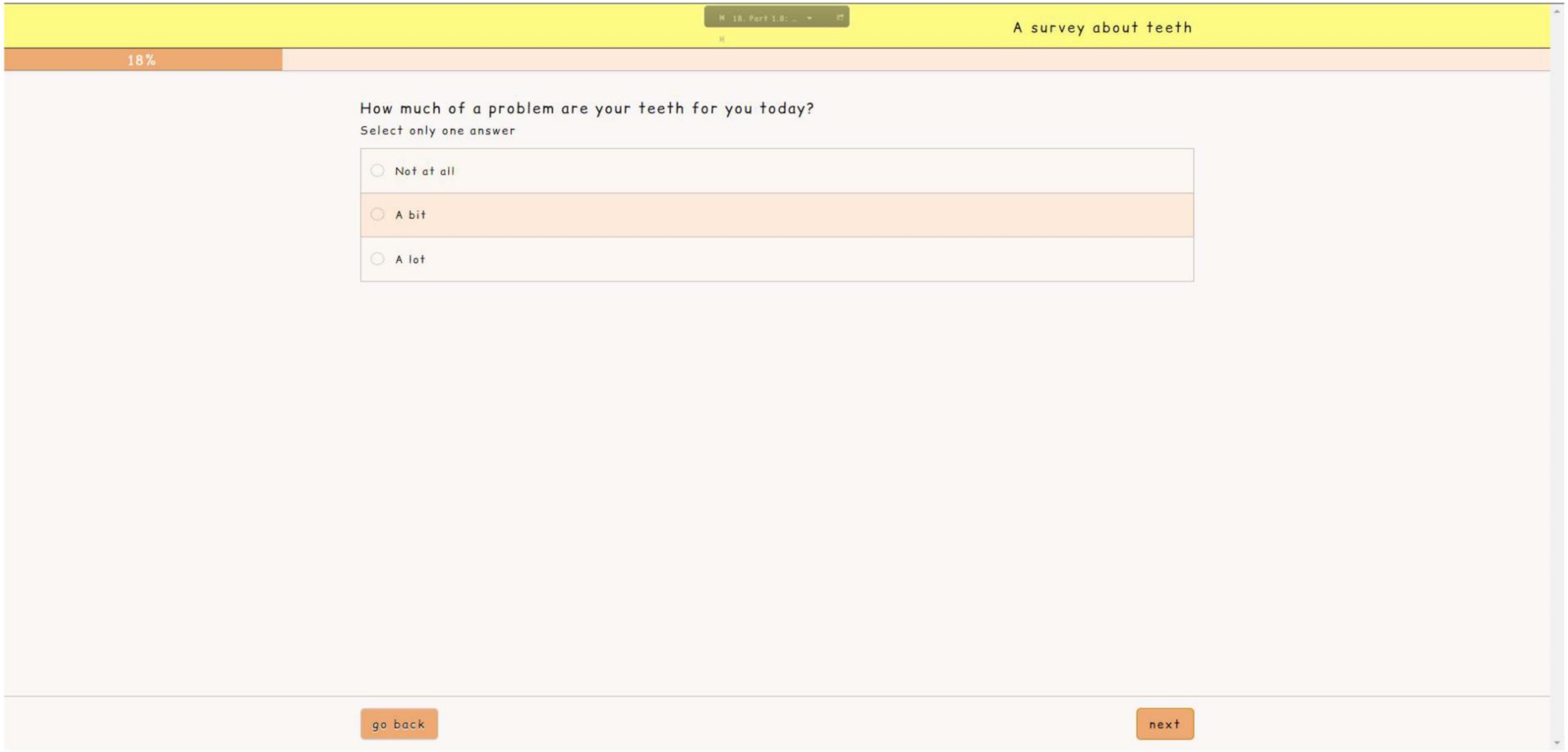
○ In employment or self-employment
○ Retired
○ Housework
○ Student
○ Seeking work
○ Unemployed
○ Long term sick
○ Prefer not to say
○ Other



In general, how would you rate your health today?

Select only one answer

- ○ Very good
- ○ Good
- ○ OK
- ○ Bad
- ○ Very bad



How much of a problem are your teeth for you today?

Select only one answer

○ Not at all

○ A bit

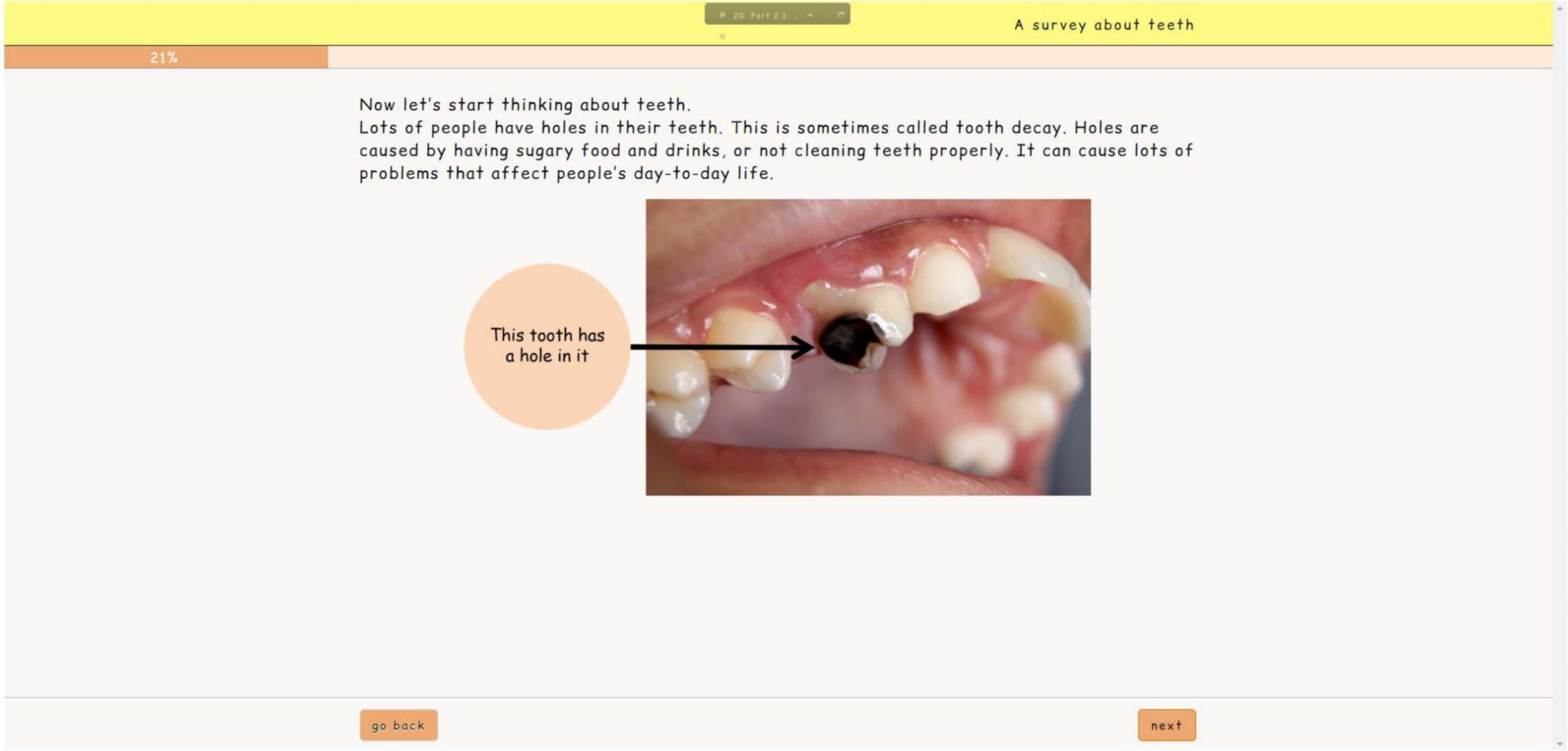
○ A lot

Have you ever had a filling or a tooth taken out because it had a hole or cavity?

Select only one answer

○ Yes

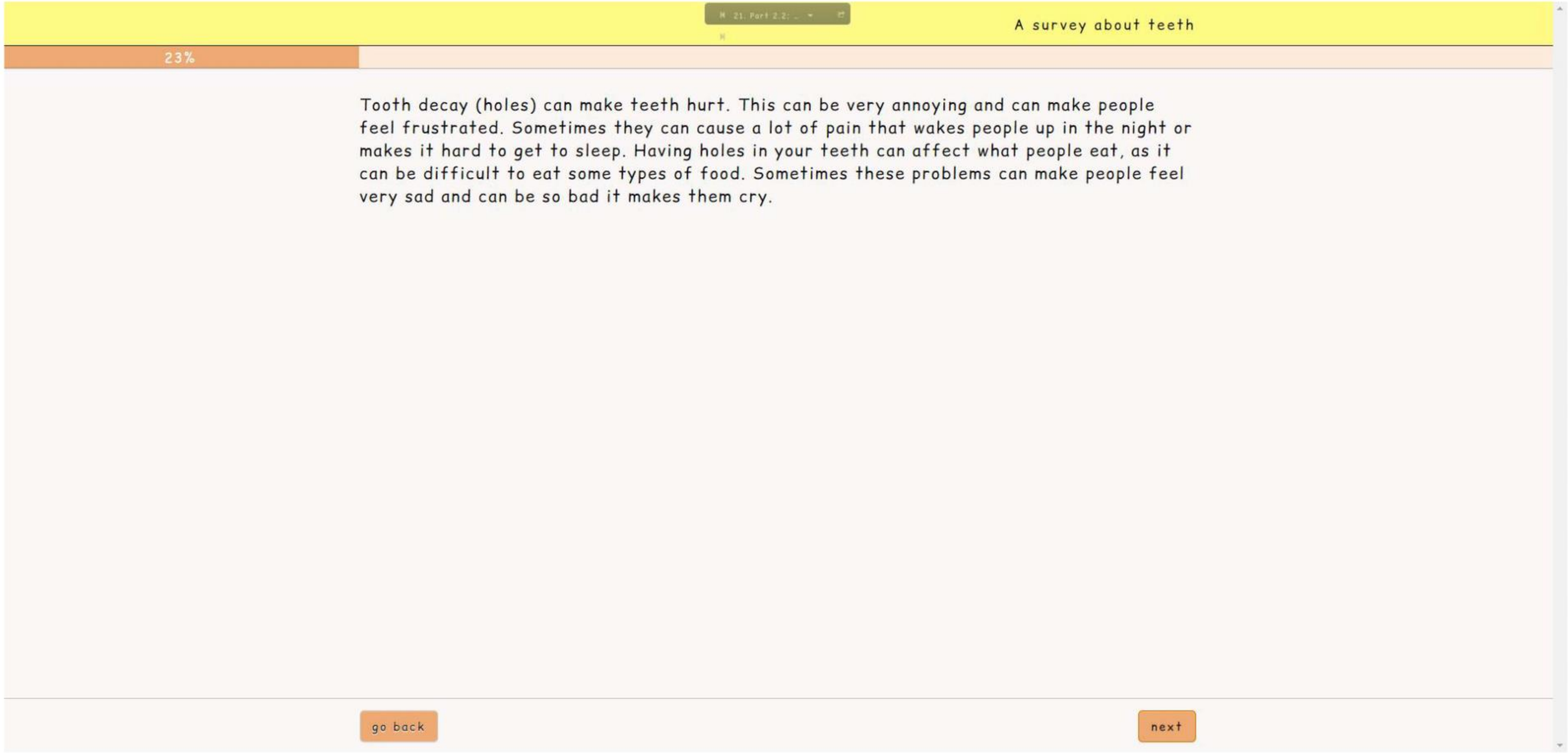
○ No



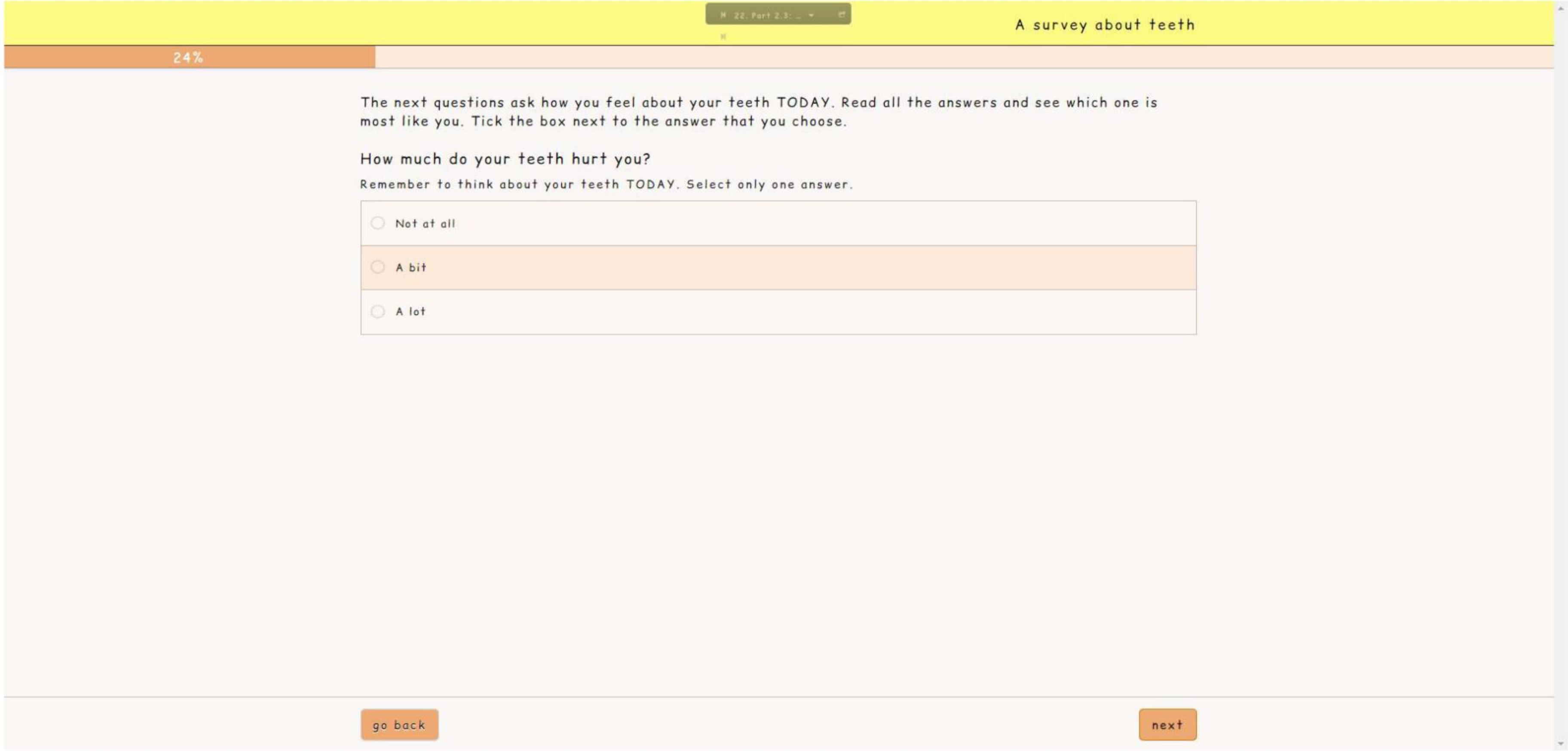
A survey about teeth

Now let's start thinking about teeth.

Lots of people have holes in their teeth. This is sometimes called tooth decay. Holes are caused by having sugary food and drinks, or not cleaning teeth properly. It can cause lots of problems that affect people's day-to-day life.



Tooth decay (holes) can make teeth hurt. This can be very annoying and can make people feel frustrated. Sometimes they can cause a lot of pain that wakes people up in the night or makes it hard to get to sleep. Having holes in your teeth can affect what people eat, as it can be difficult to eat some types of food. Sometimes these problems can make people feel very sad and can be so bad it makes them cry.

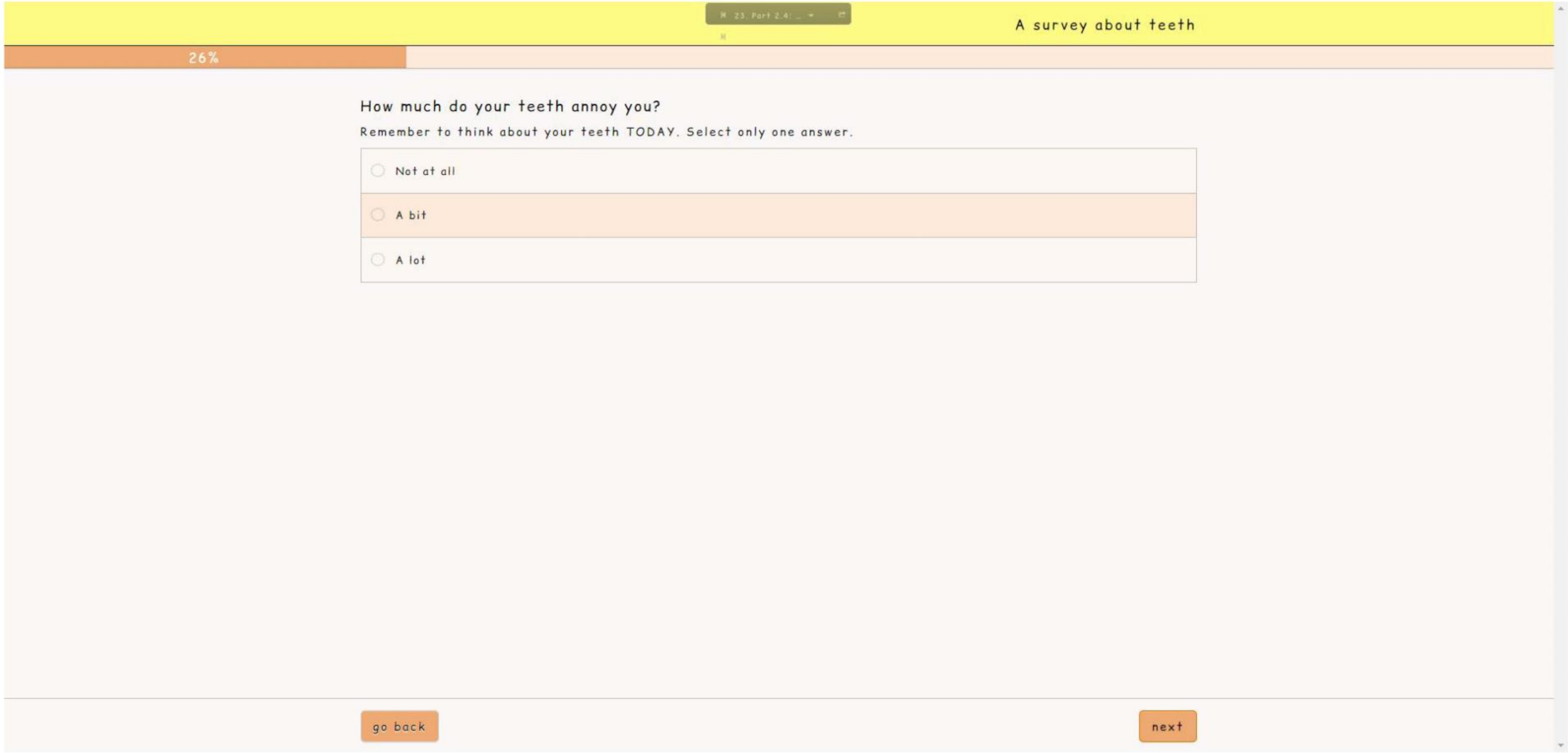


The next questions ask how you feel about your teeth TODAY. Read all the answers and see which one is most like you. Tick the box next to the answer that you choose.

How much do your teeth hurt you?

Remember to think about your teeth TODAY. Select only one answer.

○ Not at all
○ A bit
○ A lot



How much do your teeth annoy you?

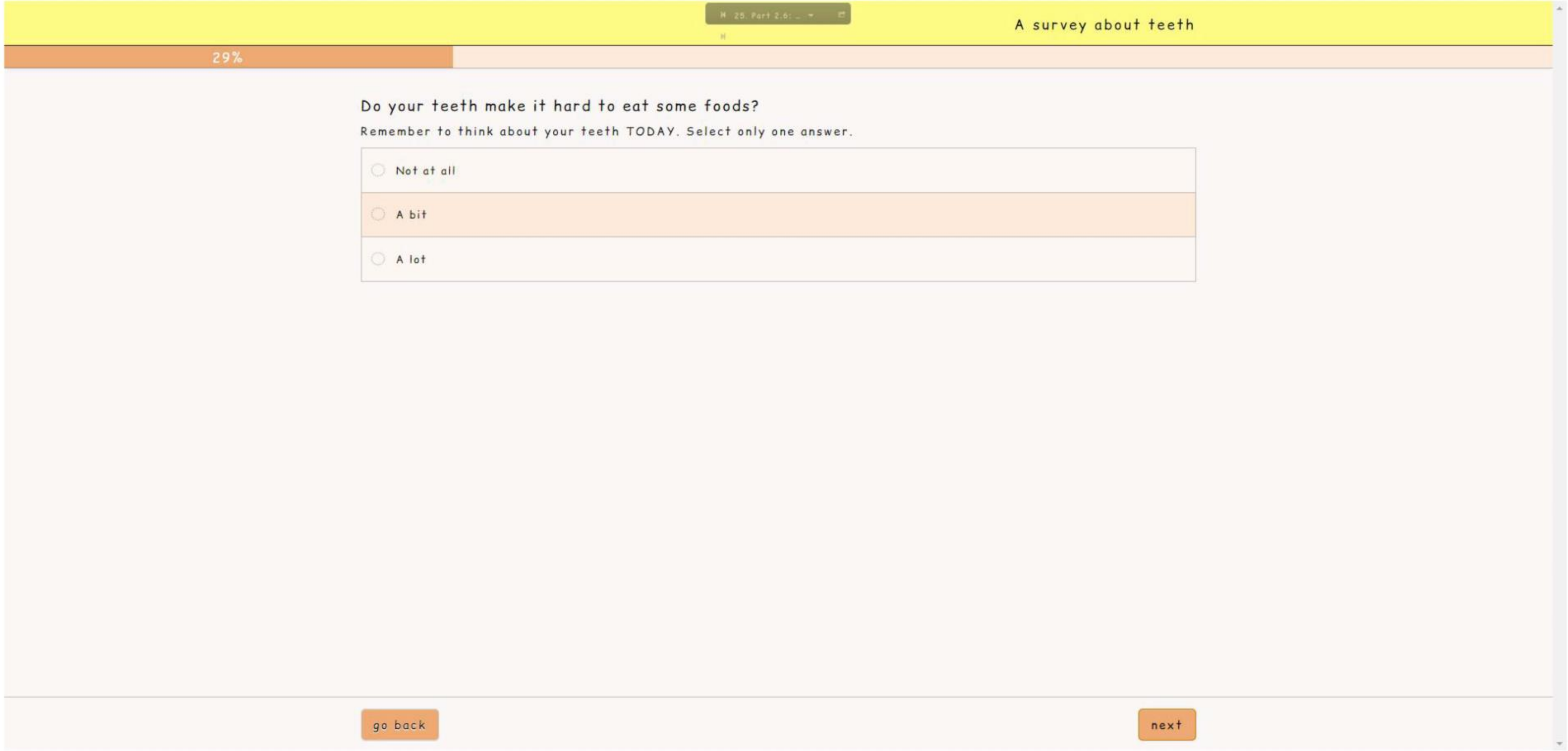
Remember to think about your teeth TODAY. Select only one answer.

- ○ Not at all
- ○ A bit
- ○ A lot

How much do you get kept awake by your teeth?

Remember to think about your teeth TODAY. Select only one answer.

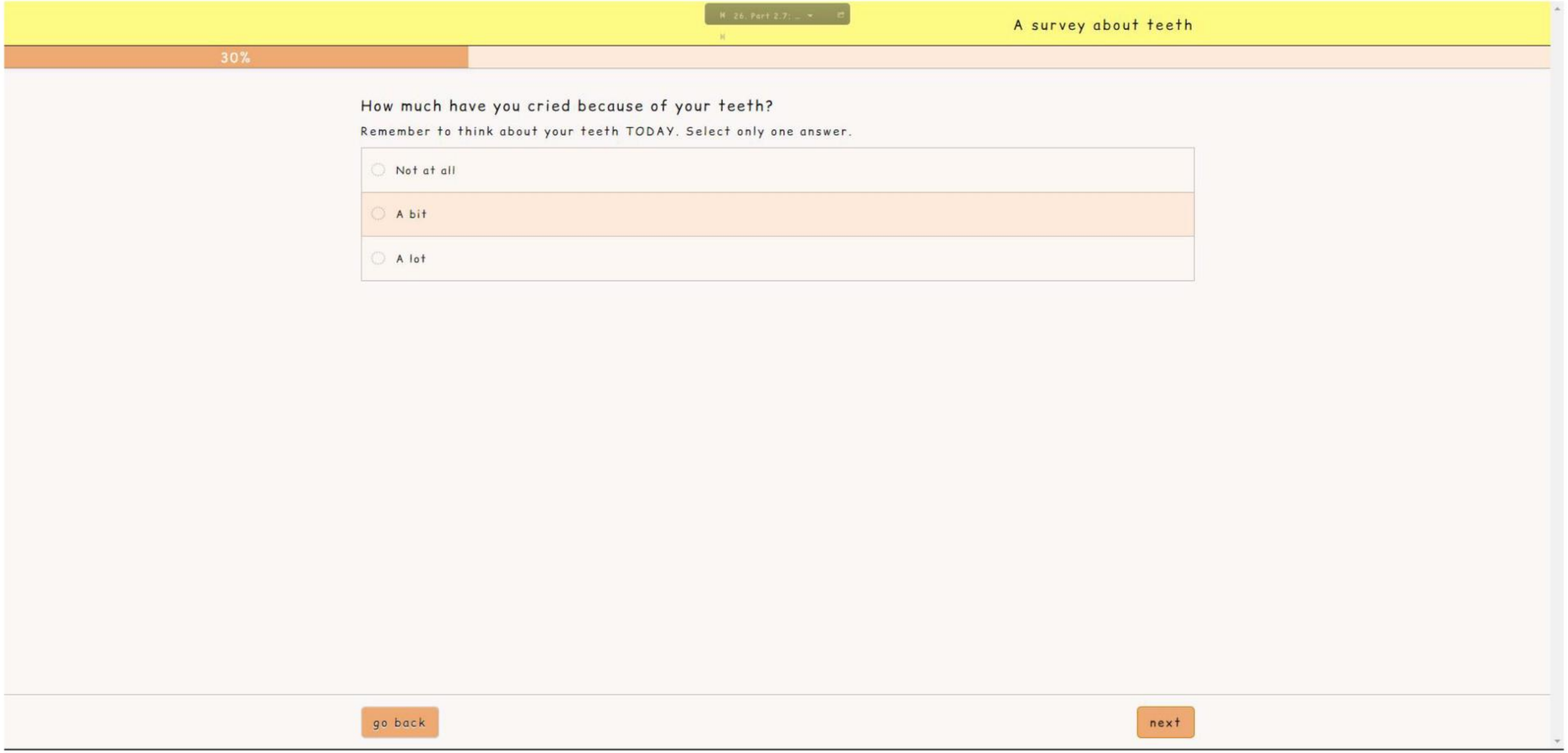
○ Not at all
○ A bit
○ A lot



Do your teeth make it hard to eat some foods?

Remember to think about your teeth TODAY. Select only one answer.

- ○ Not at all
- ○ A bit
- ○ A lot



How much have you cried because of your teeth?

Remember to think about your teeth TODAY. Select only one answer.

○ Not at all

○ A bit

○ A lot

Now we want you to think about what it would be like to live with different problems caused by tooth decay (holes).

There will be nine main questions for you to answer. The next pages will show you how to answer the main questions using an example. There is also a practice question for you to try before you start the main questions.

Try to answer the questions thinking generally, rather than thinking about which answer is most like you.

Example question

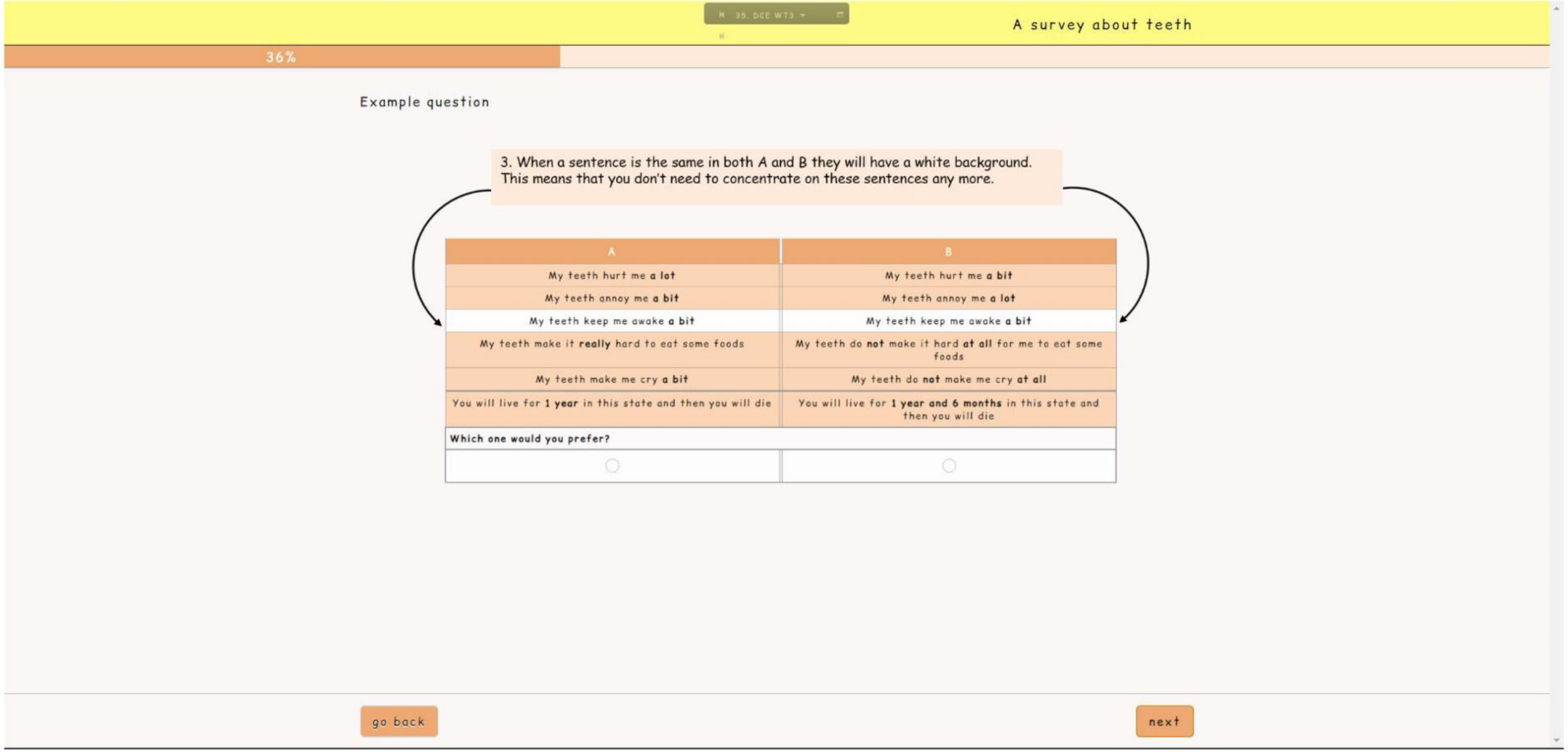
1. There are two options to choose from, A or B. Each sentence describes a different problem that you can have if you get holes in your teeth.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a bit**	My teeth keep me awake **a bit**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
You will live for **1 year** in this state and then you will die	You will live for **1 year and 6 months** in this state and then you will die
Which one would you prefer?	
○	○

Example question

2. The words **in bold** describe how bad the problem is.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a bit**	My teeth keep me awake **a bit**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
You will live for **1 year** in this state and then you will die	You will live for **1 year and 6 months** in this state and then you will die
Which one would you prefer?	
○	○

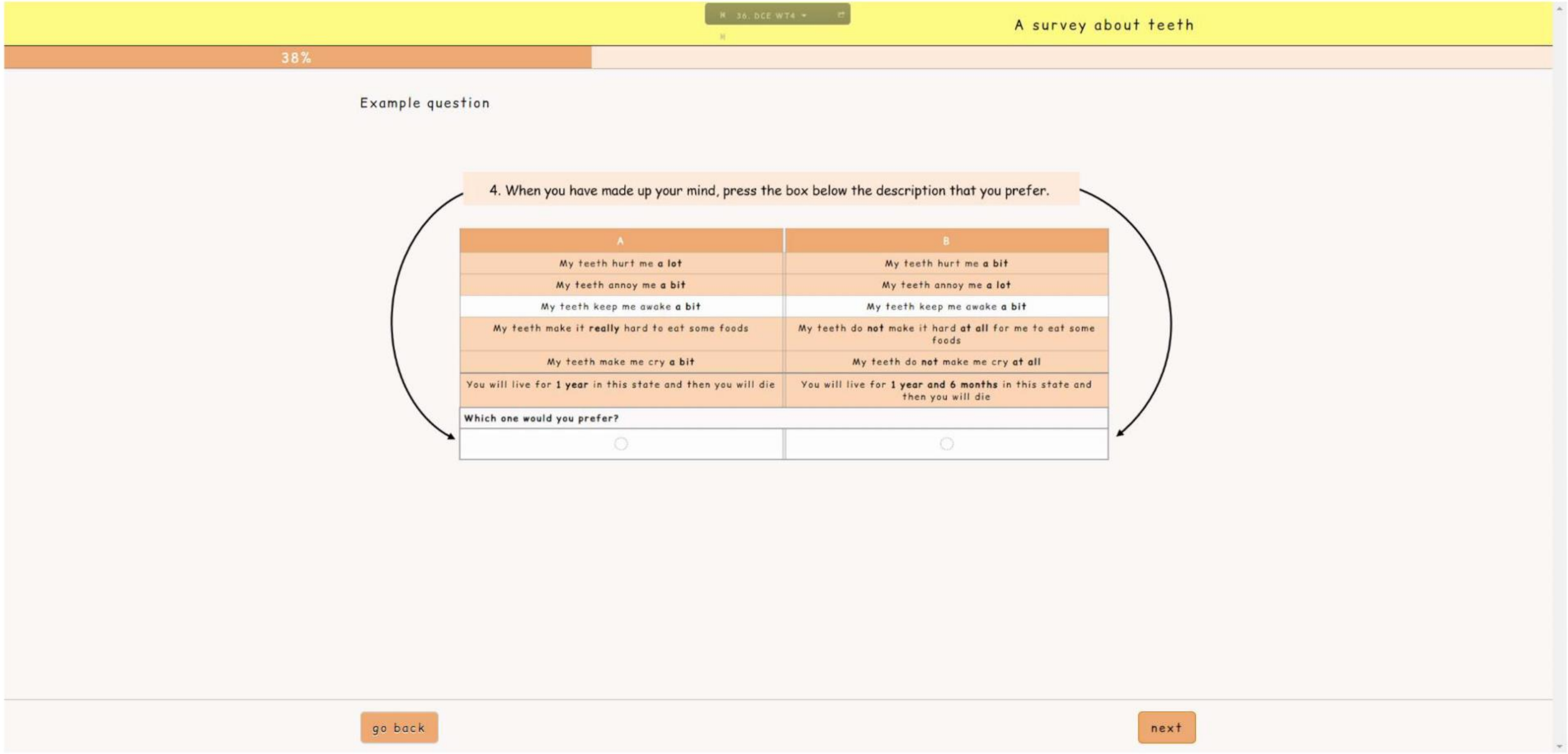


A survey about teeth

Example question

3. When a sentence is the same in both A and B they will have a white background.
This means that you don't need to concentrate on these sentences any more.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a bit**	My teeth keep me awake **a bit**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
You will live for **1 year** in this state and then you will die	You will live for **1 year and 6 months** in this state and then you will die
Which one would you prefer?	
○	○

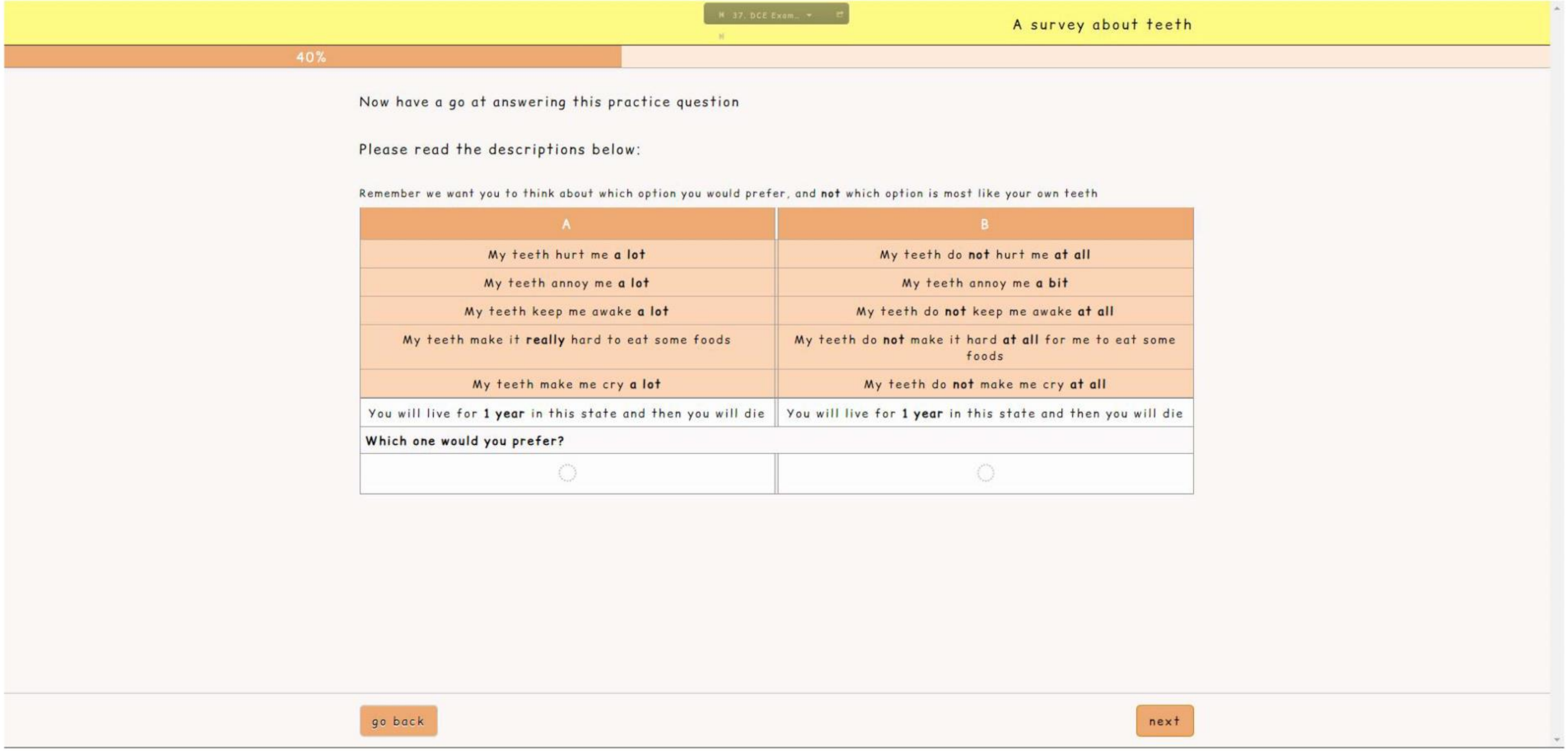


A survey about teeth

Example question

4. When you have made up your mind, press the box below the description that you prefer.

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a bit**	My teeth keep me awake **a bit**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
You will live for **1 year** in this state and then you will die	You will live for **1 year and 6 months** in this state and then you will die
Which one would you prefer?	
○	○

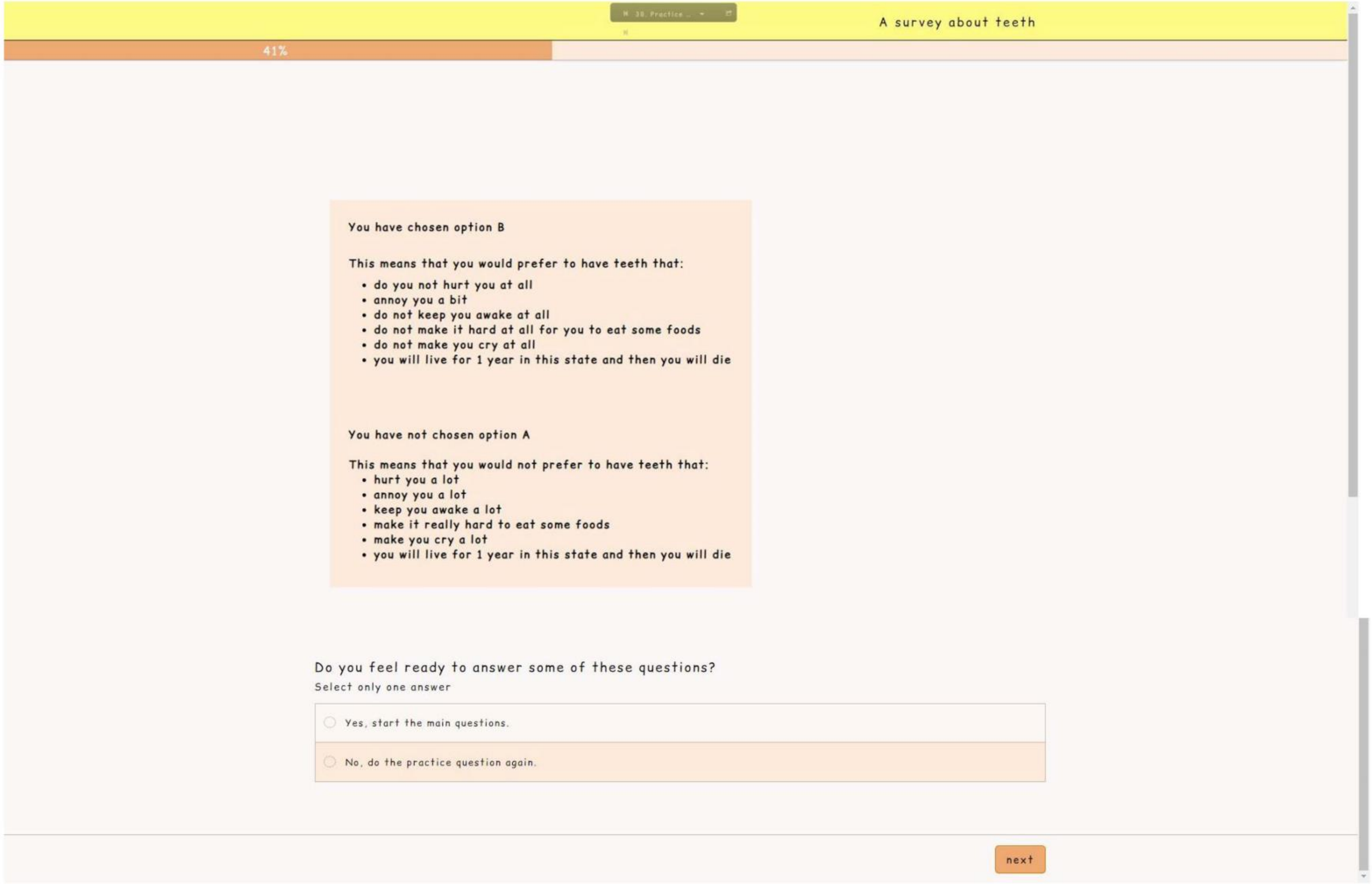


Now have a go at answering this practice question

Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a lot**	My teeth do **not** hurt me **at all**
My teeth annoy me **a lot**	My teeth annoy me **a bit**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a lot**	My teeth do **not** make me cry **at all**
You will live for **1 year** in this state and then you will die	You will live for **1 year** in this state and then you will die
Which one would you prefer?	
○	○



You have chosen option B

This means that you would prefer to have teeth that:

- do you not hurt you at all
- annoy you a bit
- do not keep you awake at all
- do not make it hard at all for you to eat some foods
- do not make you cry at all
- you will live for 1 year in this state and then you will die

You have not chosen option A

This means that you would not prefer to have teeth that:

- hurt you a lot
- annoy you a lot
- keep you awake a lot
- make it really hard to eat some foods
- make you cry a lot
- you will live for 1 year in this state and then you will die

Do you feel ready to answer some of these questions?

Select only one answer

○ Yes, start the main questions.

○ No, do the practice question again.

A survey about teeth

1. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a bit**	My teeth do **not** hurt me **at all**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth do **not** keep me awake **at all**	My teeth keep me awake **a lot**
My teeth make it **a bit** hard to eat some foods	My teeth make it **really** hard to eat some foods
My teeth make me cry **a lot**	My teeth do **not** make me cry **at all**
You will live for **1 year and 6 months** in this state and then you will die	You will live for **1 year and 6 months** in this state and then you will die
Which one would you prefer?	
○	○

2. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth do **not** hurt me **at all**	My teeth hurt me **a bit**
My teeth do **not** annoy me **at all**	My teeth annoy me **a bit**
My teeth do **not** keep me awake **at all**	My teeth keep me awake **a lot**
My teeth make it **a bit** hard to eat some foods	My teeth make it **really** hard to eat some foods
My teeth make me cry **a lot**	My teeth make me cry **a bit**
You will live for **2 years** in this state and then you will die	You will live for **2 years** in this state and then you will die
Which one would you prefer?	
○	○

3. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a bit**	My teeth do **not** hurt me **at all**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a lot**	My teeth keep me awake **a bit**
My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth make me cry **a lot**	My teeth do **not** make me cry **at all**
You will live for **2 years** in this state and then you will die	You will live for **2 years** in this state and then you will die
Which one would you prefer?	
○	○

4. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a lot**	My teeth do **not** hurt me **at all**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **a bit** hard to eat some foods	My teeth make it **really** hard to eat some foods
My teeth do **not** make me cry **at all**	My teeth make me cry **a bit**
You will live for **6 months** in this state and then you will die	You will live for **6 months** in this state and then you will die
Which one would you prefer?	
○	○

5. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth do **not** hurt me **at all**	My teeth hurt me **a lot**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a bit**	My teeth keep me awake **a lot**
My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth make me cry **a lot**	My teeth do **not** make me cry **at all**
You will live for **6 months** in this state and then you will die	You will live for **6 months** in this state and then you will die
Which one would you prefer?	
○	○

6. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth make it **a bit** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
You will live for **1 year and 6 months** in this state and then you will die	You will live for **1 year and 6 months** in this state and then you will die
Which one would you prefer?	
○	○

A survey about teeth

7. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a lot**	My teeth hurt me **a bit**
My teeth annoy me **a lot**	My teeth do **not** annoy me **at all**
My teeth keep me awake **a bit**	My teeth keep me awake **a lot**
My teeth make it **really** hard to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth do **not** make me cry **at all**
You will live for **1 year** in this state and then you will die	You will live for **1 year** in this state and then you will die
Which one would you prefer?	
○	○

8. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth do **not** hurt me **at all**	My teeth hurt me **a lot**
My teeth annoy me **a bit**	My teeth annoy me **a lot**
My teeth keep me awake **a lot**	My teeth do **not** keep me awake **at all**
My teeth do **not** make it hard **at all** for me to eat some foods	My teeth make it **a bit** hard to eat some foods
My teeth do **not** make me cry **at all**	My teeth make me cry **a lot**
You will live for **2 years** in this state and then you will die	You will live for **2 years** in this state and then you will die
Which one would you prefer?	
○	○

9. Please read the descriptions below:

Remember we want you to think about which option you would prefer, and **not** which option is most like your own teeth

A	B
My teeth hurt me **a bit**	My teeth hurt me **a lot**
My teeth do **not** annoy me **at all**	My teeth annoy me **a lot**
My teeth keep me awake **a bit**	My teeth do **not** keep me awake **at all**
My teeth do **not** make it hard **at all** for me to eat some foods	My teeth do **not** make it hard **at all** for me to eat some foods
My teeth make me cry **a bit**	My teeth make me cry **a bit**
You will live for **1 year and 6 months** in this state and then you will die	You will live for **2 years** in this state and then you will die
Which one would you prefer?	
○	○

Thinking about the questions where you were asked to choose between options A and B, did you find them:

Select only one answer

○ Easy to understand

○ Difficult to understand

○ Somewhere in the middle

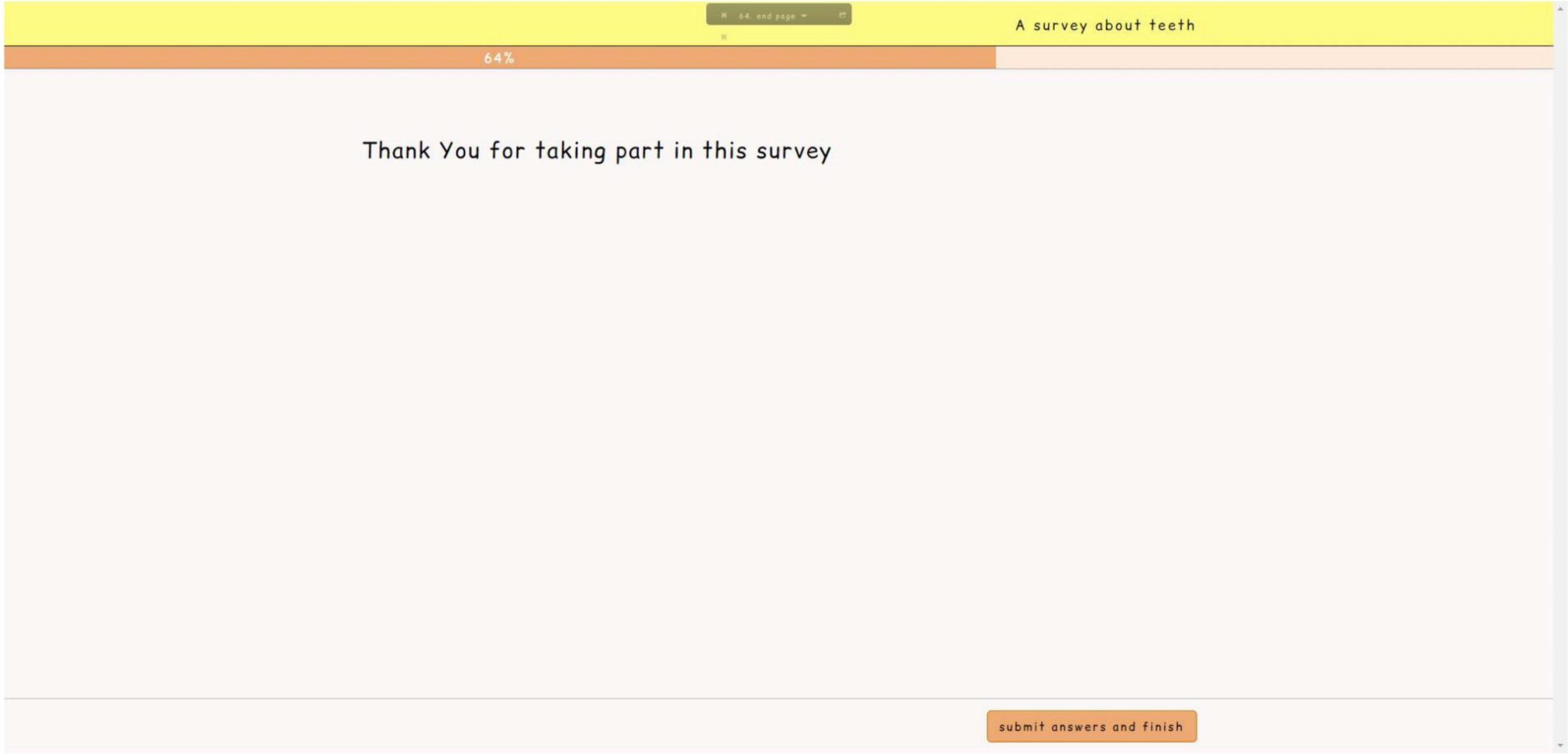
Thinking about the questions where you were asked to choose between options A and B, did you find it:

Select only one answer

○ Easy to choose an answer

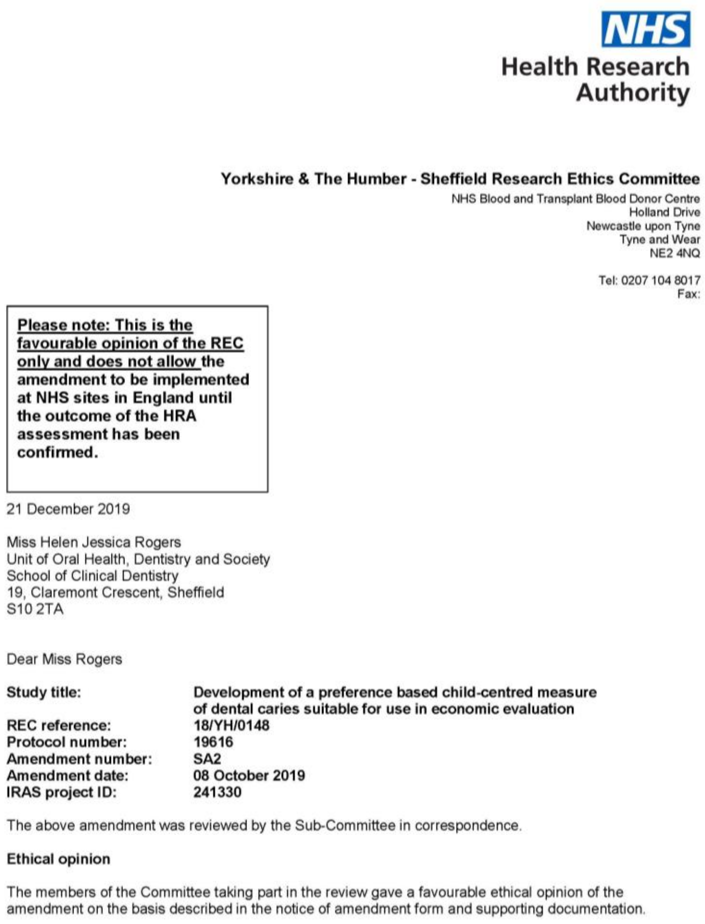
○ Difficult to choose an answer

○ Somewhere in the middle



Thank You for taking part in this survey

V. Favourable ethical opinion for second substantial amendment



NHS
Health Research
Authority

Yorkshire & The Humber - Sheffield Research Ethics Committee
NHS Blood and Transplant Blood Donor Centre
Holland Drive
Newcastle upon Tyne
Tyne and Wear
NE2 4NQ

Tel: 0207 104 8017
Fax:

Please note: This is the favourable opinion of the REC only and does not allow the amendment to be implemented at NHS sites in England until the outcome of the HRA assessment has been confirmed.

21 December 2019

Miss Helen Jessica Rogers
Unit of Oral Health, Dentistry and Society
School of Clinical Dentistry
19, Claremont Crescent, Sheffield
S10 2TA

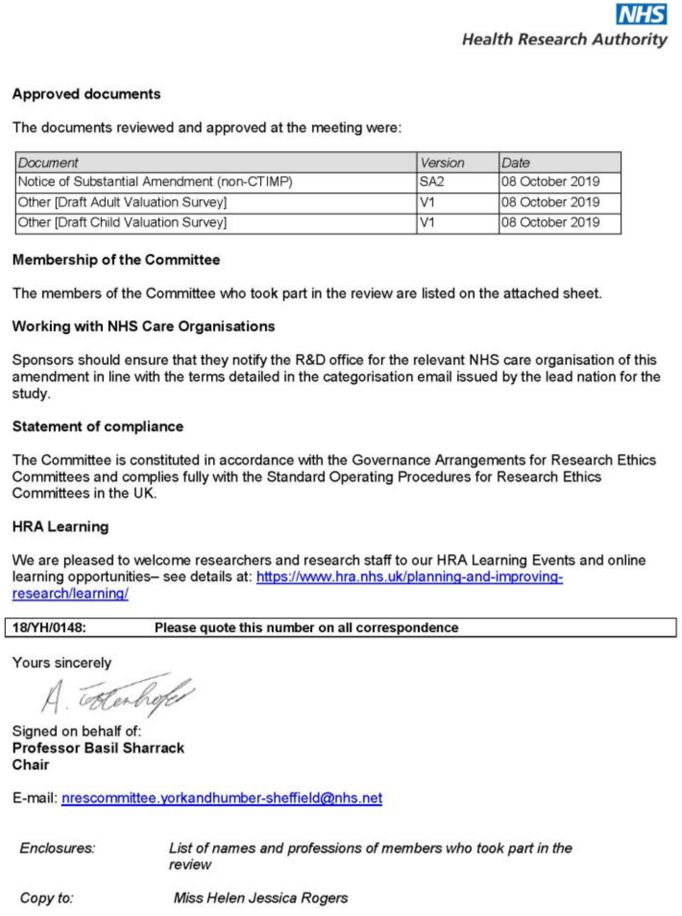
Dear Miss Rogers

Study title:	**Development of a preference based child-centred measure of dental caries suitable for use in economic evaluation**
REC reference:	**18/YH/0148**
Protocol number:	**19616**
Amendment number:	**SA2**
Amendment date:	**08 October 2019**
IRAS project ID:	**241330**

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.



Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMP)	SA2	08 October 2019
Other [Draft Adult Valuation Survey]	V1	08 October 2019
Other [Draft Child Valuation Survey]	V1	08 October 2019

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

Working with NHS Care Organisations

Sponsors should ensure that they notify the R&D office for the relevant NHS care organisation of this amendment in line with the terms detailed in the categorisation email issued by the lead nation for the study.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at: https://www.hra.nhs.uk/planning-and-improving-research/learning/

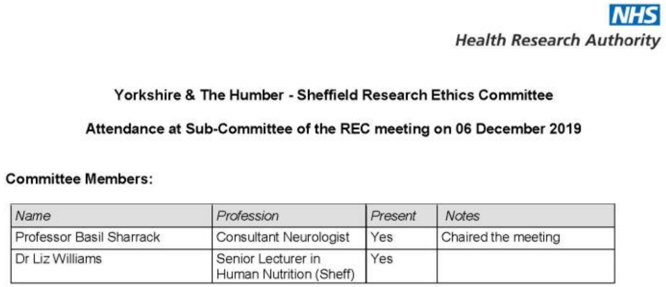
18/YH/0148:	**Please quote this number on all correspondence**

Yours sincerely

Signed on behalf of:
Professor Basil Sharrack
Chair

E-mail: nrescommittee.yorkandhumber-sheffield@nhs.net

Enclosures:	*List of names and professions of members who took part in the review*
Copy to:	*Miss Helen Jessica Rogers*



Health Research Authority

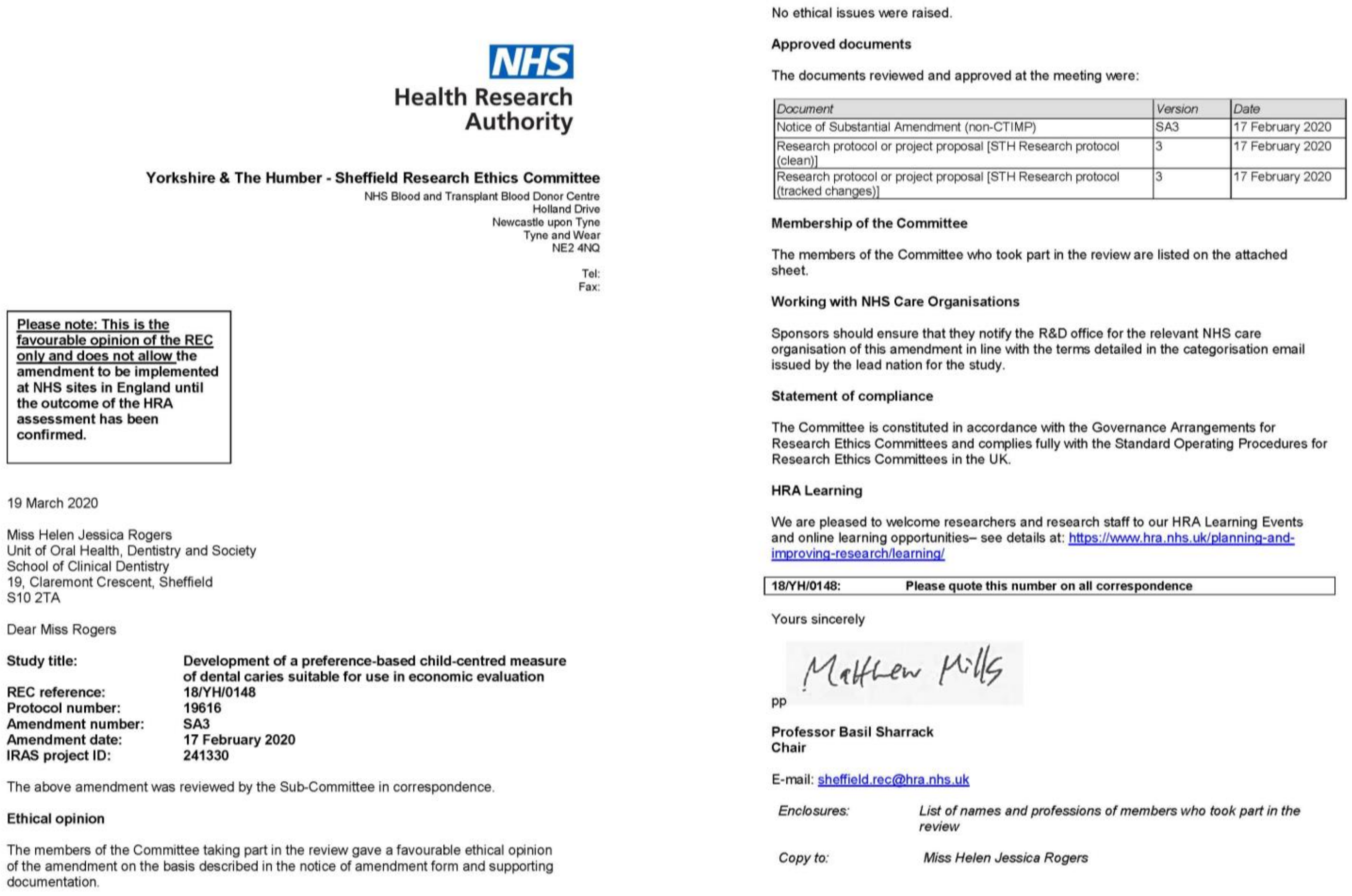
Yorkshire & The Humber - Sheffield Research Ethics Committee

Attendance at Sub-Committee of the REC meeting on 06 December 2019

Committee Members:

Name	*Profession*	*Present*	*Notes*
Professor Basil Sharrack	Consultant Neurologist	Yes	Chaired the meeting
Dr Liz Williams	Senior Lecturer in Human Nutrition (Sheff)	Yes	

W. Third substantial amendment and favourable opinion



NHS
Health Research
Authority

Yorkshire & The Humber - Sheffield Research Ethics Committee
NHS Blood and Transplant Blood Donor Centre
Holland Drive
Newcastle upon Tyne
Tyne and Wear
NE2 4NQ

Tel:
Fax:

Please note: This is the favourable opinion of the REC only and does not allow the amendment to be implemented at NHS sites in England until the outcome of the HRA assessment has been confirmed.

19 March 2020

Miss Helen Jessica Rogers
Unit of Oral Health, Dentistry and Society
School of Clinical Dentistry
19, Claremont Crescent, Sheffield
S10 2TA

Dear Miss Rogers

Study title:	**Development of a preference-based child-centred measure of dental caries suitable for use in economic evaluation**
REC reference:	**18/YH/0148**
Protocol number:	**19616**
Amendment number:	**SA3**
Amendment date:	**17 February 2020**
IRAS project ID:	**241330**

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

No ethical issues were raised.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMP)	SA3	17 February 2020
Research protocol or project proposal [STH Research protocol (clean)]	3	17 February 2020
Research protocol or project proposal [STH Research protocol (tracked changes)]	3	17 February 2020

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

Working with NHS Care Organisations

Sponsors should ensure that they notify the R&D office for the relevant NHS care organisation of this amendment in line with the terms detailed in the categorisation email issued by the lead nation for the study.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

HRA Learning

We are pleased to welcome researchers and research staff to our HRA Learning Events and online learning opportunities– see details at: https://www.hra.nhs.uk/planning-and-improving-research/learning/

18/YH/0148:	**Please quote this number on all correspondence**

Yours sincerely

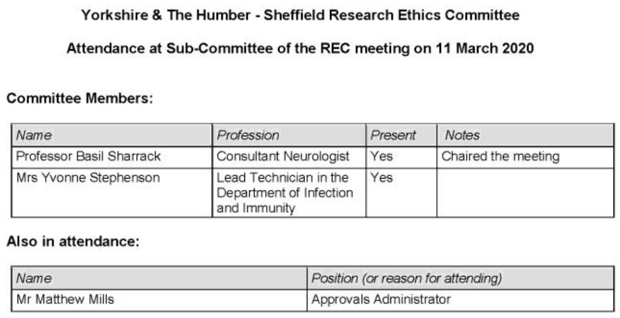
Matthew Mills
pp

Professor Basil Sharrack
Chair

E-mail: sheffield.rec@hra.nhs.uk

Enclosures: *List of names and professions of members who took part in the review*

Copy to: *Miss Helen Jessica Rogers*



Yorkshire & The Humber - Sheffield Research Ethics Committee

Attendance at Sub-Committee of the REC meeting on 11 March 2020

Committee Members:

Name	*Profession*	*Present*	*Notes*
Professor Basil Sharrack	Consultant Neurologist	Yes	Chaired the meeting
Mrs Yvonne Stephenson	Lead Technician in the Department of Infection and Immunity	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Mr Matthew Mills	Approvals Administrator

X. Exploration of heterogeneity in the BWS data

Table 10-5: Regression results when using the reduced sample approach to explore heterogeneity in the BWS data relating to age

Variables	Standard model	11-year-olds	12-year-olds	13-year-olds	14-year-olds	15-year-olds	16-year-olds
Hurt0	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
Hurt1	-2.235***	-2.434***	-2.291***	-2.084***	-2.179***	-2.508***	-2.311***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hurt2	-3.406***	-3.418***	-4.033***	-2.905***	-3.476***	-3.721***	-3.656***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy0	-0.959***	-1.056***	-1.079***	-0.815***	-0.745***	-1.192***	-1.262***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy1	-1.989***	-2.107***	-2.030***	-1.627***	-1.963***	-2.335***	-2.336***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy2	-2.720***	-2.730***	-2.720***	-2.216***	-3.139***	-3.077***	-3.365***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake0	-0.866***	-0.845***	-0.949***	-0.612***	-0.697***	-1.346***	-1.060***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake1	-2.322***	-2.521***	-2.566***	-1.977***	-2.099***	-2.654***	-2.759***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake2	-2.827***	-3.592***	-2.940***	-2.154***	-3.000***	-2.766***	-4.283***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat0	-0.949***	-1.171***	-0.787***	-0.742***	-0.896***	-1.248***	-1.261***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat1	-1.874***	-1.945***	-2.012***	-1.565***	-1.774***	-2.206***	-2.203***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Variables	**Standard model**	**11-year-olds**	**12-year-olds**	**13-year-olds**	**14-year-olds**	**15-year-olds**	**16-year-olds**
Eat2	-2.543***	-2.642***	-2.828***	-2.148***	-2.371***	-2.846***	-3.336***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry0	-0.266***	-0.226	-0.128	-0.350***	-0.064	-0.547***	-0.421***
	(0.000)	(0.106)	(0.338)	(0.001)	(0.602)	(0.000)	(0.005)
Cry1	-2.039***	-2.186***	-2.194***	-1.740***	-2.017***	-2.293***	-2.150***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry2	-3.097***	-2.900***	-3.698***	-2.814***	-3.200***	-3.192***	-3.705***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	56,870	8,480	9,920	12,160	10,080	9,840	7,200
Log likelihood	-14362	-2104	-2381	-3343	-2493	-2471	-1735
Rho-squared	0.215	0.235	0.259	0.158	0.236	0.224	0.263

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot.

Table 10-6: Regression results when using the reduced sample approach to explore heterogeneity in the BWS data relating to participant gender, self-reported general and dental health, and previous caries experience

Variables	Standard model	Male gender	Bad or very bad general health	Current dental problems	Previous caries experience
Hurt0	-	-	-	-	-
	-	-	-	-	-
Hurt1	-2.235***	-2.272***	-3.105***	-1.594***	-1.921***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hurt2	-3.406***	-3.434***	-4.914***	-2.592***	-3.199***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy0	-0.959***	-0.937***	-1.580***	-0.552***	-0.733***
	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)
Annoy1	-1.989***	-1.958***	-3.529***	-1.063***	-1.637***
	(0.000)	(0.000)	(0.005)	(0.000)	(0.000)
Annoy2	-2.720***	-2.632***	-4.028**	-1.670***	-2.215***
	(0.000)	(0.000)	(0.010)	(0.000)	(0.000)
Awake0	-0.866***	-0.830***	-1.452**	-0.460***	-0.704***
	(0.000)	(0.000)	(0.012)	(0.000)	(0.000)
Awake1	-2.322***	-2.263***	-3.439***	-1.395***	-1.892***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake2	-2.827***	-2.584***	-19.060***	-1.773***	-2.293***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat0	-0.949***	-0.937***	-1.359*	-0.430***	-0.751***
	(0.000)	(0.000)	(0.076)	(0.000)	(0.000)
Eat1	-1.874***	-1.790***	-2.100**	-1.148***	-1.534***
	(0.000)	(0.000)	(0.012)	(0.000)	(0.000)
Eat2	-2.543***	-2.360***	-3.488***	-1.678***	-2.155***
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)

Variables	**Standard model**	**Male gender**	**Bad or very bad general health**	**Current dental problems**	**Previous caries experience**
Cry0	-0.266***	-0.302***	-0.671	-0.075	-0.216***
	(0.000)	(0.000)	(0.159)	(0.390)	(0.006)
Cry1	-2.039***	-2.016***	-2.549***	-1.535***	-1.841***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Cry2	-3.097***	-2.988***	-4.593***	-2.312***	-2.857***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	56,870	30,960	720	21,840	28,000
Log likelihood	-14362	-8019	-145.8	-6277	-7535
Rho-squared	0.215	0.201	0.371	0.120	0.171

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot.

Y. Examining the robustness of the BWS model

Table 10-7: Regression results exploring model robustness to exclusion of adolescent participants who failed the dominance test, difficulty understanding or difficulty of choice

Variables	**Standard model**	**Robustness 1** Excluding participants that failed the dominance test[†] first time	**Robustness 2** Excluding participants that found it difficult to understand the tasks	**Robustness 3** Excluding participants that failed the dominance test[†] and found it difficult to understand the tasks	**Robustness 4** Excluding participants that found it difficult to choose an answer	**Robustness 5** Excluding participants that failed the dominance test[†] and found it difficult to choose an answer
Hurt0	-	-	-	-	-	-
	-	-	-	-	-	-
Hurt1	-2.235***	-3.322***	-2.361***	-3.383***	-2.389***	-3.411***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hurt2	-3.406***	-4.918***	-3.614***	-5.132***	-3.641***	-5.178***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy0	-0.959***	-1.569***	-1.018***	-1.569***	-1.036***	-1.583***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy1	-1.989***	-3.418***	-2.146***	-3.482***	-2.172***	-3.500***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy2	-2.720***	-4.619***	-2.922***	-4.643***	-2.968***	-4.696***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake0	-0.866***	-1.517***	-0.929***	-1.513***	-0.947***	-1.532***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Variables	**Standard model**	**Robustness 1**	**Robustness 2**	**Robustness 3**	**Robustness 4**	**Robustness 5**
Awake1	-2.322***	-3.938***	-2.508***	-3.992***	-2.544***	-4.018***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake2	-2.827***	-5.519***	-3.055***	-5.606***	-3.085***	-5.683***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat0	-0.949***	-1.729***	-1.027***	-1.757***	-1.063***	-1.802***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat1	-1.874***	-3.167***	-2.016***	-3.255***	-2.037***	-3.276***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat2	-2.543***	-4.713***	-2.731***	-4.762***	-2.767***	-4.785***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry0	-0.266***	-0.577***	-0.292***	-0.573***	-0.285***	-0.576***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry1	-2.039***	-3.019***	-2.122***	-3.045***	-2.136***	-3.065***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry2	-3.097***	-4.965***	-3.306***	-5.046***	-3.345***	-5.114***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	56,870	31,120	53,760	29,920	52,400	29,520
Log likelihood	-14362	-5857	-13374	-5546	-12942	-5431
Rho-squared	0.215	0.415	0.235	0.424	0.239	0.428

Notes: Robust p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1 †Dominance test was failed if participant was unable to identify the correct 'best' and 'worst' attributes. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot.

Table 10-8: Regression results exploring model robustness to exclusion of adolescent participants who completed the survey very quickly or very slowly

Variables	**Standard model**	**Robustness 6** Excluding participants that completed the survey in < 3 mins	**Robustness 7** Excluding participants that completed the survey in > 30 mins
Hurt0	-	-	-
	-	-	-
Hurt1	-2.235***	-2.263***	-2.234***
	(0.000)	(0.000)	(0.000)
Hurt2	-3.406***	-3.495***	-3.439***
	(0.000)	(0.000)	(0.000)
Annoy0	-0.959***	-0.976***	-0.967***
	(0.000)	(0.000)	(0.000)
Annoy1	-1.989***	-2.023***	-1.999***
	(0.000)	(0.000)	(0.000)
Annoy2	-2.720***	-2.773***	-2.785***
	(0.000)	(0.000)	(0.000)
Awake0	-0.866***	-0.878***	-0.865***
	(0.000)	(0.000)	(0.000)
Awake1	-2.322***	-2.356***	-2.357***
	(0.000)	(0.000)	(0.000)
Awake2	-2.827***	-2.896***	-2.868***
	(0.000)	(0.000)	(0.000)
Eat0	-0.949***	-0.966***	-0.962***
	(0.000)	(0.000)	(0.000)
Eat1	-1.874***	-1.903***	-1.904***
	(0.000)	(0.000)	(0.000)
Eat2	-2.543***	-2.587***	-2.578***
	(0.000)	(0.000)	(0.000)
Cry0	-0.266***	-0.270***	-0.289***
	(0.000)	(0.000)	(0.000)
Cry1	-2.039***	-2.064***	-2.064***
	(0.000)	(0.000)	(0.000)
Cry2	-3.097***	-3.174***	-3.121***
	(0.000)	(0.000)	(0.000)
Observations	56,870	56,230	54,400
Log likelihood	-14362	-14080	-13678
Rho-squared	0.215	0.222	0.219

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth

---

make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot.

Z. Linearity of the duration attribute in the DCE$_{TTO}$

Table 10-9: Regression results from the test of linearity for the DCE$_{TTO}$ duration attribute

Variables	Estimated regressions
Hurt1_LY	-0.372*** (0.000)
Hurt2_LY	-1.217*** (0.000)
Annoy1_LY	0.008 (0.829)
Annoy2_LY	-0.262*** (0.000)
Awake1_LY	-0.207*** (0.000)
Awake2_LY	-0.633*** (0.000)
Eat1_LY	-0.127*** (0.000)
Eat2_LY	-0.354*** (0.000)
Cry1_LY	-0.214*** (0.000)
Cry2_LY	-0.566*** (0.000)
LY12m	1.114*** (0.000)
LY18m	2.114*** (0.000)
LY24m	3.271*** (0.000)
Observations	13,086
Log likelihood	-3467
Rho-squared	0.235

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

AA. Exploration of heterogeneity in DCE_{TTO} survey sample

Table 10-10: Unanchored regressions exploring interaction effects for gender, dental problems and caries experience in the adult DCE_{TTO} data

	Male participants	Self-reported current dental problems	Previous caries experience
Hurt1_LY	-0.413***	0.512***	-0.425***
	(0.000)	(0.000)	(0.000)
Hurt2_LY	-1.334***	1.456***	-1.187***
	(0.000)	(0.000)	(0.000)
Annoy1_LY	0.010	-0.117**	-0.076
	(0.863)	(0.042)	(0.375)
Annoy2_LY	-0.238***	0.407***	-0.253***
	(0.000)	(0.000)	(0.002)
Awake1_LY	-0.173***	0.252***	-0.185**
	(0.000)	(0.000)	(0.016)
Awake2_LY	-0.741***	0.779***	-0.504***
	(0.000)	(0.000)	(0.000)
Eat1_LY	-0.136***	0.215***	-0.057
	(0.008)	(0.000)	(0.476)
Eat2_LY	-0.409***	0.484***	-0.442***
	(0.000)	(0.000)	(0.000)
Cry1_LY	-0.259***	0.329***	-0.257***
	(0.000)	(0.000)	(0.001)
Cry2_LY	-0.639***	0.746***	-0.552***
	(0.000)	(0.000)	(0.000)

Male participants		**Self-reported current dental problems**		**Previous caries experience**	
LY	2.407***		2.692***		2.071***
	(0.000)		(0.000)		(0.000)
male_hurt1_LY	0.072	**problems_hurt1_LY**	0.258***	**caries_hurt1_LY**	0.062
	(0.282)		(0.000)		(0.459)
male_hurt2_LY	0.218***	**problems_hurt2_LY**	0.412***	**caries_hurt2_LY**	-0.045
	(0.007)		(0.000)		(0.647)
male_annoy1_LY	-0.005	**problems_annoy1_LY**	0.228***	**caries_annoy1_LY**	0.104
	(0.945)		(0.003)		(0.279)
male_annoy2_LY	-0.040	**problems_annoy2_LY**	0.270***	**caries_annoy2_LY**	-0.015
	(0.573)		(0.000)		(0.870)
male_awake1_LY	-0.064	**problems_awake1_LY**	0.075	**caries_awake1_LY**	-0.030
	(0.358)		(0.282)		(0.725)
male_awake2_LY	0.201***	**problems_awake2_LY**	0.241***	**caries_awake2_LY**	-0.164*
	(0.007)		(0.002)		(0.074)
male_eat1_LY	0.023	**problems_eat1_LY**	0.157**	**caries_eat1_LY**	-0.089
	(0.742)		(0.029)		(0.317)
male_eat2_LY	0.108	**problems_eat2_LY**	0.225***	**caries_eat2_LY**	0.109
	(0.137)		(0.002)		(0.235)
male_cry1_LY	0.084	**problems_cry1_LY**	0.209***	**caries_cry1_LY**	0.049
	(0.243)		(0.004)		(0.579)
male_cry2_LY	0.136*	**problems_cry2_LY**	0.322***	**caries_cry2_LY**	-0.021
	(0.059)		(0.000)		(0.811)
male_LY	-0.462**	**problems_LY**	0.934***	**caries_LY**	0.118
	(0.012)		(0.000)		(0.613)

Male participants		Self-reported current dental problems		Previous caries experience	
Observations	13,086	**Observations**	13,086	**Observations**	13,086
Log likelihood	-3457	**Log likelihood**	-3439	**Log likelihood**	-3461
Rho-squared	0.238	**Rho-squared**	0.242	**Rho-squared**	0.237

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

Table 10-11: Unanchored regressions examining the interaction effects for employment, marital status and general health in the adult DCE_{TTO} data

	Employed participants	Married participants	Participants with bad or very bad general health
Hurt1_LY	-0.341***	0.500***	0.352***
	(0.000)	(0.000)	(0.000)
Hurt2_LY	-1.295***	1.331***	1.197***
	(0.000)	(0.000)	(0.000)
Annoy1_LY	-0.007	-0.004	0.018
	(0.909)	(0.953)	(0.649)
Annoy2_LY	-0.312***	0.224***	0.257***
	(0.000)	(0.000)	(0.000)
Awake1_LY	-0.225***	0.162***	-0.212***
	(0.000)	(0.003)	(0.000)
Awake2_LY	-0.694***	0.547***	0.625***
	(0.000)	(0.000)	(0.000)
Eat1_LY	-0.183***	0.126**	0.138***
	(0.001)	(0.028)	(0.000)
Eat2_LY	-0.461***	0.393***	0.365***
	(0.000)	(0.000)	(0.000)
Cry1_LY	-0.296***	0.182***	0.222***
	(0.000)	(0.001)	(0.000)
Cry2_LY	-0.699***	0.623***	0.561***
	(0.000)	(0.000)	(0.000)
LY	2.357***	2.178***	2.133***
	(0.000)	(0.000)	(0.000)

Employed participants		Married participants		Participants with bad or very bad general health	
employed_hurt1_LY	-0.059	**married_hurt1_LY**	0.206***	**health_hurt1_LY**	0.354**
	(0.381)		(0.003)		(0.016)
employed_hurt2_LY	0.131	**married_hurt2_LY**	0.182**	**health_hurt2_LY**	0.352**
	(0.109)		(0.028)		(0.038)
employed_annoy1_LY	0.026	**married_annoy1_LY**	0.018	**health_annoy1_LY**	-0.131
	(0.735)		(0.816)		(0.386)
employed_annoy2_LY	0.082	**married_annoy2_LY**	-0.067	**health_annoy2_LY**	-0.063
	(0.257)		(0.361)		(0.654)
employed_awake1_LY	0.025	**married_awake1_LY**	-0.073	**health_awake1_LY**	0.050
	(0.725)		(0.301)		(0.729)
employed_awake2_LY	0.096	**married_awake2_LY**	-0.139*	**health_awake2_LY**	-0.131
	(0.207)		(0.068)		(0.392)
employed_eat1_LY	0.089	**married_eat1_LY**	0.006	**health_eat1_LY**	0.204
	(0.217)		(0.934)		(0.152)
employed_eat2_LY	0.179**	**married_eat2_LY**	0.071	**health_eat2_LY**	0.193
	(0.016)		(0.341)		(0.195)
employed_cry1_LY	0.140*	**married_cry1_LY**	-0.056	**health_cry1_LY**	0.078
	(0.053)		(0.444)		(0.580)
employed_cry2_LY	0.230***	**married_cry2_LY**	0.085	**health_cry2_LY**	-0.079
	(0.002)		(0.247)		(0.610)
employed_LY	-0.331*	**married_LY**	-0.026	**health_LY**	0.490
	(0.077)		(0.891)		(0.208)
Observations	13,086	**Observations**	13,086	**Observations**	13,086
Log likelihood	-3458	**Log likelihood**	-3456	**Log likelihood**	-3462
Rho-squared	0.238	**Rho-squared**	0.238	**Rho-squared**	0.237

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

Table 10-12: Unanchored regressions examining the interaction effects for age in the adult DCE$_{TTO}$ data

Participants aged 65 and over		Participants aged 18-24	
Hurt1_LY	-0.361***		-0.356***
	(0.000)		(0.000)
Hurt2_LY	-1.157***		-1.225***
	(0.000)		(0.000)
Annoy1_LY	0.013		0.021
	(0.761)		(0.601)
Annoy2_LY	-0.229***		-0.260***
	(0.000)		(0.000)
Awake1_LY	-0.182***		-0.199***
	(0.000)		(0.000)
Awake2_LY	-0.611***		-0.645***
	(0.000)		(0.000)
Eat1_LY	-0.130***		-0.119***
	(0.001)		(0.002)
Eat2_LY	-0.325***		-0.325***
	(0.000)		(0.000)
Cry1_LY	-0.163***		-0.214***
	(0.000)		(0.000)
Cry2_LY	-0.490***		-0.565***
	(0.000)		(0.000)
LY	2.008***		2.154***
	(0.000)		(0.000)
older_hurt1_LY	-0.079	**younger_hurt1_LY**	-0.206*
	(0.367)		(0.077)

Participants aged 65 and over		Participants aged 18-24	
older_hurt2_LY	-0.374***	**younger_hurt2_LY**	0.012
	(0.001)		(0.929)
older_annoy1_LY	-0.023	**younger_annoy1_LY**	-0.137
	(0.825)		(0.279)
older_annoy2_LY	-0.199**	**younger_annoy2_LY**	-0.054
	(0.042)		(0.656)
older_awake1_LY	-0.165*	**younger_awake1_LY**	-0.104
	(0.073)		(0.363)
older_awake2_LY	-0.174*	**younger_awake2_LY**	0.083
	(0.093)		(0.513)
older_eat1_LY	-0.002	**younger_eat1_LY**	-0.068
	(0.983)		(0.571)
older_eat2_LY	-0.182*	**younger_eat2_LY**	-0.288**
	(0.070)		(0.021)
older_cry1_LY	-0.314***	**younger_cry1_LY**	-0.020
	(0.001)		(0.862)
older_cry2_LY	-0.462***	**younger_cry2_LY**	-0.009
	(0.000)		(0.941)
older_LY	0.946***	**younger_LY**	0.180
	(0.000)		(0.567)
Observations	13,086	**Observations**	13,086
Log likelihood	-3447	**Log likelihood**	-3460
Rho-squared	0.240	**Rho-squared**	0.237

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

BB. Examining the robustness of the DCE$_{TTO}$ model

Table 10-13: Regression results exploring model robustness to exclusion of adult participants who failed the dominance test or had difficulty understanding the DCE$_{TTO}$ tasks

Variables	Standard model	Standard model anchored	Robustness 1	Robustness 1 anchored	Robustness 2	Robustness 2 anchored	Robustness 3	Robustness 3 anchored
			Excluding participants that failed the dominance test		Excluding participants that found it difficult to understand the tasks		Excluding participants that failed the dominance test and found it difficult to understand the tasks	
Hurt1_LY	-0.373***	-.1728429	-0.425***	-.1742738	-0.389***	-.1729981	-0.398***	-.1733476
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hurt2_LY	-1.217***	-.5635363	-1.354***	-.5548113	-1.283***	-.5708403	-1.315***	-.5722786
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy1_LY	0.009	.004016	-0.034	-.0137991	-0.032	-.014017	-0.031	-.0135733
	(0.820)	(0.821)	(0.414)	(0.409)	(0.460)	(0.455)	(0.472)	(0.468)
Annoy2_LY	-0.262***	-.1212449	-0.313***	-.1280649	-0.289***	-.1287591	-0.302***	-.1312883
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake1_LY	-0.209***	-.0967463	-0.234***	-.0959131	-0.214***	-.0952338	-0.216***	-.0938763
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake2_LY	-0.634***	-.2933218	-0.723***	-.2960955	-0.682***	-.3034562	-0.694***	-.3022111
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat1_LY	-0.126***	-.0582052	-0.149***	-.0609352	-0.136***	-.060538	-0.138***	-.0599877
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Eat2_LY	-0.354***	-.1636369	-0.396***	-.1624236	-0.353***	-.1568854	-0.357***	-.1553178
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

	Standard model	Standard model anchored	Robustness 1	Robustness 1 anchored	Robustness 2	Robustness 2 anchored	Robustness 3	Robustness 3 anchored
Cry1_LY	-0.215***	-.0994839	-0.235***	-.0962163	-0.209***	-.0930287	-0.220***	-.0956501
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry2_LY	-0.565***	-.2615666	-0.618***	-.2532607	-0.580***	-.2580093	-0.584***	-.2543272
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LY	2.160***	-	2.440***	-	2.248***		2.298***	
	(0.000)	-	(0.000)	-	(0.000)		(0.000)	
Observations	13,086		12,240		10,890		10,710	
Log likelihood	-3468		-3118		-2830		-2760	
Rho-squared	0.235		0.265		0.250		0.256	

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

Table 10-14: Unanchored and anchored regression results exploring model robustness to exclusion of adult participants that had difficulty choosing a response to the DCE$_{TTO}$ tasks

Variables	Standard model	Standard model anchored	Robustness 4	Robustness 4 anchored	Robustness 5	Robustness 5 Anchored
			Excluding participants that found it difficult to choose a response		Excluding participants that found it difficult to choose a response and failed the dominance test	
Hurt1_LY	-0.373***	-.1728429	-0.390***	-.1762983	-0.393***	-.1761347
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hurt2_LY	-1.217***	-.5635363	-1.275***	-.5755977	-1.278***	-.572701
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy1_LY	0.009	.004016	-0.003	-.0012436	-0.008	-.003385
	(0.820)	(0.821)	(0.945)	(0.945)	(0.851)	(0.851)
Annoy2_LY	-0.262***	-.1212449	-0.285***	-.1288312	-0.291***	-.1304021
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake1_LY	-0.209***	-.0967463	-0.205***	-.092646	-0.203***	-.0908825
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake2_LY	-0.634***	-.2933218	-0.672***	-.303432	-0.674***	-.3020873
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat1_LY	-0.126***	-.0582052	-0.132***	-.0594476	-0.130***	-.0581255
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Eat2_LY	-0.354***	-.1636369	-0.363***	-.1638541	-0.363***	-.1627197
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry1_LY	-0.215***	-.0994839	-0.222***	-.1003128	-0.227***	-.1018175
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry2_LY	-0.565***	-.2615666	-0.598***	-.2700924	-0.599***	-.2683327

	Standard model	**Standard model anchored**	**Robustness 4**	**Robustness 4 anchored**	**Robustness 5**	**Robustness 5 Anchored**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LY	2.160***	-	2.214***		2.232***	
	(0.000)	-	(0.000)		(0.000)	
Observations	13,086		12,348		12,222	
Log likelihood	-3468		-3218		-3180	
Rho-squared	0.235		0.248		0.249	

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

Table 10-15: Unanchored and anchored regression results exploring model robustness to exclusion of adult participants that completed the DCE_{TTO} survey very quickly or very slowly

Variables	Standard model	Standard model anchored	Robustness 6	Robustness 6 anchored	Robustness 7	Robustness 7 anchored
			Excluding participants that completed the survey in < 3 mins		Excluding participants that completed the survey in > 30 mins	
Hurt1_LY	-0.373***	-.1728429	-0.400***	-.1793527	-0.379***	-.175584
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hurt2_LY	-1.217***	-.5635363	-1.266***	-.5679314	-1.228***	-.5683242
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Annoy1_LY	0.009	.004016	0.010	.0043796	0.010	.004768
	(0.820)	(0.821)	(0.803)	(0.803)	(0.791)	(0.792)
Annoy2_LY	-0.262***	-.1212449	-0.270***	-.1210805	-0.262***	-.1213109
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake1_LY	-0.209***	-.0967463	-0.210***	-.0940191	-0.205***	-.09466
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Awake2_LY	-0.634***	-.2933218	-0.657***	-.2946578	-0.639***	-.2957129
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Eat1_LY	-0.126***	-.0582052	-0.124***	-.055495	-0.130***	-.060248
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Eat2_LY	-0.354***	-.1636369	-0.358***	-.1603906	-0.356***	-.164553
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cry1_LY	-0.215***	-.0994839	-0.217***	-.0972072	-0.216***	-.1001281
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Variables	Standard model	Standard model anchored	Robustness 6	Robustness 6 anchored	Robustness 7	Robustness 7 anchored
Cry2_LY	-0.565***	-.2615666	-0.590***	-.2645564	-0.566***	-.2621079
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	2.160***	-	2.229***	-	2.161***	-
LY	(0.000)	-	(0.000)	-	(0.000)	-
Observations	13,086		12,798		12,618	
Log likelihood	-3468		-3346		-3342	
Rho-squared	0.235		0.246		0.236	

Notes: p-values are in parentheses, where *** p<0.01, ** p<0.05, * p<0.1. An underscore (_) represents an interaction between variables i.e. Hurt1_LY is Hurt1 multiplied by LY. Hurt1:my teeth hurt me a bit; Hurt2: my teeth hurt me a lot; Annoy1: my teeth annoy me a bit; Annoy2: my teeth annoy me a lot; Awake1: my teeth keep me awake a bit; Awake2: my teeth keep me awake a lot; Eat1: my teeth make it a bit hard to eat some foods; Eat2: my teeth make it really hard to eat some foods; Cry1: my teeth make me cry a bit; Cry2: my teeth make me cry a lot; LY: duration

CC. Mapped utility estimates

Table 10-16: Mapped estimates for each health state defined by the CARIES-QC-U classification system, based upon the quadratic mapping model

Health state	DCE$_{TTO}$ prediction	BWS prediction	Mapped utility
00000	1.000	-3.04	1.000
00001	1.000	-4.813	0.937
00002	0.942	-5.871	0.946
00010	0.942	-3.965	0.861
00011	0.903	-5.738	0.903
00012	0.903	-6.796	0.814
00020	0.901	-4.634	0.876
00021	0.901	-6.407	0.785
00022	0.877	-7.465	0.877
00100	0.845	-4.496	0.823
00101	0.845	-6.269	0.727
00102	0.843	-7.327	0.795
00110	0.843	-5.421	0.696
00111	0.836	-7.194	0.891
00112	0.836	-8.252	0.801
00120	0.827	-6.09	0.836
00121	0.827	-7.863	0.741
00122	0.819	-8.921	0.796
00200	0.804	-5.001	0.745
00201	0.804	-6.774	0.642
11112	0.412	-11.517	0.453
11120	0.411	-9.355	0.270
11121	0.411	-11.128	0.133
11122	0.403	-12.186	0.369
11200	0.403	-8.266	0.238
11201	0.388	-10.039	0.481
11202	0.388	-11.097	0.358
11210	0.385	-9.191	0.230
11211	0.380	-10.964	0.634
11212	0.380	-12.022	0.522
11220	0.377	-9.86	0.340
11221	0.377	-11.633	0.207
11222	0.371	-12.691	0.473
12000	0.371	-7.036	0.349
12001	0.356	-8.809	0.244
12002	0.353	-9.867	0.341
12010	0.347	-7.961	0.159
12011	0.345	-9.734	0.085
12012	0.341	-10.792	0.577
12020	0.341	-8.63	0.461

00202	0.780	-7.832	0.746
00210	0.778	-5.926	0.715
00211	0.769	-7.699	0.751
00212	0.769	-8.757	0.649
00220	0.746	-6.595	0.652
00221	0.746	-8.368	0.542
00222	0.739	-9.426	0.762
01000	0.739	-4.07	0.660
01001	0.739	-5.843	0.782
01002	0.739	-6.901	0.682
01010	0.737	-4.995	0.731
01011	0.737	-6.768	0.627
01012	0.730	-7.826	0.699
01020	0.730	-5.664	0.593
01021	0.728	-7.437	0.667
01022	0.728	-8.495	0.558
01100	0.722	-5.526	0.654
01101	0.720	-7.299	0.621
01102	0.713	-8.357	0.733
01110	0.707	-6.451	0.860
01111	0.707	-8.224	0.767
01112	0.704	-9.282	0.669
01120	0.681	-7.12	0.563
01121	0.681	-8.893	0.693
01122	0.681	-9.951	0.586
01200	0.672	-6.031	0.604
01201	0.672	-7.804	0.490
01202	0.670	-8.862	0.570
12021	0.339	-10.403	0.542
12022	0.339	-11.461	0.424
12100	0.323	-8.492	0.382
12101	0.321	-10.265	0.315
12102	0.315	-11.323	0.544
12110	0.312	-9.417	0.233
12111	0.306	-11.19	0.182
12112	0.306	-12.248	0.039
12120	0.288	-10.086	0.030
12121	0.283	-11.859	0.473
12122	0.283	-12.917	0.350
12200	0.283	-8.997	0.402
12201	0.283	-10.77	0.274
12202	0.281	-11.828	0.436
12210	0.281	-9.922	0.310
12211	0.280	-11.695	0.140
12212	0.274	-12.753	0.562
12220	0.274	-10.591	0.445
12221	0.274	-12.364	0.323
12222	0.274	-13.422	0.190
20000	0.272	-6.446	0.255
20001	0.272	-8.219	0.117
20002	0.265	-9.277	0.266
20010	0.257	-7.371	0.437
20011	0.254	-9.144	0.108
20012	0.248	-10.202	0.256
20020	0.242	-8.04	0.372
20021	0.242	-9.813	0.241

01210	0.670	-6.956	0.453
01211	0.664	-8.729	0.686
01212	0.664	-9.787	0.578
01220	0.649	-7.625	0.777
01221	0.649	-9.398	0.677
01222	0.646	-10.456	0.571
02000	0.642	-4.801	0.639
02001	0.642	-6.574	0.527
02002	0.640	-7.632	0.582
02010	0.640	-5.726	0.466
02011	0.631	-7.499	0.510
02012	0.631	-8.557	0.389
02020	0.623	-6.395	0.458
02021	0.616	-8.168	0.583
02022	0.616	-9.226	0.606
02100	0.614	-6.257	0.548
02101	0.608	-8.03	0.695
02102	0.608	-9.088	0.588
02110	0.607	-7.182	0.512
02111	0.605	-8.955	0.475
02112	0.584	-10.013	0.696
02120	0.584	-7.851	0.539
02121	0.584	-9.624	0.420
02122	0.575	-10.682	0.624
02200	0.575	-6.762	0.511
02201	0.573	-8.535	0.402
02202	0.573	-9.593	0.273
02210	0.567	-7.687	0.531
20022	0.218	-10.871	0.373
20100	0.216	-7.902	0.334
20101	0.215	-9.675	0.204
20102	0.215	-10.733	0.062
20110	0.184	-8.827	0.255
20111	0.184	-10.6	0.117
20112	0.183	-11.658	-0.067
20120	0.178	-9.496	0.394
20121	0.178	-11.269	0.265
20122	0.177	-12.327	0.420
20200	0.177	-8.407	0.293
20201	0.175	-10.18	0.355
20202	0.175	-11.238	0.223
20210	0.160	-9.332	0.257
20211	0.160	-11.105	0.177
20212	0.158	-12.163	0.215
20220	0.151	-10.001	0.356
20221	0.151	-11.774	0.090
20222	0.149	-12.832	0.014
21000	0.145	-7.476	0.521
21001	0.145	-9.249	0.401
21002	0.119	-10.307	0.144
21010	0.119	-8.401	0.306
21011	0.119	-10.174	0.171
21012	0.110	-11.232	0.113
21020	0.110	-9.07	-0.034
21021	0.092	-10.843	-0.043
21022	0.087	-11.901	0.414

02211	0.567	-9.46	0.411
02212	0.566	-10.518	0.555
02220	0.566	-8.356	0.437
02221	0.564	-10.129	0.495
02222	0.564	-11.187	0.372
10000	0.558	-5.275	0.504
10001	0.550	-7.048	0.599
10002	0.550	-8.106	0.485
10010	0.549	-6.2	0.403
10011	0.547	-7.973	0.365
10012	0.544	-9.031	0.713
10020	0.544	-6.869	0.607
10021	0.541	-8.642	0.496
10022	0.535	-9.7	0.648
10100	0.535	-6.731	0.537
10101	0.526	-8.504	0.601
10102	0.519	-9.562	0.443
10110	0.517	-7.656	0.378
10111	0.508	-9.429	0.298
10112	0.508	-10.487	0.449
10120	0.508	-8.325	0.324
10121	0.485	-10.098	0.507
10122	0.478	-11.156	0.462
10200	0.478	-7.236	0.338
10201	0.476	-9.009	0.549
10202	0.476	-10.067	0.430
10210	0.470	-8.161	0.386
10211	0.470	-9.934	0.256
21100	0.087	-8.932	0.286
21101	0.080	-10.705	0.238
21102	0.080	-11.763	0.099
21110	0.078	-9.857	0.166
21111	0.078	-11.63	0.023
21112	0.061	-12.688	0.014
21120	0.055	-10.526	0.168
21121	0.054	-12.299	0.197
21122	0.052	-13.357	0.125
21200	0.046	-9.437	0.309
21201	0.046	-11.21	0.174
21202	0.022	-12.268	0.311
21210	0.022	-10.362	0.113
21211	0.022	-12.135	-0.033
21212	0.014	-13.193	0.220
21220	0.014	-11.031	0.079
21221	-0.004	-12.804	0.071
21222	-0.012	-13.862	0.189
22000	-0.012	-8.207	0.046
22001	-0.013	-9.98	-0.143
22002	-0.018	-11.038	0.332
22010	-0.018	-9.132	0.198
22011	-0.036	-10.905	0.190
22012	-0.043	-11.963	-0.005
22020	-0.045	-9.801	-0.084
22021	-0.077	-11.574	0.074
22022	-0.083	-12.632	0.019
22100	-0.083	-9.663	-0.133

10212	0.467	-10.992	0.319
10220	0.467	-8.83	0.185
10221	0.461	-10.603	0.330
10222	0.452	-11.661	0.426
11000	0.450	-6.305	0.177
11001	0.446	-8.078	0.585
11002	0.446	-9.136	0.469
11010	0.444	-7.23	0.526
11011	0.444	-9.003	0.406
11012	0.444	-10.061	0.321
11020	0.443	-7.899	0.348
11021	0.441	-9.672	0.281
11022	0.438	-10.73	0.728
11100	0.438	-7.761	0.623
11101	0.435	-9.534	0.452
11102	0.435	-10.592	0.327
11110	0.427	-8.686	0.398
11111	0.421	-10.459	0.527
22101	-0.101	-11.436	-0.142
22102	-0.109	-12.494	-0.025
22110	-0.115	-10.588	0.171
22111	-0.115	-12.361	0.028
22112	-0.118	-13.419	0.097
22120	-0.118	-11.257	-0.050
22121	-0.135	-13.03	-0.059
22122	-0.141	-14.088	0.099
22200	-0.174	-10.168	0.042
22201	-0.174	-11.941	-0.108
22202	-0.206	-12.999	-0.246
22210	-0.238	-11.093	-0.079
22211	-0.240	-12.866	-0.160
22212	-0.279	-13.924	-0.054
22220	-0.279	-11.762	-0.210
22221	-0.297	-13.535	-0.220
22222	-0.402	-14.593	-0.326

Notes: The position of the numbers within the health state denotes the related item within the classification system, which in order is: ‘hurt’, ‘annoy’, ‘awake’, ‘hard to eat’ and ‘cried’. Each number within the health state represents the severity level of that item: 0 is ‘not at all’, 1 is ‘a bit’ and 2 is ‘a lot’. The mapped utility for health state 00000 has been capped at 1, as described in section 7.4.5.

DD. Email exchange regarding survey complaint

10/02/20

Dear Sir/Madam,

About 15 minutes ago I started Survey P20-10951 in [redacted]. It stated that it had been funded by the National Institute for Health Research and approved by the South Yorkshire Research Ethics Committee (reference number for the survey being 241330). But it contained **very upsetting options** - such as whether I would prefer to live one year or 1 1/2 years and then die! I cannot believe such a survey was approved - especially in a survey about dental health. I have almost never been so appalled by a survey, and I have taken many. I stopped taking the survey because I was so appalled.

Please give this feedback to whatever parties are relevant, and respond at your earliest convenience.

Thank you,

Regards,

[redacted]

12/02/20

Dear [redacted],

Thank you for taking the time to contact me. I am very sorry to hear your concerns around the content of this survey and the distress it has caused you. My apologies for the delay in responding to you - I have been on leave.

I just wanted to inform you that I am in the process of discussing your concerns with the project team, and will send a full response within the next seven days.

Kind Regards,

Helen

15/02/20

Dear [redacted],

Thank you for taking the time to email me regarding the survey you commenced about dental health in [redacted].

I am very sorry to hear that you found the content of this survey upsetting. As you mentioned in your email, this survey had gone through all the necessary checks and approvals before its launch, though I appreciate that this may not be much comfort to you. Whilst you may have read the information provided before commencing the survey, I wanted to take this opportunity to give you a bit more detail about why the questions are phrased in such a way, and how this type of question has been used in previous surveys.

It is hoped that this project will help us to develop a tool that can find out which treatments for children with tooth decay are the best value-for-money, through generating something called quality adjusted life years (QALYs). To do this, it is necessary to ask survey participants to make certain decisions, or trade-offs, between different hypothetical scenarios. This is a technique that has been used in health economics studies for a number of years, for medical conditions such as asthma and cancer, though this is the first time it has been used in a study about dental health. As strange as it seems, the questions have to consider how long the respondent would be prepared to 'live' with the conditions stated. The questions have to state that the respondent will die at the end of the condition otherwise the technique doesn't work properly, as it requires the trading of years of life. We have launched two surveys at the same time – one for children, and one for adults. The survey for children has no mention of length of life, or death in it, as it is widely acknowledged that it is not appropriate to cover this topic with this age group. Unfortunately, to make the results from the child survey meaningful, it is necessary for us to direct these more complex questions to our adult respondents.

We are sorry that these questions have caused you distress, and understand why you stopped the survey. Please note that I have documented your concerns and informed the rest of the project team. In the meantime, if this survey has raised any issues for you, please call NHS 111 for further advice.

Kind Regards.

Helen

18/02/20

Thank you for the explanation. I did suspect that this was the reason, but if there was some way of wording it less... startlingly... that would be far better. Possibly something like, "The rest of my life would be ____ years, and all of them lived with this condition." That states it in a more positive way while still keeping the intent. I hope that helps. :)

Thank you for documenting my concerns! I appreciate your taking them seriously.

Best wishes!

[redacted]

19/02/20

Dear [redacted],

Many thanks for your email. I really like your suggestion about changing the wording and the example you provided. We are planning to write up the findings of this study to publish in a scientific journal that will be read by other researchers who are conducting similar surveys. I wondered whether you would be happy for us to mention your suggestion of altering the wording (anonymously of course) in this publication, as it could help to prompt other researchers to consider this when designing their surveys?

Do let me know your thoughts.

Best Wishes,

Helen

19/02/20

Certainly! I'm so glad to help! :D