# Using Cognitive Interviewing to assess the Validity of Acceptance and Commitment Therapy (ACT) Questionnaires in individuals with Chronic Pain.

Holly Victoria Castle

Submitted in accordance with the requirements for the degree of
Doctor of Clinical Psychology (D. Clin. Psychol.)

The University of Leeds
School of Medicine
Academic Unit of Psychiatry and Behavioural Sciences

July, 2019

# *Acknowledgements*

Firstly, I would like to thanks every person who volunteered their time as participants, your generosity allowed this study to happen.

To my supervisors Dr Chris Graham and Dr Gary Latchford, thank you for all your encouragement throughout. Your confidence in me and the study has kept me focused and motivated- through the tough times especially! I am grateful for all the guidance and knowledge you have instilled in me, your commitment to the work and all the time you have given me. Your boundless enthusiasm has rubbed off on me and will influence my future work; clinical and research, in the years to come. Thank you to Professor Stephen Morley, I was privileged to have received your support in the early stages of the study, you are missed. Thank you to Dr Vivienne Laidler for all your patience and support with recruitment to the study.

I am also grateful to all the ACT clinicians involved in the initial survey, their time and thoughtful comments were a real testament to the ACT community and helped to shape the main study. Special thanks also go to the British Pain Society, Pain Management Programme - Special Interest Group, for your contributions to this, and to Everybody's Voice (Service User and Carer reference group) with your support during the study design.

Thank you to my marvellous cohort(s), especially to the library crew for keeping me going and the fun times to counteract the tough. I have made lifelong friends along the way, special thanks to Lucy O'Neill and Becky Anderson, your encouragement has kept me going this year and hopefully I will soon join up with the pack!

Lastly, my eternal thanks go to my family, especially to Anokh, Nahla and Sophie. The patience, interest and understanding that you have shown is so appreciated. You have listened to me and encouraged me and I wouldn't be in the position I am without that.

*"As in most of science, the true beauty lies within the story."*

**Brian Cox (2019), The Planets, BBC.**

# *Abstract*

**Introduction:** To enable clinicians to adapt treatment methods during therapy and to further refine ACT theory through research, we must be able to measure processes of change. Existing research has supported the psychometric quality of ACT questionnaires used within the general population and for those designed for individuals experiencing chronic pain. Correlation studies have demonstrated evidence for ACT questionnaire reliability (internal-consistency, test-retest reliability) and for aspects of construct validity (discriminant, convergent validity). However, studies using quantitative methodology (factor analysis or multiple regression) to assess validity do not tell us *how* individuals understand questionnaire items or whether they are interpreted as intended. No existing study provides a comprehensive, empirical investigation of this important aspect of construct validity; *content validity*.

**Aim:** The current study investigated whether commonly used ACT questionnaires captured their intended processes in individuals experiencing chronic pain; to explore content validity.

**Method:** The study was conducted in two parts: 1) to establish the most frequently used ACT measures within research and clinical practice; 2) the main study: assessing ACT questionnaire validity. Participants receiving ACT for chronic pain were recruited. Cognitive interviewing methodology investigated how individuals made sense of questionnaires as they completed them. A taxonomy of problem classification was used to objectively classify errors made in completing items.

**Results:** Logical, lexical, conceptually inconsistent and ‘response’ errors were found within each ACT measure assessed (CPAQ-8, PIPS and AAQ-II). Patterns of error making demonstrated how accurately the items were understood; a number of items were identified as the most problematic.

**Discussion:** Findings showing where problems lie in respondent understanding could influence how future research may evaluate processes occurring during treatment, thus refining ACT theory itself. The findings are also important for clinical practice as valid measures are needed to understand how therapy works, enabling clinicians to adapt treatment, targeting specific sub-processes, thus improving outcomes.

# *Abbreviations*

AAQ = Acceptance and Action Questionnaire

ACT = Acceptance and Commitment Therapy

BOS = Bristol Online Survey

BPS = British Pain Society

CBT = Cognitive Behavioural Therapy

CI = Cognitive Interviewing

CompACT = Comprehensive Assessment of ACT processes

CRPS = Complex Regional Pain Syndrome

CPAQ = Chronic Pain Acceptance Questionnaire

DSM-5 = Diagnostic and Statistical Manual

FND = Functional Neurological Disorder

IASP = International Association for the Study of Pain

ICC = Intra-class Correlations Coefficients

MS = Multiple Sclerosis

PSEQ = Pain Self Efficacy Questionnaire

PF = Psychological Flexibility

PIPS = Psychological Inflexibility in Pain Scale

PMP = Pain Management Programme

RCT = Randomised Controlled Trial

RFT = Relational Frame Theory

TMD = Temporomandibular Disorder

# *Table of Contents*

# *List of Figures*

# *List of Tables*

# CHAPTER ONE

## Chapter overview

This chapter introduces the context of the study and provides a comprehensive review of the relevant literature. It provides background information on chronic pain, its prevalence, psychological factors and treatment options. It then focuses on ACT and its use for the treatment of chronic pain, measuring change using ACT process measures and the literature surrounding their validity. The chapter finishes with identifying the problem this study wishes to address through the research- to investigate whether these measures are understood by people experiencing chronic pain, thus establishing their content validity.

# GENERAL INTRODUCTION

## Background on Chronic Pain

### *What is Pain and defining 'Chronic Pain'*

Originally deriving from the Latin *poena*, meaning 'penalty' or 'punishment', the meaning of the word pain evolved to signify a *response* in our bodies, serving a function of alerting us to danger, hurt and suffering and protecting us from further harm. Although it can be a transient experience, for some individuals, pain continues past the point of being useful and can impact greatly on quality of life. The definition endorsed by the International Association for the Study of Pain (IASP) (updated from Merskey & Bogduk, 1994) views pain as an *"unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage."* (IASP, 2011). Pain is not purely a physical experience, but exists along a continuum of individual subjective experience. This definition also hints at the idea that pain can exist without the presence of physical damage. Williams and Craig (2016) proposed a revised definition which includes the *"sensory, emotional, cognitive and social components"* of pain, adding to our understanding of the functional and adaptive nature of this process. Furthermore, they suggest that referring to pain as merely '*unpleasant*' trivialises this experience for many and *'distressing'* is proposed as a more appropriate terminology. No matter the origin of pain, there is no denying its impact on an individual's life.

There have been attempts to categorise types of pain, for example the British Pain Society (BPS) define acute pain as that lasting no longer than 12 weeks, whilst chronic pain has an extended duration of over 12 weeks (BPS, 2014). Unlike for acute pain, there appears to be no definitive definition of chronic pain. The consensus however, is that it is persistent pain which occurs when the physiological *"process of repair is apparently ended"* (IASP, 2011). Although we have moved on from labelling non-anatomical pain as *'psychopathological'* the current Diagnostic and Statistical Manual (DSM-5) combination of pain disorder with hypochondriasis to form a 'somatic symptom disorder' has been criticised for perpetuating stigma and its lack of specificity, leading to over diagnosis (Katz, Rosenbloom, & Fashler, 2015). Attempting to classify does not always aid our understanding of pain experience; both acute and chronic pain can be caused by potential physical damage through injury, yet can also occur without actual physical causation. For example, in the case of recurrent acute pain in individuals with multiple sclerosis (MS), painful spasms have been linked - using magnetic resonance imaging - to lesions in the brain (Spissu, Cannas, Ferringno, Pelaghu & Spissu, 1999). Yet, pain related to a different kind of episodic pain, osteoarthritis, has weak associations with radiographic findings in a large-scale systematic review conducted by Bedson and Croft (2008). These conditions do not fit neatly into either category and despite also being classed as examples of chronic pain they are not likely to complete the healing *'process of repair'*.

Our understanding of the causal and maintaining factors in pain is complicated by the fact that it is not a single experience. Rather, multiple, individual factors have been implicated in the transition from acute to chronic pain, all of which incorporate a complex interaction between biological, sensory, social, emotional, environmental, behavioural and psychological components (Lavand'homme, 2017). In literature surrounding chronic postsurgical pain, these risk factors determine its development, for example certain genetic polymorphisms (predisposing genes) and central sensitisation (heightened nervous system sensitivity to pain) have been associated with chronic pain (Fingleton, Smart, Moloney, Fullen, & Doody, 2015; Katz & Seltzer, 2009). We can begin to see what a complex phenomenon chronic pain is; firstly in defining it but also in treating it.

## *Prevalence and Impact of Chronic Pain*

These definitional problems make it challenging to estimate the prevalence of chronic pain. Despite varying diagnostic criteria for conditions such as fibromyalgia and reliance on an individual’s subjective report of pain symptoms, its prevalence has been estimated at 5.4% of the population (Jones et al., 2015). However, a recent systematic review suggests that chronic and chronic widespread pain affects between one-third to one-half of the UK’s population (Fayaz, Croft, Langford, Donaldson, & Jones, 2016). Here, chronic pain was found to occur across the lifespan increasing up to 62% in those over 75 years old, suggesting it may likely continue increasing as our population ages. Chronic pain also varies between the sexes with certain conditions including migraine and tension-type headaches, temporomandibular disorder (TMD) pain and abdominal pain being more common in women than in men (LeResche, Mancl, Drangsholt, Saunders & Von Korff, 2005). Despite these variances, it is clear that chronic pain can impact any individual and has been associated with limitations in physical, social, emotional and occupational functioning (Breivik, Collett, Ventafrida, Cohen, & Gallagher, 2006).

There are direct and indirect costs to experiencing chronic pain. Depression and anxiety have been consistently associated with chronic pain throughout the literature (Dersh, Polatin & Gatchel, 2002; Magni, Caldieron, Rigatti-Luchini & Merskey, 1990; McWilliams, Goodwin & Cox, 2004; Rayner et al., 2016). Yet the direction of this relationship is unclear and may be reciprocal in nature, with depression and anxiety impacting on levels of pain and vice versa (Lerman, Rudich, Brill, Shalev & Shahar, 2015). The mediational effects of other comorbidities with chronic pain may be at play. For example, the impact of chronic pain on mood has been found to be mediated by levels of fatigue, sleep difficulties and anxiety in people with MS (Amtmann et al., 2015). According to a large-scale meta-analysis with over 110 studies, individuals with chronic pain are more likely to report psychological distress with the largest effect sizes found for anxiety related to the pain ($d$= -1.15) and somatisation ($d$= -1.2) (Burke, Mathias, & Denson, 2015). Although anxiety and depression were also reported, these findings provide more insight into the mechanism by which individuals living with chronic pain are affected through their fear of the pain. Chronic pain can also interfere with social functioning and consequently have direct financial costs, should an individual be out of work. Back pain alone has been previously estimated to have direct health, care and production costs of around £10.7 billion (Maniadakis & Gray, 2000). More recent figures for patients with back pain have been found to be double those of matched controls in a large-scale UK study, with estimations for total nationwide health costs of up to £2.8 billion (Hong, Reed, Novick, & Happich, 2013). Thus the economic burden of chronic pain is not only for the individual, but also health services and wider society.

## *The Psychology of Pain*

It is often difficult to separate out the direct impact of chronic pain itself with other difficulties experienced by individuals living with chronic pain. The previously discussed comorbidities linked to chronic pain may act as both antecedents and consequences. For example, if a person is not able to participate in previously enjoyable activities because of the pain, this may itself lead to social isolation and impact on mood and feelings of self-worth, which in turn could further an individual's experience of pain. The next section will focus on understanding the mechanisms underlying pain, an important factor in providing effective treatment.

The biomedical model explains pain in relation to neurophysiological causes and as a result the focus of medical interventions include surgery and medication. However, as previously mentioned, this conceptualisation of chronic pain is limited to its focus on the observable and does not take into account the existence of pain in the absence of physical damage. Given this complexity it is unsurprising that direct biomedical interventions to reduce the pain are often unsuccessful (Reid et al., 2011). Chronic pain is now also understood as a perceptual experience with an individual's interpretation of pain that can determine its impact. Early psychological models such as the gate control theory by Melzack and Wall (1965) were the beginnings of understanding how psychological processes might impact on how an individual experiences pain. More recent literature has suggested that although over 50 years old, the 'gate' metaphor is still recognised as a useful way to explain pain and is commonly used in chronic pain management groups today (Katz & Rosenbloom, 2015). The model suggests that the spinal dorsal horn 'gating mechanism' controls electrical nerve impulses that are sent to the brain. It can be 'opened' by physical injury or by the brain itself through many different psychological factors influencing pain perception, including anxiety, fear and focus on the pain. This theory widely influenced how we understand pain, its dependence on context and meaning, and how we might treat it, however it does not explain the transition from an individual's experience of acute to chronic pain (Moayedi & Davis, 2012).

There have been a variety of psychological explanations for the experience of pain. Attempts to control pain have been found to increase hypervigilance and attention towards the pain (Notebaert et al., 2011; Van Damme, Crombez, Eccleston & Koster, 2006). Attentional-bias towards controlling pain could have a detrimental impact on how someone engages with their environment, their enjoyment of activities and their pursuit of meaningful goals. Tabor et al. (2016) conducted a study based on the notion that chronic pain changes the way an individual perceives their environment, with findings demonstrating that pain impacts on decisions to 'engage with their environment' and how much effort they attribute towards certain tasks. Although this process may be reciprocal in nature, it is clear that chronic pain emerges in relation to the way an individual perceives pain, whether it is evaluated as dangerous and therefore whether it requires protective behaviour. When pain is interpreted as threatening and 'catastrophised', this fear of the pain can lead to an avoidance of pain related activities, an over awareness of bodily sensations and eventually disuse of certain muscles. These patterns of thinking eventually become engrained over time and the cycle of fear and avoidance can lead to low mood, increased disability and further perpetuate chronic pain. Pain-related fear (e.g. the fear of re-injury or fear of movement) can ultimately lead to an avoidance of behaviours associated with that pain (Crombez, Vlaeyen, Heuts, & Lysens, 1999; Riley, Ahern & Follick, 1988). Psychosocial factors can also influence an individual's experience of pain as shown in a systematic review carried out by Krahé, Springer, Weinman and Fotopoulou (2013). Their findings demonstrated that the presence or perception of social support influence how we interpret potential 'threats' in our environment, which in turn impact on the perception of pain signals. They argued that these effects were also a product of individual differences in attachment and coping methods.

A Cognitive-Behavioural Fear-Avoidance model proposed by Vlaeyen and Linton (2000) explains how avoidance behaviours such as not partaking in certain activities due to pain, can lead to the development and maintenance of chronic pain. An individual's beliefs about the meaning of their symptoms and their ability to control pain have been suggested as mechanisms for chronic pain (Turk & Okifuji, 2002). A recent study conducted by Van

Ryckeghem, Noel, Sharpe, Pincus and Van Damme (2019) also investigated the role of cognitive biases in the maintenance of pain. They suggest that interpretation biases towards pain related information increase the risk of acute pain transitioning to chronic pain. Here, it is not only what individuals fear about pain, but also their methods for coping and interpretations that are important. A development on this approach to explaining pain is based on the notion of ‘psychological inflexibility’; which locates the cause of psychological distress in relation to external environmental factors or ‘patterns of behaving’ as opposed to internal constructs (Hayes, 1995). The role of the adaptive response of ‘psychological flexibility’ will later be discussed in detail, but for now can be summarised as a shift in the way an individual behaves or acts, despite the presence of unwanted pain-related internal thoughts or feelings.

## *Treatment for Chronic Pain*

When pain has persisted beyond acute into chronic pain and is negatively impacting on a person’s functioning and quality of life, it is recommended that they are offered access to a Pain Management Program (PMP; NICE guidelines, 2018). These are group-based therapeutic sessions and can be underpinned by a variety of psychological models. They provide individuals with the opportunity to meet other people living with long-term pain. Given the breadth of cognitive, social and emotional aspects in chronic pain, a range of psychological (group based or one-to-one interventions) are now often used. For example, Cognitive Behavioural Treatments (CBTs) have been trialled extensively and found to reduce pain and improve physical and emotional functioning in chronic pain (McCracken & Turk, 2002; Morley, Eccleston & Williams, 1999). However, other studies have also shown that these effect sizes are modest (Williams, Eccleston & Morley, 2012). A newer type of Cognitive Behavioural Therapy (CBT) called Acceptance and Commitment Therapy (ACT; Hayes, Strosahl & Wilson, 1999) is now emerging amongst psychological approaches as the treatment of choice.

There is a growing evidence base for ACT for improving outcomes in chronic pain. (Dahl, Wilson & Nilsson, 2004; Vowles, McCracken & O’Brien, 2011; Wicksell, Melin, Lekander & Olsson, 2009). Two recent meta-analyses show that ACT appears to be effective as a treatment for chronic pain, with beneficial effects on physical and emotional functioning (Hann & McCracken, 2014; Veehof, Trompetter, Bohlmeijer & Schreurs, 2016). A study conducted by Wagener and Zettle (2011) highlighted a need to distinguish between CBT and ACT based treatments effects. Findings showed that participants in an acceptance-based condition progressed further on the Perceived-Threat Behavioural Approach Test (PT-BAT; Cochrane, Barnes-Holmes & Barnes-Holmes, 2008) in comparison to a control-based group intervention. Wagener and Zettle (2011) attributed these changes to the acceptance-based groups focus on relating to difficult thoughts and feelings, as opposed to using control-based techniques which try to minimise unwanted thoughts. However, it is highlighted that this study does not tell us which specific mechanisms of change in the acceptance-group can be attributed to the increased efficacy of ACT therapy. Although these results are preliminary, they do provide initial support and more importantly highlight a need for further research surrounding ACT process measures. Firstly however, I will now describe in more detail what ACT therapy is and its theoretical foundations.

# ACT Model of Chronic Pain

ACT is underpinned by the behavioural theory of human language and cognition (RFT; Hayes, Barnes-Holmes, & Roche, 2001) which in turn rests on the philosophical foundations of functional contextualism (Biglan & Hayes, 1996; Hayes, Hayes, Reese, & Sarbin 1993). This way of thinking views the context in which behaviours occur as important. Our thoughts and feelings are considered to be *"...ongoing actions of the whole organism interacting in and with historically and situationally defined contexts"* (Hayes, Luoma, Bond, Masuda, & Lillis, 2006, p.5). This worldview focuses on the 'function' or purpose of behaviours and includes an a-ontological stance, where a therapist does not seek to align perception or behaviour with an external 'truth' or reality, but rather to help a participant to align their behaviour with their own overarching goals and values. Ergo, here, and in ACT, "true" is what achieves the goals of the analysis.

These underlying theories have informed the focus of ACT, which aims to improve outcomes (e.g. quality of life, psychological distress and pain interference) by targeting an earlier mentioned process called psychological flexibility. Psychological flexibility can be defined as: "*ability to contact the present moment more fully as a conscious human being, and to change or persist in behaviour when doing so serves valued ends"* (Hayes et al., 2006, p.7). ACT is clearly a process-orientated therapy which aims at increasing psychological flexibility through six core interrelated processes: cognitive defusion, acceptance, present moment awareness, contact with self-as-context, moving towards values and making committed actions (as illustrated in Figure 1.1).

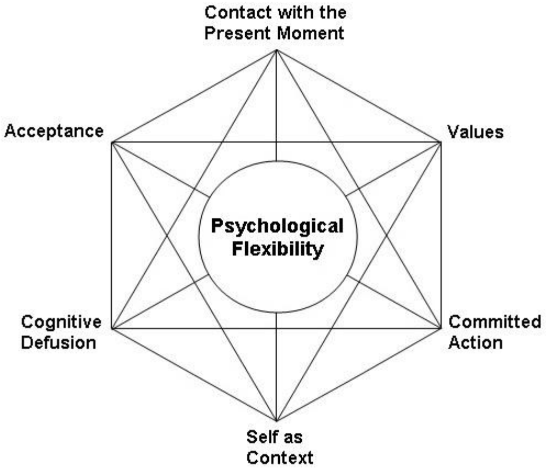


**Figure 1.1:** 'Hexaflex' model of core psychological processes strengthened through ACT *(taken from Hayes et al., 2006).*

The Hexaflex model conceptualises psychological flexibility as a product of these distinct yet related sub-processes. Demonstrating psychological flexibility thus tends to involve accepting painful feelings and thoughts, with a focus on opportunities that may arise in the current situation rather than being lost in the past, whilst actions and behaviour become more in line with what really matters (values) instead of attempts to control pain (McCracken & Vowles, 2014). The six key psychological skills which make up psychological flexibility can be divided into two overlapping processes, with mindfulness and acceptance based constructs on one side and behaviour change and committed action constructs on the other (Hayes et al., 2006).

In relation to chronic pain, *Acceptance* is targeted through establishing willingness to experience pain. Indeed, the evidence suggests that patients showing a willingness to have pain, and engage in activities regardless, leads to better functioning (McCracken, Vowles & Eccleston, 2005). The individual learns to stay with the unpleasant and unwanted experience of pain without attempting to avoid or control it, whilst negative thoughts associated with pain are targeted through exposure work rather than trying to change or stop them (Dahl, Wilson & Nilsson, 2004). The process of *Defusion* brings awareness to thought processes 'as thoughts'. Whilst we can acknowledge their power, defusion creates a space for stepping back from these pain related thoughts and experiences. Establishing an individual's values whilst encouraging movement towards them even with pain is achieved through increasing values-based activities or *committed actions. Self as context,* or 'self as observer' refers to the ability to become familiar with and separate from their conceptualised self. In an individual with chronic pain, this may involve attaining a perspective in which they are not defined by their own thoughts and feelings – and can shift perspective on themselves and their experiences. Through becoming more aware of their experiences, and consequently creating defusion from negative thoughts and feelings about their chronic pain, the individual can begin to view these as transient events that do not reflect their 'true self' or reality (McCracken, Gutiérrez-Martínez, & Smyth, 2013). Lastly, the process of *present moment awareness* refers to observation of experiences, of physical pain sensations, thoughts and emotions attached to pain. The mindfulness-based methods used as a treatment for chronic pain try to increase present focused awareness and encourage the individual to not react in unhelpful ways to pain in the body or related emotions and thoughts (McCracken, Gauntlett-Gilbert, & Vowles, 2007).

Whilst the Cognitive Behavioural Therapy (CBT) approach shares common active components with ACT, for example (behavioural activation) its emphasis is on symptom reduction and associating perception with an external reality. ACT instead, focuses on aligning behaviour with overarching values. There are many theoretical differences between ACT and CBT accounts of chronic pain. As CBT purports that pain is mediated by our cognitions, it would imply that pain could be reduced by changing these inaccurate beliefs. ACT focuses on acceptance, the unavoidable nature of pain and encourages individuals to learn skills in living with the pain and having a full life (McAndrews, Richardson & Stopa, 2018). It has been suggested that ACT is particularly relevant to treating chronic pain because unlike traditional CBT it does not try to reduce behaviours linked to pain. Instead, the acceptance of pain determines improvements in well-being (Reneman, Dijkstra, Geertzen & Dijkstra, 2010). In contrast, with its focus on changing the content of an individual's thoughts and beliefs regarding pain and its consequences (Vlaeyen & Linton, 2000), it is more difficult to determine which components of CBT transpose from one treatment to another. Although the effectiveness of CBT for treating chronic pain is well evidenced (Vowles & McCracken, 2008) it is unclear which specific elements of this approach are agents of change (Morley, 2004). Yet, as ACT was developed from CBT they do contain very similar methods in their approach. A large randomised control trial conducted by Wetherell et al. (2011) found equivalence in each of these therapies effectiveness in improving pain interference, depression and pain-related anxiety in people with chronic pain. Although, it was noted that ACT was rated as more satisfactory and enjoyable by participants. To help us better understand the mechanisms of change underlying effective treatment methods it may be more beneficial for future research to focus on exploring processes-orientated therapies such as ACT (McCracken & Vowles, 2014).

## ACT Process Variables

McCracken and Vowles (2014) highlight the importance of looking at the theory underlying ACT. Understanding therapeutic mechanisms of change in ACT interventions will provide information on how treatment works, in this process-orientated therapy. They define these process variables as: *'directly targeted, theoretically based, psychological elements deemed to affect improvements in treatment outcome variables'*. Research investigating ACT processes will establish whether interventions for chronic pain work via the appropriate 'mechanism of action'. Understanding the treatment processes and the route by which changes are made will help to further refine the ACT theory itself. Furthermore, clinicians will be better equipped to adapt interventions and improve treatment efficacy if we can determine whether treatment effects are due to the intervention itself, as opposed to other non-specific variables such as emotional reactions and common sense, which are often misleading (Schulte & Eifert, 2002).

The accurate recording of ACT process variables will help to discern ACT from other intervention models. For example, to discriminate from CBT, which is theorised to work not via improving psychological flexibility but via changes in cognitive distortion/coping. Hayes, Hope and Hayes (2007) note the importance of recording process variables in psychological intervention studies as these can give us insight into processes involved in recovery and drop out from treatment. Focusing on processes of change is important to determine the active ingredients in therapy. Clinical decision making depends on our knowledge of which components are needed for treatment and in what order they are most effectively delivered (Levin, Herbert & Forman, 2017). Table 1.1 outlines the six core ACT processes which make up psychological flexibility.

**Table 1.1 Brief Description of the 6 core ACT processes (adapted from Harris, 2009)**

| Process | Description |
|---|---|
| Acceptance | Opening up and willingness to make room for painful private experiences (thoughts, feelings, emotions, sensations or urges). |
| Cognitive defusion | Stepping back and taking perspective from thoughts, images and memories. Noticing them for what they are instead of getting caught up in them. |
| Contacting the present moment | Being psychologically present; connecting and engaging with what is happening in the present moment. |
| Self-as-context | Awareness of the "observing self," or self-as-context. |
| Values | Describe how we want to behave, our life directions-in other words; what we want our life to stand for. |
| Committed action | Behaving in line with valued directions- the actions taken to create a full and meaningful life. |

## Evaluating the validity and reliability of ACT questionnaires

Questionnaires have been developed which capture ACT processes. These can let us know whether our interventions are leading to changes in the targeted processes – for example in psychological flexibility. Thus, they are very useful in clinical and research work. Investigating the degree to which responses on these measures are indeed reflective of these targeted processes, is an important factor in capturing change. In order to establish whether ACT measures do accurately record the ACT process variables, we must be sure of their *reliability*. Reliability concerns the consistency of a measurement tool, i.e. the extent to which the measurement of a phenomenon is stable and repeatable (Cohen, & Swerdlik, 1999). For example, a questionnaire designed to measure acceptance of pain, the Chronic Pain Acceptance Questionnaire- Revised 8 item version (CPAQ-8), has been shown to demonstrate good scale score reliability (internal consistency) with Cronbach's alphas ranging from .77 to .89 for subscales of Activity engagement and Pain Willingness, respectively (Fish, McGuire, Hogan, Morrison, & Stewart, 2010). Here, items assessing each scale reliably yielded similar scores. A revised version of the Acceptance and Action Questionnaire (AAQ-II, Bond et al., 2011) was also found to be a reliable measure of psychological flexibility with a mean alpha coefficient of .84 and a 12 month test-retest reliability of .79. This study also demonstrated the AAQ-II to hold discriminant validity; factor analysis revealed it was not significantly associated with 'theoretically distinct constructs'.

*Validity* must also be established in order to say whether ACT measures do record the processes they were originally designed to capture. Validity refers to the extent to which a test measures what it claims to measure and ultimately determines how useful inferences made from scores are (Cohen & Swerdlik, 1999). A process to evaluate whether an ACT questionnaire measures what it is designed to would be to establish its *construct validity* which has been defined as - *"a judgement about the appropriateness of inferences drawn from test scores regarding individual standings on a variable"* (Cohen & Swerdlik, 1999, p. 197). For example, analysis has supported the construct validity of the Chronic Pain Values Inventory (CPVI) which significantly correlated with other measures of avoidance and acceptance of pain (McCracken & Yang, 2006). Less promising findings for ACT questionnaire validity were found in a systematic review conducted by Reneman et al. (2010). They applied psychometric quality criteria (see Table 1.2; Terwee et al., 2007) to 23 individual studies using ACT questionnaires designed to measure the construct acceptance of chronic pain. Four ACT questionnaires were examined, including the CPAQ. Findings showed that none of these questionnaires met all identified criteria for internal consistency, criterion validity or construct validity. It was proposed that an active involvement from the target population (individuals who are experiencing chronic pain) is required to create valid questionnaire items and limited evidence of this was found. Consulting the target population can be useful in ensuring *content validity*- which is defined as the extent to which items on a measure *"reflect the content universe to which the instrument will be generalized"* (Straub, Boudreau & Gefen, 2004). It tells us how well the individual questionnaire items create the construct and thus measure the behaviour. Construct validity however, refers to the test as a whole- it concerns the extent to which overall questionnaire scores relate to other measures also consistent with the theoretically driven hypotheses surrounding the measured constructs. In all but one of the 23 studies reviewed by Reneman et al. (2010) was evidence presented which could be used to confirm the ACT questionnaires' content validity. Therefore, relying solely on the interpretation by health professionals appears to be problematic in the research base to date. As already touched upon, a measure can only assess the construct it is deemed to measure if the questionnaire is understood by the respondent (Strauss & Smith, 2009).

The valid measurement of processes has significant implications for continued ACT research and is part of the explicit research agenda for ACT (Hayes, Barnes-Holmes & Wilson, 2012). The authors of this review article refer to truth as 'pragmatic', coming from a contextual science perspective, as opposed to classical test theory which makes (ontological) assumptions about test validity. They purport that aim of ACT research (and the functional contextualism philosophy underpinning it) is to explore how we might 'predict-and-influence' the processes of interest, e.g. defusion, as opposed to explaining how constructs are represented internally. Hayes, et al. (2012) note that the most important aspect of a measure is in its applicability to treatment and research. In other words, can it guide clinical practice and future research? Improved outcomes rely on the targeting of *"behavioral processes suggested by contextual principles and models*" and understanding which aspects of the intervention work and how (Hayes, et al., 2012). This improvement of care for people experiencing chronic pain using ACT- relies on the use of valid measures which are able to accurately reflect change as therapy progresses. Establishing the content validity of measures is clinically important as without attaining a sense of how questionnaire items are understood by the target population, we cannot be certain that the domains of interest (ACT processes) are being accurately sampled by the questionnaire items. Assessing how accurately a measure represents change is also important for future research investigating processes of change and will further refine the theory itself.

A predominant focus of the research literature to date has been on issues with the Acceptance and Action Questionnaire (AAQ-II; Bond et al., 2011) (Francis, Dawson, & Golijani-Moghaddam, 2016; Gámez, Chimielewski, Kotiv, Ruggero, & Watson, 2011; Gámez et al., 2014; Rochefort, Baldwin, & Chmielewski, 2018). The AAQ-II has been criticised for overly focusing on single sub-processes (acceptance and defusion processes). Although purported as a unidimensional measure of psychological inflexibility (Bond et al., 2011), Francis et al. (2016) suggest that the AAQ-II focuses on experiential avoidance and fusion process and does not capture changes in contact with the present moment, values or committed action. Furthermore, other aspects of psychological flexibility are not examined through the questionnaire items at all (self as context or committed actions). It was suggested that this over emphasis on developing ACT questionnaires which measure single sub-processes may be problematic as although useful for exploring the impact of the individual ACT processes, it does not capture the process of psychological flexibility as a whole. This could cause problems in data interpretation as we do not know to what degree different ACT questionnaires may overlap in terms of the constructs they are measuring. Although the Francis et al. (2016) study raised important questions surrounding the content validity of the AAQ-II it did not empirically test this, rather it used an exploratory factor analysis to develop a new 'general measure of ACT processes'- a 23-item global measure of ACT processes: the Comprehensive assessment of ACT processes (CompACT). On this measure, subscale scores are obtained for each of the three 'dyadic' main processes: 'Openness to Experience', 'Behavioural Awareness' and 'Valued Action', which correspond with definitions of psychological flexibility (Hayes et al., 2006). The CompACT also obtains a total score to indicate psychological flexibility as a whole.

The discriminant validity of the AAQ-II was also criticised by Wolfgast (2014) for certain items overlapping with distress outcome variables, e.g. item 2 'I'm afraid of my feelings' could be interpreted as an actual fear of feelings as opposed to an indication of experiential avoidance. This study also used an exploratory factor analysis on individual AAQ-II items amongst items designed specifically to measure distress and acceptance/non-acceptance. As some of the items were found to not load onto their intended constructs of measurement it was argued that this makes it difficult to determine whether ACT questionnaire responses indicate an individual's level of psychological inflexibility *or* another process, for example distress or worry. This example demonstrates why poorly defined items and hence issues with the questionnaire's content validity may be problematic and ultimately impact on what clinicians can interpret from research trials or clinical therapeutic practice. However, neither of these studies, nor any found amongst the literature to date, have empirically tested the content validity of these questionnaires. There remains a gap in the existing research to examine the content validity of ACT measures in a robust way, as the current study will go on to investigate. An additional factor contributing towards the poor content and face validity of ACT questionnaires may also be attributed to the esoteric concepts and complicated language used in their items. However, as previously mentioned there has been limited evidence in the literature to date demonstrating the involvement of the target population in developing or testing the content validity of these items- i.e. the extent to which respondents actually understand them. Therefore, it remains unknown whether socialisation to the ACT model might impact on how an individual might respond to the questionnaires as they progress through therapy.

The research literature to date has largely focused on examining changes in psychological flexibility following ACT/ACT-based treatment for chronic pain using self-report measures. It has been purported that this reliance on self-report measures may be problematic as they are susceptible to demand characteristics (Levin, Herbert & Forman, 2017). Measurement error is potentially compounded by the social desirability biases of respondents who may want to provide the researcher with positive answers. Furthermore, these self-report measures also require a certain degree of respondent awareness to accurately convey levels on the ACT processes intended (Levin, et al., 2017). To establish a change in ACT processes a questionnaire must be administered before the intervention and then at least after the intervention. As alluded to earlier, difficulties in accurately measuring change occur when there are potential changes in responding once participants are socialised to the model. Scott, Hann and McCracken (2016) suggest that future studies would benefit from utilising other assessment methods as opposed to relying only on self-report measures, for example, a measurement of directly observable patterns of behaviour following treatment. However, in light of the current reliance on self-report measures amongst the existing evidence base, issues related to the validity of these measures will now be highlighted.

What can be inferred from the ACT process research (e.g. on whether psychological inflexibility is a core pathological process) may also be limited due to issues with the questionnaires themselves. Historically, the validation of measures of psychological constructs was problematic due to a limited foundation on which to build future knowledge. Strauss and Smith (2009) noted that an established knowledge base is required in order to validate scores on measures; i.e. to test whether the measure of a psychological construct relates to measures of other constructs defined by the theory. The idea being that theory driven research has a reciprocal influence on the developing knowledge base and also provides a way of validating measures which can further test this theory. As mentioned earlier, this process of theory testing is otherwise known as *construct validity*. They proposed that "*to validate one's claim that scores on a measure play a certain role in a network of psychological processes, one needs valid measures of the different components of the specified process.*" (Strauss & Smith, 2009, p.2). According to Messick (1995) the overarching theory of construct validity encompasses numerous forms of evidence to support it which includes theory to, "*support the adequacy and appropriateness of interpretations and actions on the basis of test scores…*" (Messick 1995, p. 174). *Content validity* is included in this and is important in determining whether a questionnaire measures what the author originally intended it to measure.

## *Assessing content validity in ACT process measures*

Exploring aspects of ACT questionnaire content validity will give us an indication of their accuracy and may also help us to improve them. Yet, as mentioned there has been limited research exploring individual's experience of these questionnaires. Exploring this component of validity may help us to understand what their responses might be based upon and how this may impact the validity of the questionnaire. Terwee et al. (2007) developed quality criteria for use in systematic reviews for reviewing health status questionnaires, as shown in Table 1.2.

**Table 1.2: Measure content validity criteria based on Terwee et al. (2007)**

| Aspect of content validity | Definition |
|---|---|
| *Measurement aim* | There should be a clearly stated aim for the measure e.g. is it discriminative, evaluative, or predictive? This is crucial because different items are valid for different aims. For example, values clarification vs assessing changes in values. |
| *Target Population* | The population for which the measure was developed is stated so that the relevance and comprehensiveness of the items can be assessed. For example, is the measure for a specific clinical population such as |

| | |
|---|---|
| | chronic pain (as in the CPAQ) or does it target the more general population? (as in the AAQ). This is important as allows the user to assess whether the measure is applicable to the respondent. |
| *Intended concepts are being measured* | So that the usefulness of a measure for its intended purpose can be assessed, the questionnaire authors must have provided a clear framework detailing the concepts to be measured. For example, the particular ACT components of psychological flexibility it intends to measure. |
| *Item selection/ reduction* | The population for which the measure targets should be involved in the design of and ultimately the selection or exclusion of items on it. For example, through the use of pilot studies to investigate readability and comprehension and relevance of items to that. This aspect of content validity is not considered crucial for content validity as it is acknowledged that a comprehensive set of items can still be achieved without item reduction. |
| *Interpretability of the items.* | Items should be short and simple and not contain difficult words or jargon terms. This will ensure the item is readable. Item should also not consist of two questions at the same time to avoid confusion. |

To summarise, the inaccurate measurement of psychological flexibility processes has implications for what we can infer from clinical trials of ACT for chronic pain. Without being certain of how accurately changes in psychological flexibility are measured we cannot confidently attribute this to treatment effects, or adapt interventions accordingly. Indeed, Wolgast (2014) argue that previous findings linking psychological flexibility to improved psychological well-being may actually be a product of the poor operationalisation and measurement of ACT sub-processes. The research summarised here raises concerns that current ACT questionnaires do not accurately measure what they intend to measure, and consequently may not truly reflect changes in process variables throughout treatment. Studies such as the ones conducted by Wolfgast (2014) and Francis et al. (2016) highlight the focus of existing research literature on using quantitative methodology to assess validity. Although they provide evidence for issues with the construct validity of psychological flexibility measures and what is actually being measured, the potential issues with content validity are only rarely commented on. Another example from the literature is a factor analysis conducted by Rochefort, Baldwin, and Chmielewski (2018), who examined the construct validity of the AAQ-II through assessing its convergent validity and discriminant validity. Their findings suggested that this process measure was actually more correlated with other established measures of neuroticism. Although the suboptimal content validity of this process questionnaire is implied this was not examined empirically.

It is these concerns which lead me to the present study which hopes to address a need for further research exploring issues with content validity with certain items on ACT questionnaires, in relation to how these are understood by those completing the measure. This will in turn allow us to examine the extent to which the constructs of interest (psychological flexibility variables) are adequately sampled by the items in the ACT questionnaires. It is not the aim of the present study to validate the psychological constructs which make up psychological flexibility. Previous literature provides evidence for the existence of clearly established psychological flexibility processes (Gloster, Klotsche, Chaker, Hummel, & Hoyer, 2011; McCracken & Morley, 2014). Instead, as described, the focus will be on establishing the content validity of ACT measures. This study will explicitly test the content validity of ACT process measures using cognitive interviewing methodology which will now be described.

# Methodology to explore problems with ACT questionnaire validity

## *Cognitive Interviewing methodology*

Cognitive interviewing has been used extensively to help identify problems experienced by respondents when answering questionnaires (Ericsson & Simon, 1998; Schuman, 1966; Willis, DeMaio & Harris-Kojetin, 1999). This technique asks respondents to complete individual questionnaires items whilst verbalising their thought processes as they answer the question. Beatty and Willis (2007) defined this method as *"the administration of draft survey questions while collecting additional verbal information about the survey responses, which is used to evaluate the quality of the response or to help determine whether the question is generating the information that its author intends."* (Beatty & Willis, 2007, p.1). The cognitive interviewing approach was developed in the 1980s and has been used in the design of survey questions through evaluating sources of response error. The theory underpinning this approach attributes numerous processes to respondents' successful understanding and completion of questionnaire items (Tourangeau, 1984) - namely problems with: comprehension of questions, retrieval of relevant information, and decision and response processes in arriving at the question answer. Cognitive interviewing focuses on revealing these hidden cognitive processes used by respondents when answering questions.

Two variants of cognitive interviewing exist: ‘thinking-aloud’ and ‘probing’. According to Beatty and Willis (2007) the former involves minimal intervention as the respondent arrives at their answer whilst the latter requires the interviewer to ‘guide’ the process, with the use of direct questions regarding answers provided. The probing technique however, has been subject to criticism due to the confounding influence it has on the stream of thoughts respondents experience as they attempt to answer question (Conrad, Blair & Tracy, 2000). The ‘pure’ think-aloud method however, has been considered less susceptible to biases as data is collected at the very point respondents consciously think through and verbalise their answers (Van der Veer, Hak & Jansen, 2002, as cited in Beatty & Willis, 2007). This method may produce a more accurate representation of thought processes as it does not rely on the respondent holding information in their memory whilst they respond to a ‘probe’. Instead they attend to and verbalise their thought processes as they come to mind. Although limited, there are examples of previous research utilising the think-aloud methodology to reveal more about respondent’s level of understanding, indicating where errors in answering the questions are likely to occur and identify problematic questions on psychological measures. For example, Van Oort, Schroder and French (2011) used this variant of cognitive interviewing to investigate the difficulties patients experienced as they completed the Brief Illness Perception Questionnaire (Brief IPQ). Findings showed evidence of questions being misinterpreted (responses showed that a different question was being answered), suggesting problems with the content validity of this measure.

### *What cognitive interviewing adds to the assessment of ACT questionnaire validity*

This method may help contribute towards our understanding of ACT questionnaire content validity as it examines the extent to which the constructs of interest (psychological flexibility variables) are interpreted as intended by respondents. Individuals are required to verbalise their current thoughts, without providing rationale or reasoning for them. Drennan (2003) suggests using this method for potentially complex or sensitive questions and for use in clinical groups where questionnaire completion has been problematic. This could involve issues with comprehension, making decisions, recall, or being able to respond in answering the question effectively.  According to the literature to date there is no evidence to suggest that cognitive interviewing methods have been used extensively to evaluate ACT questionnaires in particular. The ‘think-aloud’ method is able examine various issues in responses to questionnaire items as it accesses the underlying cognitive processes. As summarised by Drennan (2003) it can provide insight into the degree to which respondents comprehend the words and concepts used in items. Secondly, it can tell us whether the respondent understands the question i.e. assesses coherence with concepts under study. Lastly, it can be used to tell us what words or phrasing lends to problems in understanding. Therefore, we can know whether the measures access the constructs they deem to measure -i.e. assess the questionnaires *content validity*. This will indicate about what can be done to improve their validity. Finally, Conrad and Blair (1996) recognised a need to categorise these problems in understanding into a framework or ‘taxonomy of problem classification’ to avoid the limitations of adopting a purely qualitative approach towards analysing the responses. They proposed a systematic way of objectively analysing this qualitative data from the cognitive interviews which was later developed by Drennan (2003) (and will be described in detail in the present study method section). Different types of difficulties will be explored but in a systematic way e.g. difficulties with understanding and response formatting.

## Summary, Research Question and Aims

Clinicians need to understand more about the process of change in a psychological intervention to enable them to adapt treatment methods during therapy and develop more effective approaches. At the moment, we rely on self-report questionnaires to assess these processes. We know little about how valid such questionnaires are, particularly those focusing on ACT. Existing research has provided partial evidence to suggest that ACT questionnaires meet some criteria for psychometric quality. For example, the review conducted by Reneman et al. (2010) on a number of process measures found that cumulatively, the CPAQ held the strongest psychometric properties, with a number of studies reporting good internal consistency and construct validity. However, as they concluded, no measures met all the psychometric criteria applied. More importantly, studies using this quantitative methodology to assess validity do not tell us *how* individuals understand the items in the questionnaires, and how this influences their response. Crucially, we don’t know whether they interpret the questions in the way intended by the questionnaire design. There appears to be no existing research *directly* investigating how individuals interpret then respond to ACT questionnaires. Therefore, we remain unclear as to whether the questionnaires demonstrate content validity. Research is needed to explore patient’s understanding of the questions so that conclusions can be made around the ability of the ACT questionnaire to accurately reflect the intended process variables. In order to address the gaps identified above, this study aims to use a mixed methodology to explore the validity of a variety of ACT process variable questionnaires. It is hoped that the findings from this study will provide a more comprehensive understanding of where problems lie in the use of ACT questions and in turn influence the ability for future research to evaluate the processes occurring during treatment. As a result, clinical psychologists working within chronic pain services may be better equipped to effectively capture change in clients undergoing ACT for chronic pain.

**Overarching Aim:** To investigate the content validity of commonly used ACT questionnaires in individuals experiencing chronic pain.

**Research Questions:** Do commonly used ACT questionnaires capture their intended processes, in people experiencing chronic pain? More specifically: do individuals understand the questionnaire items as intended? For refinement of questionnaires: can we establish whether there are any particular patterns in misunderstanding of questionnaire items?

**There were two parts to the study:**

**Study Part One-** To establish the most frequently used ACT process measures in chronic pain clinical research and practice.

**Study Part Two-** A questionnaire validation study utilising cognitive interviewing, which aimed to evaluate how individuals make sense of the most used questionnaires as they complete them.

# PART ONE- REVIEWING RESEARCH AND CLINICAL USE OF ACT QUESTIONNAIRES

## Chapters overview

Before embarking on the questionnaire validation component of the study, use of ACT questionnaires in research trials and the clinical context was first established. We wanted to confirm which ACT process questionnaires are used most frequently in chronic pain research and clinical practice. This was so we could focus our Cognitive Interviewing assessment of content validity on the most common or influential questionnaires in the field. Therefore, the following three chapters: 2, 3 and 4 comprise a search across ACT intervention research and an online survey amongst practicing ACT clinicians.

## INTRODUCTION

### Identifying most commonly used ACT questionnaires

As described in Chapter 1, questionnaires have been developed to capture sub-processes of Psychological Flexibility (PF) in chronic pain. For example, the Chronic Pain Values Inventory (CPVI) is a 12-item measure used to determine individuals' values and assess to what degree they are successfully following and living by those values (McCracken & Yang, 2006). The individual components making up psychological flexibility have been researched using ACT questionnaires (Hann & McCracken, 2014). However, the use of these specific questionnaires has varied across ACT intervention research according to several meta-analyses (Hann & McCracken, 2014; Hughes, Clark, Colclough, Dale & McMillan, 2017; Ost, 2014; Simpson, Mars & Esteves, 2017; Veehof, 2011; Veehof 2016). The Randomised Control Trials (RCTs) reviewed in these studies identified which ACT process measures are most frequently used in the research context, however they do not establish clinical usage of ACT questionnaires in recording changes following therapy.

# Aims of the research review and clinician survey

In order to establish the frequency of use for ACT process questionnaires in both the research literature and clinical practice, this preliminary study was conducted in two parts:

> ***Study 1: Systematic review*** - A systematic review to collate the RCT studies or existing systematic reviews exploring the use of ACT for chronic pain. This aimed to establish the ACT process measures most frequently used in clinical research.
>
> ***Study 2: Online questionnaire study*** - An online questionnaire of clinicians who use ACT for chronic pain in clinical practice to enable us to see which are used most often in clinical practice.

In addition to this, we also took the opportunity via the online questionnaire to obtain feedback on clinicians' experiences of using these questionnaires in clinical practice. To date there has been no large-scale attempt to investigate clinicians' experiences of ACT processes questionnaires. We wanted to establish whether clinicians notice problems in practice and whether there are patterns in the emergent strengths and weaknesses of existing measures which are considered most useful or problematic.

The findings from these two studies were then used to develop an item pool to explore questionnaire validity in the main study.

# CHAPTER TWO: (STUDY 1) Literature Review using Systematic Search

## METHOD

### Design

This study comprised a literature review using a systematic search. This initial literature review informed the selection of questionnaires to be included in the clinician survey.

### Inclusion and exclusion criteria

The literature search to establish usage of ACT questionnaires across empirical research considered studies for inclusion that a) were existing meta-analysis or systematic research reviews; b) were RCTs; c) assessed ACT interventions for the treatment of chronic pain using ACT outcome measures; d) included adult populations (≥ 18 years) and e) were published in the English language.

### Procedure

The literature review was performed in three online databases (EMBASE, PsychINFO and MEDLINE) and Google Scholar with searches from the earliest available date until the end of January 2017. Abstracts of studies were read and if suitable, a full-text copy was retrieved and assessed against the inclusion/exclusion criteria described above. The review used a subject and text word strategy with (Chronic pain) and (ACT) and (Acceptance and Commitment Therapy) and (Randomised controlled trial) as the primary search terms, as shown in Table 2.1 (see Appendix D for full search strategy terms and screenshots).  Furthermore, the reference lists and cited articles of included studies were examined for additional potentially eligible studies. Systematic reviews were checked for overlap with separately identified RCT studies so that an accurate total for each ACT outcome measure could be calculated. The literature search was repeated on write-up (April 2019) so that any recently added studies could also be acknowledged. A PRISMA flow chart (Moher, Liberati, Tetzlaff & Altman, 2009) was used to document the various stages of the most recent literature search, as illustrated in figure 2.1.

**Table 2.1 Example search strategy (MEDLINE -20/04/19)**

| | Search terms | Search result: number of papers |
|---|---|---|
| 1 | Chronic pain | 11667 |
| 2 | Chronic pain.mp | 32922 |
| 3 | Acceptance and commitment therapy.mp | 527 |
| 4 | Acceptance commitment therapy.mp | 21 |
| 5 | ACT.mp | 231540 |
| 6 | Acceptance and commitment therapy/ or acceptance | 92608 |
| 7 | Random* controlled trials.mp | 152662 |
| 8 | Random* controlled trial.mp | 606351 |
| 9 | Meta analysis.mp or meta-analysis | 135327 |
| 10 | RCT.mp | 132695 |
| 11 | 1 OR 2 | 32922 |
| 12 | 3 OR 4 OR 5 OR 6 | 322689 |
| 13 | 7 OR 8 OR 9 OR 10 | 726341 |
| 14 | 11 AND 12 AND 13 | 132 |
| 15 | Limit 14 to English language | 127 |

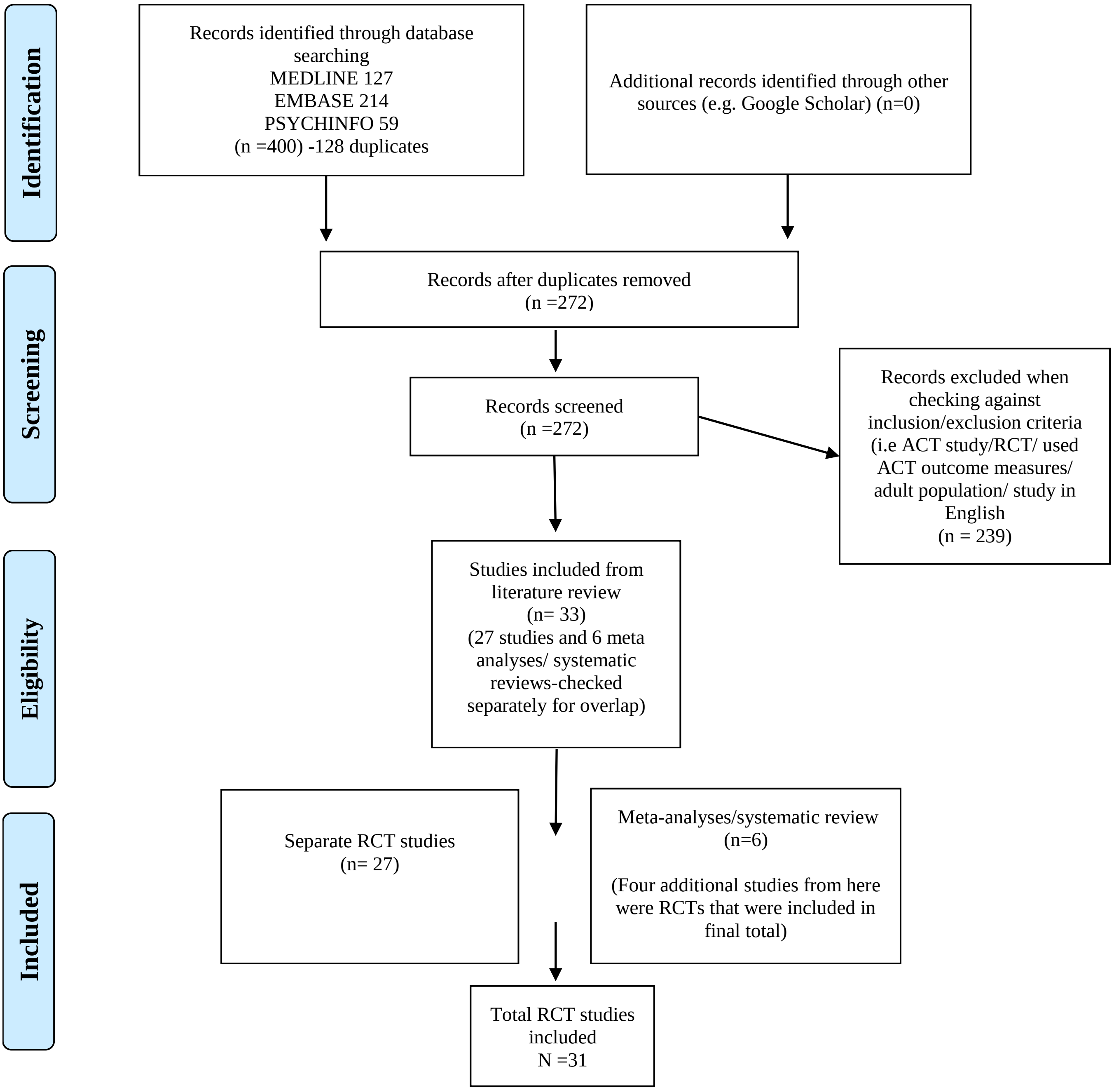


**Figure 2.1 PRISMA Flow diagram to demonstrate search strategy**

# RESULTS

The literature search informing the selection of ACT measures for the main study was originally conducted during January 2017. The table below demonstrates total frequencies of use for each ACT measure (Table 2.2).

**Table 2.2 Frequency of ACT outcome measures used in RCT studies (January 2017)**

| | Studies included | Measure used* | Measure frequency | |
|---|---|---|---|---|
| **RCTs identified in literature review** | Wicksell et al., 2008a; | PIPS | *CPAQ/ (R, 8)* | *= 11* |
| | Johnston et al., 2010; | CPAQ/CPVI | *PIPS* | *= 5* |
| | Wetherall et al., 2011; | CPAQ-R | *AAQ-II* | *= 3* |
| | Thorsell et al., 2011; | CPAQ | *CPVI* | *= 1* |
| | Wicksell et al., 2013; | PIPS | *VLQ* | *= 1* |
| | Buhrman et al., 2013; | CPAQ | *ELS* | *= 1* |
| | McCracken et al., 2013a; | CPAQ/AAQ-II | *FFMQ* | *= 1* |
| | Dahl et al., 2004; | VLQ | | |
| | Lin et al., 2015; | CPAQ/ AAQ-II | | |
| | Kemani et al., 2016; | PIPS | | |
| | Cederberg et al., 2016; | CPAQ | | |
| | Trompetter et al., 2015a; | PIPS | | |
| | Kemani et al., 2015; | CPAQ | | |
| | Trompetter et al., 2015b; | PIPS / ELS/ FFMQ | | |
| | Hayes et al., 2014; | CPAQ-8 | | |
| | McCracken et al., 2014a; | CPAQ/ AAQ-II | | |
| | Sullivan et al., 2012. | CPAQ | | |
| **Additional RCTs identified from systematic reviews within the literature** | Steiner et al., 2013; | CPVI | *CPAQ* | *= 1* |
| | Luciano et al., 2014; | CPAQ | *PIPS* | *= 1* |
| | Plumb Vilardaga, 2012; | PIPS/FFMQ/BE | *AAQ* | *= 1* |
| | Alonso et al., 2013. | AAQ/ CPVI | *FFMQ* | *=1* |
| | | | *BE* | *=1* |
| | | | *CPVI* | *= 2* |
| **TOTALS** | | | ***CPAQ/ (R, 8)*** | ***= 12*** |
| | | | ***PIPS*** | ***= 6*** |
| | | | ***AAQ-I/II*** | ***= 4*** |
| | | | ***CPVI*** | ***= 3*** |
| | | | *FFMQ* | *=2* |
| | | | *VLQ* | *= 1* |
| | | | *ELS* | *= 1* |
| | | | *BE* | *=1* |

**Measures used: PIPS = Psychological Inflexibility in Pain scale; CPAQ = Chronic Pain Acceptance Questionnaire; CPAQ-R = Chronic Pain Acceptance Questionnaire- Revised; CPAQ-8 = 8-item Chronic Pain Acceptance Questionnaire; CPVI = Chronic Pain Values Inventory; AAQ-I/II = Acceptance and Action Questionnaire/ version two; VLQ = Valued Living Questionnaire; ELS = Engaged Living Scale; FFMQ = Five Facet Mindfulness Questionnaire; BE= Bulls-Eye Values Assessment.*

The electronic search identified 21 RCTs utilising a range of ACT outcome measures. Four additional studies were not found directly within the literature search and were instead accounted for within the systematic reviews (Veehof et al., 2011; Hann & McCracken, 2014; Ost, 2014; Veehof et al., 2016). The four most frequently used ACT outcome measures included the CPAQ, AAQ-II, PIPS and CPVI.

### *Characteristics of included studies*

The 21 studies included in the literature review evaluated adults who had received ACT or ACT based interventions for the treatment of chronic pain in a predominantly group setting. The studies used male and female participants with varying lengths of pain duration. The studies used patients with varying types of chronic pain for e.g. those with unspecified chronic pain, fibromyalgia, site specific chronic pain such as headache or back pain and rheumatoid arthritis. Study sizes ranged from small pilot studies for RCTs to more large-scale RCT studies.

### *April 2019 Literature search update*

The literature search was repeated in April 2019 to acknowledge any more recently published research. Two additional systematic reviews were found as they were published post the initial search (Hughes, Clark, Colclough, Dale & McMillan, 2017; Simpson, Mars & Esteves, 2017). Ten additional RCTs were also found in the recent review (see Table 2.3). Importantly however, this search did not make any difference to the comparative frequency of use of ACT process questionnaires; the most frequently used measures remained the CPAQ, AAQ-II, PIPS and CPVI. Therefore, the ACT outcome measures selected for the clinician survey remained relevant.

**Table 2.3 Additional ACT outcome measures used in recent RCTs (April 2019)**

| Studies included | Measure used | Measure frequency |
|---|---|---|
| Casey et al., 2018; | CPAQ-8 | CPAQ/ (8) = 8 |
| Lin et al., 2017: | CPAQ/AAQ-II | AAQ-II = 3 |
| Lin et al., 2018; | CPAQ/AAQ-II | PIPS = 1 |
| Herbert et al., 2017; | CPAQ | CPVI = 1 |
| Nes et al., 2017; | CPAQ | VLQ = 1 |
| Wicksell et al., 2010; | PIPS | FFMQ = 1 |
| Dindo et al., 2018; | CPAQ/ CPVI | |
| Probst et al., 2019; | AAQ-II | |
| Simister et al., 2018; | CPAQ/VLQ/FFMQ | |
| Alonso-Fernandez et al., 2016. | CPAQ | |

## Summary

The four most frequently used ACT process questionnaires from RCTs evaluating the impact of ACT on chronic pain included the CPAQ (McCracken, Vowles & Eccleston, 2004), CPVI (McCracken & Yang, 2006), PIPS (Wicksell, Renöfält, Olsson, Bond & Melin, 2008) and the AAQ-II (Bond et al., 2011). This literature review informed the questionnaires to be included in the online survey.

# CHAPTER THREE: (STUDY 2) Clinician survey

## METHOD

## Design

An online survey for ACT clinicians was constructed following the literature review. This asked clinicians about their experiences of using ACT outcome measures and how frequently they used different measures with chronic pain patients (see appendix A). The most frequently used measures from the RCT research were included in the survey. An 'other' option was also included to account for any questionnaires not identified in the literature that may also have been used commonly in a clinical setting. The first part of the survey gathered background information, asking each respondent whether they were a practicing ACT clinician, which country they practiced in, and elicited use of any other therapeutic models (three items). The second part was around use of ACT outcome measures and asked respondents whether they regularly used ACT with clients experiencing chronic pain, which questionnaires they used, how often ('with none of my clients, some of my clients, about half of my clients, most of my clients or with all of my clients'), which questionnaires they thought were useful and which were problematic and how useful in general they found ACT questionnaires (five items) (See Appendix B).

## Ethical clearance

Ethical approval for the online survey was received from the University of Leeds School of Medicine and Research Ethics Committee (SoMREC) on 13/03/2017 (reference number: MREC16-077) (see Appendix A).

## Inclusion and exclusion criteria

Participants for the online survey were considered for inclusion if they: a) were a practicing ACT clinician working with people with chronic pain; b) had the capacity to consent and make one’s own decisions; c) were able to read and write in English (due to questionnaire validation issues).

## Participants

Participants for the clinician review of ACT questionnaire usage were recruited online via a questionnaire created with the Online Surveys system (Online Surveys webpage, 2019; formerly known as the Bristol Online Surveys system-BOS). Based on interest following preliminary contact with the British Pain Society (BPS), the aim was to involve at least 30 participants, however there was no target sample size to answer the research questions – analysis was intended to be purely descriptive and used to inform the main component of the study. The Pain Management Programme (PMP) Special Interest Group (SIG) of the BPS expressed interest in taking part in the study and agreed to send out an email invitation to their members. Other associated forums were also contacted on Twitter, Facebook and the Association for Contextual Behavioural Science (ACBS) online ACT for professionals’ group. As such, clinicians from all over the UK (and potentially globally) could participate.

## Procedure

Recruitment and data collection occurred online from April until June 2017. Participants were recruited via an email advertisement through the British Pain Society, a post on the ACBS members online forums for ‘Pain’ or ‘ACT professionals’ or on Twitter or Facebook (see appendix A). After following the link, they were presented with participant information and consent information which included information related to the purpose of the study and use and reuse of data (see appendix A for combined participant information, consent form and online survey). Participants indicated their consent by checking buttons within the consent form page and entering their e-mail address. They were not able to proceed to the survey without indicating yes to all points of consent. Participants were made aware that they could withdraw at any point during survey completion and their data would not be retained should they not submit the final page of survey.

# RESULTS

Forty-four individuals, all practicing ACT clinicians, completed the online survey. Length of time practicing as an ACT clinician varied from three months to 18 years. Countries included parts of the UK (Wales, England, Northern Ireland and Scotland), Jersey, Cyprus, Canada, USA, Sweden and Brazil. All but one individual regularly used ACT with clients experiencing chronic pain. Eight respondents practiced only ACT; the other 36 also used other therapeutic models in addition, including: CBT, Behavioural Therapy (BT), Cognitive Analytic Therapy (CAT), Motivational Interviewing (MI), Functional analytic psychotherapy (FAP), Schema Therapy, Compassion Focused Therapy (CFT), Eye Movement Desensitization and Reprocessing (EMDR), Interpersonal Therapy (IPT), Solution Focused Therapy, Neuropsychological rehabilitation, Psychodynamic approaches, Mindfulness based approaches, Systemic approaches, Clinical Hypnosis, physical exercise and psychopharmacological treatment. Figure 3.1 displays the frequency of use for each ACT outcome measure amongst the ACT clinicians.

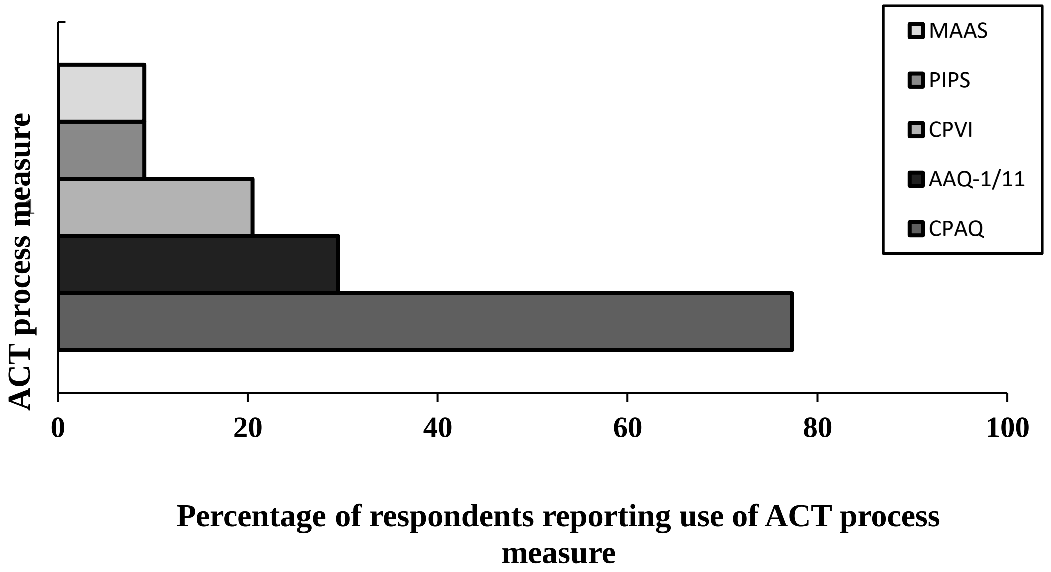


**Figure 3.1 Bar chart to display use of ACT process measures amongst clinicians**

**One extra PIPS measure was incorporated into the final figures; data check revealed one PIPS had been included in the 'other' option in error, thus omitted from question 4(a).*

Results showed that the measure most frequently used in clinical practice was the **CPAQ** (Chronic Pain Acceptance Questionnaire; McCracken et al., 2004) with 77.3% of respondents citing its use. The **CPVI** (Chronic Pain Values Inventory; McCracken & Yang, 2006) and **AAQ-I/II** (Acceptance and Action Questionnaire; Bond et al., 2011; Hayes et al., 2004) were the next most commonly used by 20.5% and 29.5% of respondents, respectively. Clinical use of the **PIPS** (Psychological Inflexibility in Pain Scale; Wicksell et al., 2008b) matched that of the **MAAS** (Mindful Attention Awareness Scale; Brown & Ryan, 2003) at 9.1%. However, the MAAS was included in the survey in error and did not feature at all during the literature search. Therefore, as its inclusion in the survey did not impact on final results, this measure was not included in the main component of the study. Sixteen respondents also selected 'other' in answering the question 'which ACT questionnaires have you used with your chronic pain patients?' This question accounted for the possibility that additional questionnaires not cited in the literature review were also being used clinically. Findings revealed a variety of additional measures (non-ACT based) were also used clinically. These included the **DASS** (Depression Anxiety Stress Scale; Parkitny & McAuley, 2010); **SCS** (Self Compassion Scale; Neff, 2003); **BPI** (Brief Pain Inventory; Cleeland & Ryan, 1994); **BCPI-2** (Brief Pain Coping Inventory; Vowles, McCracken, Sowden & Ashworth, 2014); **MAIA** (The Multidimensional Assessment of Interoceptive Awareness; Mehling, Price, Daubenmier, Acree, Bartmess & Stewart, 2012); **PSEQ** (Pain Self Efficacy Questionnaire; Nicholas, 2007); **HADS** (Hospital Anxiety and Depression Scale; Zigmond & Snaith, 1983); **GAD-7** (tool for assessing Generalised Anxiety Disorder; Spitzer, Kroenke, Williams & Löwe, 2006); **PHQ-9** (Patient Health Questionnaire; Kroenke, Spitzer & Williams, 2001); **ODI** (Oswestry Disability Inventory; Fairbank, Couper, Davies & O'brien, 1980); **PCS** (Pain Catastrophizing Scale; Sullivan, Bishop & Pivik, 1995). Their clinical use was minimal and ranged from only 2.3- 6.8% of the total frequency. Table 3.1 displays ACT /ACT-based process measures used by clinicians in addition to those most commonly cited across empirical RCT studies.

**Table 3.1 Additional measures selected in by clinicians for 'other measures used'**

| Measure | Author | % Frequency of use |
|---|---|---|
| Committed Action Questionnaire (CAQ) | McCracken (2013) | 6.8 |
| Cognitive Fusion Questionnaire (CFQ) | Gillanders et al., (2014) | 6.8 |

| | | |
|---|---|---|
| Experiences Questionnaire (EQ) | Fresco et al., (2007) | 6.8 |
| Valued Living Questionnaire (VLQ) | Wilson, Sandoz, Kitchens, Roberts, (2010) | 2.3 |
| Valuing Questionnaire (VQ) | Smout, Davies, Burns and Christie (2014) | 4.5 |
| Comprehensive assessment of Acceptance and Commitment Therapy process (CompACT) | Francis, Dawson & Golijani-Moghaddam, (2016) | 2.3 |
| Cognitive and Affective Mindfulness scale (CAMS) | Feldman, Hayes, Kumar, Greeson & Laurenceau, (2007) | 2.3 |

As illustrated in Table 3.1, the CAQ, CFQ and EQ were most commonly cited, however as they were also not identified in the RCT literature search they were not selected for inclusion in the main questionnaire validation study. Furthermore, the highest frequency of use with these measures was only 6.8% which is lower than the use of the most common questionnaires from the RCTs. The bar chart below (as illustrated in Figure 3.2) displays usage of ACT outcome measures when clinicians were asked how frequently they used each measure with 'all of my clients'.

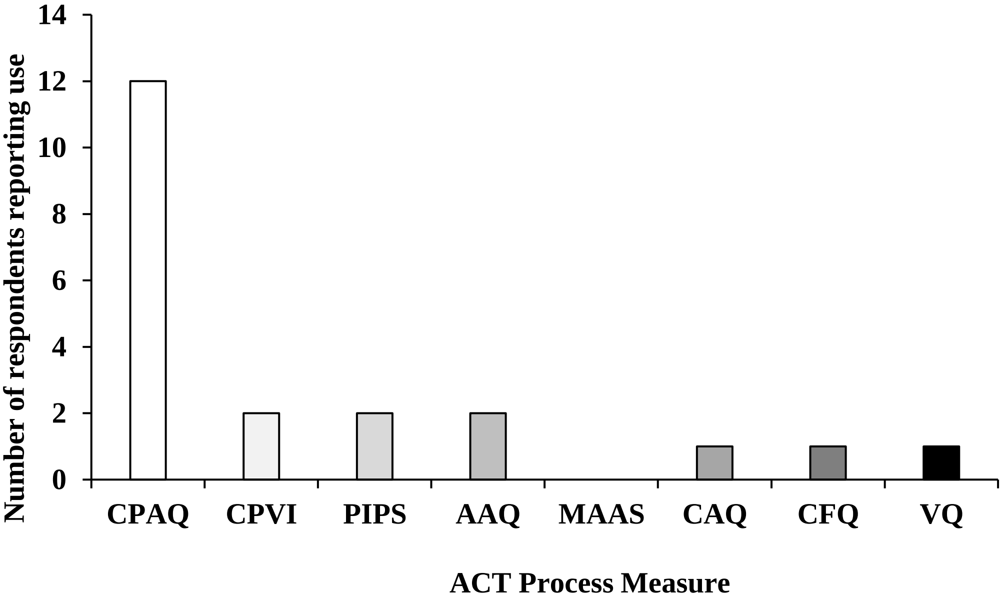


**Figure 3.2 Clinician use of ACT measures 'with all of my clients'**

Again, the CPAQ was the most commonly used, followed by the CPVI, PIPS and AAQ measures**.**

Respondents also provided qualitative answers to questions around which questionnaires they found 'most useful and why', 'most problematic and why' and 'how useful in general do you find the questionnaires?' A framework qualitative analysis was used to note common themes amongst the clinicians' responses and are displayed in the table below (Table 3.2).

**Table 3.2 Framework qualitative analysis to explore common themes around clinicians' experiences of using ACT questionnaires.**

| Main themes | Subthemes | Supporting Quotes |
|---|---|---|
| **Usefulness of questionnaires** | Ease of administration | *'We use the CPAQ as it is quick and patients find it acceptable.'*<br>*'CPAQ & PIPS- well validated questionnaires with easy administration and quick calculation of responses.'*<br>*'I find the AAQ useful as it can be 'eyeballed' quickly and gives useful indications about acceptance/fusion.'*<br>*'I like the CPVI most, it's quick to complete and often insightful.'*<br>*'Really useful, easy to administer and for patients to understand. They are relevant, capture changes.'*<br>*'Very useful measure of change and highlights particular areas of concern for the individual.'* |
| | Face validity | *'CPVI- most face validity for addressing the domains of people's lives that matter to them.'*<br>*'Very useful. Like all questionnaires, some changes do not get picked up by the questionnaires when looking at a single individual, but this is an issue relevant to all such measures.'*<br>*'[All questionnaires] seem reasonably sensitive to change.'*<br>*'CPAQ and CPVI are useful in decision making about whether patients might benefit from a group programme. All of the questionnaires are useful as outcome measures.'*<br>*'Clinically it does seem to reflect progress on this important dimension.' (CPAQ)* |
| | Seems to assess key constructs of | *'AAQ is useful as a proxy measure of changes in psychological flexibility processes.'*<br>*'CompACT -it looks at all 6 areas.'* |

|  |  |  |
|---|---|---|
|  | psychological flexibility | *'CPVI- helps to highlight discrepancy in terms of values importance and success to clients.'*<br>*'The CPAQ is very useful as it addressed the key dimension of avoidance and engagement, and it is now widely appreciated.'*<br>*'…increased acceptance often precedes increase depression or "creative hopelessness" it is useful to be able to capture this process as it happens and as predictors of change at follow up.' (CPAQ)* |
| **Problems with questionnaires** | Wording is not understood by patients | *'They are too wordy and they are well above the average adult national literacy level. It's bizarre that for such an experientially based therapy, the questions are so wordy! We need simple tools which people can easily make sense of.'*<br>*'I find that client sometimes get muddled with the wording on the CPAQ because there are some double negatives.'*<br>*'CPAQ- patients struggle with the language and concepts in part i.e. "who would ever be willing to have pain?'*<br>*'Many patients struggle with the CPAQ, especially when completing it as part of the assessment process as the terminology is often alien to them and the wording is quite confusing with some of the questions...'*<br>*'Many questions in for example CPAQ has a weird tone to them.'*<br>*'Patients find the CPAQ confusing sometimes.'*<br>*'The CPAQ can be difficult for patients to understand, particularly the negatively worded items.'*<br>*'I don't like the fact that the two sub-scales are derived from sets of items with different valences. I know this confuses people and can create the appearance of separate factors.'*<br>*'CPAQ is too complicatedly worded. Patients often leave items out as they don't understand some of the questions.'* |
|  | Patients struggle to understand ACT concepts | *'When you haven't socialised patients to the ACT model, some of the questions can seem a bit confusing or nonsensical.'*<br>*'Only commenting on CPAQ - it's long, and some of the items are quite abstract and have resulted in some clients saying that they don't understand the questions.'*<br>*'CPAQ- patients struggle with the language and concepts in part i.e. "who would ever be willing to have pain?'*<br>*'I find the AAQ-II and PIPS problematic because I find the content confused, heterogeneous, and much narrower than the developers intended.'* |

| | | |
|---|---|---|
| | | *'They are designed by intelligent, academic people for intelligent, academic people, but that is not the main demographic that they are used with in the NHS so there is a misfit.'*<br>*'The language is not always easy for patients to understand but I think it is precise language and helps shape up patient behavior over time in line with the key processes*.' |
| | Issues related to scoring due to patients not understanding the questions | *'There are issues with the CPVI. Most patients rate all domains as highly important so it is hard to discriminate.'*<br>*'I find the scoring tricky as there is no indication of the subscales on the actual questionnaire, plus the patients really struggle with the language, especially the double negatives. Many of them only partially complete it and ask for help on a number of the questions.'*<br>*'I think sometimes patients are not used to identifying their values and so will sometimes mark 'N/A' against some items.'* |
| | Questionable accuracy in measuring change | *'VLQ - its use as a tool to measure change and use for audit/research purposes is a little haphazard.'*<br>*'I don't find them to be helpful except in research contexts. They lack the validity necessary for individual feedback.'*<br>*'On a scale from 0-10 I would say, 6/10. The questionnaires are quite static. They do not assess the function of patients behaviors (i.e., how they cope with their pain situationally in their context), as well as topographically. More idiographic ways of assessing these complex behaviors (i.e., utilizing moment-to-moment changes in the context of patients) could be proven way more useful for clinicians, researchers, and patients*.'<br>*'I don't find the scores reflect the outcomes that are apparent in people's behaviour.'* |
| **Useful only in certain contexts** | Useful in guiding (or when used in combination with) clinical judgement | *'Moderately useful - have to use them in conjunction with clinical judgement in terms of change.'*<br>*'Useful for my self-guidance but as yet, in my context I have not reached a point of showing their usefulness to others*.' |

# CHAPTER FOUR: PART ONE DISCUSSION
# (Research and Clinical use of ACT Questionnaires)

The purpose of this preliminary study was to select the measures most commonly used in both research and clinical practice so that the main study could explore their content validity. Findings from the clinician survey corresponded with the literature search across ACT intervention RCTs in adults with chronic pain, whereby the measures found to be in most regular use included the CPAQ (various versions), CPVI, PIPS and AAQ-I/II. The recent literature review completed in April 2019 confirmed that these findings for use of ACT measures across RCT research remained relevant; the four most frequently cited measures matched those found during the initial literature search. Findings from the clinician survey demonstrated that although other ACT or ACT based measures were utilised by clinicians, for example the VLQ, MAAS or the CAQ, their application was not as frequent and/or did not extend to both research and clinical use.

The survey also offered some interesting themes around clinician experiences of measures. Generally, they were viewed as easy to administer and seemed to show face validity, suggesting they may be a useful tool to measure change in ACT outcome research. However, an opposing theme around the accuracy of these measures also emerged from the data. The consensus amongst clinicians appeared to be that there are issues with the questions with regards to item wording/concepts being understood by respondents, leading to issues with scoring. Clinicians indicated that the questionnaires appeared to be limited with regards to accurately measuring change, as patients did not fully understand the items.

Previous criticisms from the literature have mainly focused on a different aspect of validity: the failure of ACT questionnaires to capture *all* psychological flexibility processes. For example, it has been suggested that ACT process measures are often restricted to certain contexts of application e.g. in the chronic pain population, and are often limited to measuring individual ACT sub-processes (Francis et al., 2016). Francis et al. (2016) noted that the combined use of single-ACT process measures may present problems for interpretation of data and resultant discriminant validity of the measures, whereby their combined use does not give an accurate indication of the overarching process of psychological flexibility. They also highlighted issues with the AAQ-II, specifically in relation to its focus on the sub-processes of acceptance and defusion, despite being purported as a ‘unidimensional measure’ of psychological inflexibility (Bond et al, 2011). It could be suggested that not capturing all of the core ACT process may indeed lead to issues with the questionnaires content validity (they do not give us an accurate representation of an individuals’ current level of psychological inflexibility), however this aspect of validity was not empirically tested. In another study conducted by Wolfgast (2014) it was also suggested that issues surrounding the measure validity may be related to the wording of AAQ-II items. For example, certain items appear to measure general distress outcome variables e.g. difficult emotions, rather than specifically target psychological inflexibility. This is problematic as ACT treatments are designed to make changes with regards to specific ACT constructs, which although related to distress constructs are theoretically separate. Merging the distinct ACT processes with other outcome variables and potentially not capturing all constructs within psychological flexibility may create difficulties in interpreting ACT measures. Importantly however, it should be noted that this study investigated discriminant validity. Although the questionnaires’ content validity was speculated it was not investigated here in a systematic way. Both of these studies illustrate the focus of current research on using quantitative, correlative methodology (factor analysis, multiple regression) to assess validity. However, this does not tell us *how* people understand the questionnaires and what impact this has on responses. These issues will now be investigated in the second (and main) part of the study; the ACT questionnaire validation study. The preliminary study has identified the most frequently used ACT process measures in chronic pain clinical research and practice and these will be included in the main study.

# PART TWO: ACT QUESTIONNAIRE VALIDATION STUDY

## Chapter overview

The following chapter outlines the main study which aimed to understand how individuals make sense of questionnaires as they complete them; the aim being to establish their content validity. Following a brief recap of the points raised so far, Chapter 5 then outlines the methodology and results for the main study exploring ACT questionnaire validity. The cognitive interviewing method was used with patients accessing a chronic pain service. The development of the problem classification matrix is introduced followed by the procedure for the interrater reliability checks. The sample and examples of error classification are described followed by the results for the preliminary interrater reliability check. Results for the main study are then described, including post-hoc analysis.

# CHAPTER FIVE

# INTRODUCTION

As referred to in the general introduction, there appears to be no existing research *directly* investigating how individuals interpret and then respond to ACT questionnaires. The themes from the clinician survey gave insight into the possibility that the questionnaires may have poor content validity related to the understandability of the items. It was felt this may be because of complex item wording, concepts, language and the abstract nature of the questionnaires. Previous findings from the literature also indicate there may be issues surrounding the psychometric validity of ACT process measures, however respondents' experiences of completing items have not been systematically investigated. These issues will be addressed in Part Two of the study; the *first* study to comprehensively test content validity. This main study aimed to test the hypothesis: 'commonly used ACT questionnaires are able to capture their intended processes in people experiencing chronic pain'. This was a questionnaire validation study which aimed to evaluate aspects of content validity in the questionnaires. More specifically it aimed to explore whether individuals understand questionnaire items as intended and establish any particular patterns in misunderstanding, with the purpose of creating recommendations for the future refinement of the questionnaires.

# METHOD

## Design

This study conducted a cognitive interviewing ‘think-aloud’ method to identify difficulties in respondent understanding of the questionnaire items. In order to appreciate how clinically useful ACT questionnaires are, we have to know whether respondents understand the individual questionnaire items to respond to them in the ways intended by the authors, thus establishing content validity. A taxonomy of problem classification enabled errors in individual questionnaire items to be explored through analysing respondents’ ability to interpret, comprehend and retrieve answers, and avoid making errors in arriving at their response. This will enable conclusions to be made around the ability of the ACT questionnaires to accurately reflect the intended process variables.

## Ethical clearance

Ethical approval for the study was obtained by the Yorkshire & Humber-South Yorkshire Research Ethics Committee on 22/06/2017 (IRAS reference: 222447, REC reference: 17/YH/0165). The study was also registered with the Research and Innovation department at Leeds Teaching Hospitals NHS Foundation Trust.

## Participants

The target sample for this study were people with chronic pain who were accessing ACT based treatment within an NHS hospital setting. The following criteria were applied:

## Inclusion criteria:

1. Accessing the Chronic Pain psychology service at NHS St James Hospital, Leeds and receiving ACT therapy (1:1 sessions or ACT based group).
2. Aged 18 years and over.
3. Capacity to consent and make one’s own decisions.

# Exclusion criteria:

1. As questionnaires are written in English, individuals without adequate English language reading and speaking ability were excluded.
2. Those whom the clinician in charge of their care was aware of having recently disclosed risk.

# Sample size

The subsequent analysis relied not on inferential statistical tests of numeric data but instead from the qualitative interpretation of notes from the cognitive interview and the proposed coding framework. A similar study assessing the content validity of the Illness Perceptions Questionnaire- Revised (IPQ-R) using 'think-aloud' methodology on adults with type 2 diabetes, used a sample of 36 participants (McCorry, Scullion, McMurry, Houghton & Dempster, 2013). However, Willis and Artino (2013) have argued that samples size of between 10–30 participants are sufficient for cognitive interviewing. Therefore, this study aimed to recruit an overall sample of between 20-30.

# Measures

To minimise burden to participants, they were asked to respond to the three most frequently used questionnaires (according to previously cited research and clinical findings). This consisted of 31 items in total. These psychometrically-validated and widely used self-report questionnaires are listed below; each questionnaire assessed different ACT process variables/constructs. On all questionnaires, respondents rate items on a seven-point scale (either 0-6 or 1-7) ranging from 0/1 indicating the option 'never true' to 6/7 'always true'. Copies of the measures discussed in the following section can be found in the Appendix B. All items were typed out as individual statements on separate cards in the exact format as they are written in each of the questionnaires. Items from the Pain Self-Efficacy Questionnaire (PSEQ, Nicholas, 1989) were also incorporated into the test material and an item from this was administered prior to the main task (Appendix B). This practice item minimised the chance of participants not understanding the test instructions, a potential confounding variable for subsequent data analysis.

### *Chronic Pain Acceptance Questionnaire (CPAQ-8)*

The CPAQ-8 (Fish, McGuire, Hogan, Morrison & Stewart, 2010) is an eight-item measure and abbreviated version of the full CPAQ (McCracken, Vowles & Eccleston, 2004). In their development and preliminary validation study Fish et al. (2010) found this short-form version to have the same factor structure and psychometric properties as the original CPAQ. It too consisted of a two-factor structure reflecting pain willingness (the pursuit of life activities despite pain) and activity engagement (recognition that avoidance and control are unworkable strategies in adapting to life with chronic pain). The pain willingness scale includes items 2, 4, 7, and 8 (reverse scored) and the activity engagement scale includes items 1, 3, 5, and 6. Scores from both scales are added together to provide the total score. Higher scores indicate greater activity engagement and pain willingness. The Fish et al. (2010) study found the CPAQ-8 demonstrated adequate to good scale score reliability with Cronbach's alpha scores between .77 and .89 for an online version or paper and pencil version of the questionnaire administered to individuals with chronic pain.

### *Psychological Inflexibility in Pain Scale (PIPS)*

The PIPS (Wicksell et al., 2008a) is a 16 item measure which consists of two subscales: avoidance of pain (withdrawal from valued activities in response to or the expectation of pain) and fusion with pain thoughts (difficulty in detaching from thoughts about pain and it's cause) (Wicksell et al., 2008). It is used to measure psychological (in)flexibility (e.g. acceptance, avoidance, cognitive fusion, lack of values/contact) in the context of chronic pain. The 'avoidance of pain' component includes items 2, 3, 7, 8, 9, 11, 13, 14, 15, and 16 whilst the 'fusion with pain thought' scale includes items 1, 4, 5, 6, 10 and 12. Higher scores indicate greater levels of psychological inflexibility. In their development and preliminary validation study Wicksell et al. (2008a) found that the PIPS demonstrated good internal consistency, measured by Cronbach's alphas, of .90 (avoidance), .75 (fusion) and .89 (total scale). An intercorrelation of .46 between the two subscales showed that although associated, they did provide distinctive information about psychological inflexibility. The study demonstrated the concurrent criterion validity of the total scale as well as both the avoidance and fusion subscales. Results showed that the PIPS correlated significantly with 'mental/physical functioning' subscales of the Short Form-12 Health

Survey (SF-12) (Ware, Kosinski, & Keller, 1996), an additional ‘quality of life’ item and with all subscales on the Multidimensional Pain Inventory (MPI), with the exception of ‘support’ (Bergstrom et al., 1998).

### *Acceptance and Action Questionnaire-II (AAQ-II)*

The AAQ-II (Bond et al., 2011) is a seven-item measure and abbreviated version of the full AAQ (Hayes et al., 2004) which is used to measure the construct of psychological (in)flexibility (however not specifically in the context of chronic pain). It is purported to be a unidimensional measure of psychological inflexibility, with higher scores reflecting greater experiential avoidance and decreased acceptance. In their preliminary validation study Bond et al. (2011) found the AAQ-II to measure the same construct as the original measure (correlation coefficient of .97). However, the AAQ-II was more reliable and results showed a mean alpha coefficient of .84 and a 12 month test-retest reliability of .79.

### *Wechsler Test of Adult Reading; the Test of Premorbid Functioning (TOPF)*

An updated version of the Wechsler Test of Adult Reading; the Test of Premorbid Functioning (TOPF; Wechsler, 2011) was administered prior to the main task. The TOPF is used to provide an estimation for premorbid cognitive functioning and has high reliability (.96-.99; Holdnack & Whipple Drozdick, 2009). This brief assessment comprises a list of 70 phonetically irregular words which go in order of increasing difficulty. The TOPF assesses vocabulary through ability to pronounce irregularly spelled words and scores provides an estimation for level of intellectual functioning in relation to verbal fluency.

## Procedure

### *Recruitment*

The study was promoted to patients accessing ACT through the chronic pain clinic at NHS St James Hospital, Leeds. The local Leeds Teaching Hospital (LTHT) PMP based group for people with chronic pain is based on ACT principles and is called the ‘Living with Pain’ group. A recruitment flowchart illustrates the stages of participant recruitment and levels of attrition (see Figure 5.1). Individuals expressing interest in taking part received a study invitation leaflet (Appendix A) and their contact details were collected. Individuals expressing interest were then contacted by the principal investigator to answer any questions and to arrange a convenient time/location to attend the study.

Participant characteristics were recorded to enable future comparisons to be made, based on age, gender, type of chronic pain experienced, stage of therapy (session number), method of therapy (‘Living with Pain’ ACT group or 1:1 therapy) and current medication. All participants had experience of completing the CPAQ as a baseline outcome measure during their initial therapy session with the service. Levels of attrition were recorded from the beginning of the study to identify a cut-off point during recruitment, which may have indicated an inadequate sample size. To increase potential recruitment levels, an option to see the participant either at the chronic pain clinic or their home address was offered.

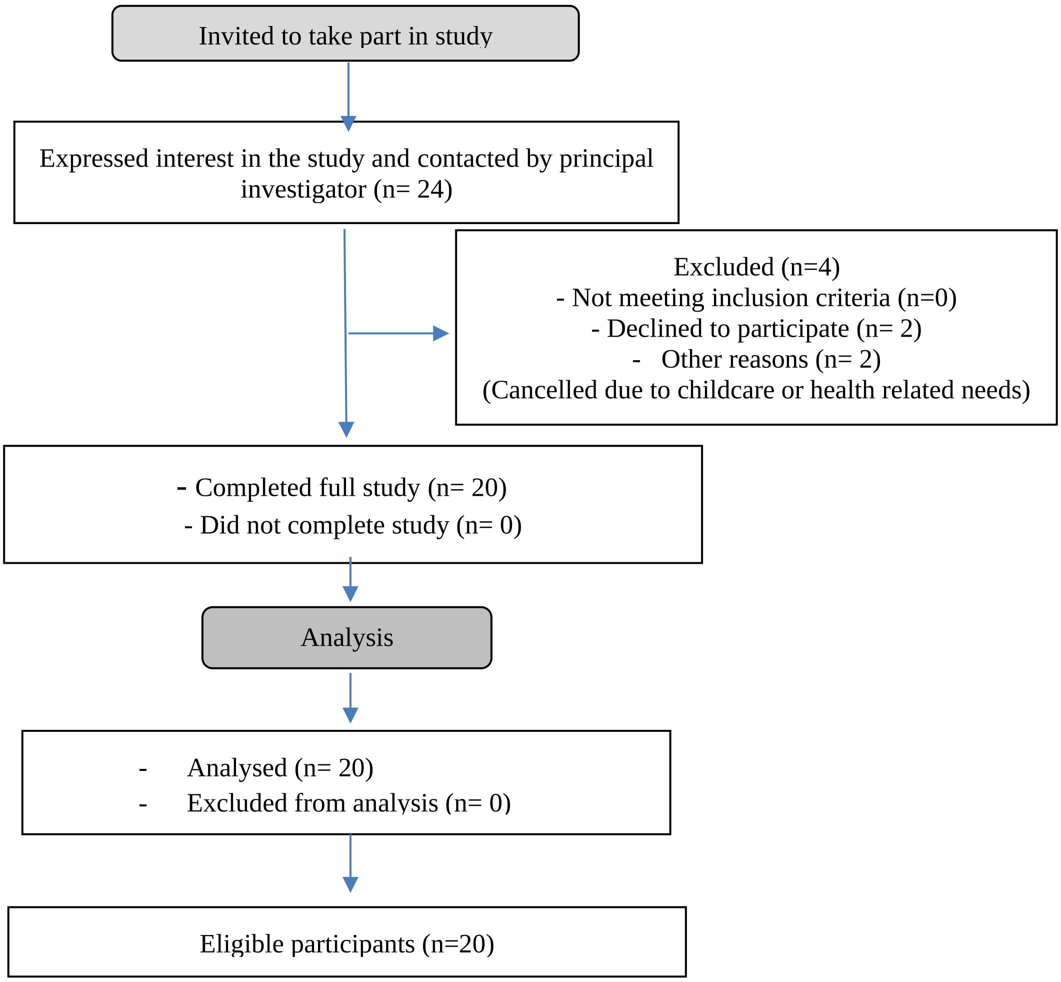


**Figure 5.1: Flowchart to illustrate the recruitment strategy**

### *Involving service users in the study design*

Individuals from the service user and carer reference group for the University of Leeds Doctorate, ‘Everybody’s Voice’, were consulted over the study design and development of the participant information sheet and consent from. Positive feedback on the consent documents was received about clarity, readability, use of paragraphs to separate the text and limited use of jargon. Comments around formatting included: removing a border which distracted away from the main text, and a suggestion to make the title font larger. Service users fed back that possible risks outlined on the participant information sheet (including possible inconvenience of participating) were outweighed by the benefits of taking part in an interesting study. They liked the flexibility offered to conduct the study at home, should this be more convenient. Service users said that the potential risk for distress to be evoked by the questionnaires was clear and they felt reassured by information provided around additional support they could access. The message around being able to withdraw from the study at any time was clear and service users said this was reassuring.

### *Main cognitive interviewing study procedure*

Prior to commencing the study, the participant was provided with an information sheet and consent form. Background information detailing participant characteristics was collected, including: age, gender, type of chronic pain experienced, stage in therapy (session number), method of therapy (ACT group or 1:1 therapy) and medication status.

An updated version of the Wechsler Test of Adult Reading; the Test of Premorbid Functioning (TOPF; Wechsler, 2011) was administered prior to the main task according to standardised procedures. This brief assessment required the participant to read from a list of 70 words phonetically irregular words which go in order of increasing difficulty. Participants were asked to pronounce each word out-loud and in order until the discontinue rule was met (5 consecutively mispronounced words). This test assessed vocabulary through ability to pronounce irregularly spelled words and scores provided an estimation for level of intellectual functioning in relation to verbal fluency. The cognitive interviewing task required participants to be verbally fluent in their articulation of responses to questionnaire items so that subsequent analysis could infer where the difficulties in understanding the items may lie. This brief assessment of intellectual functioning was therefore administered so that the impact of individual differences in intellectual functioning on task performance could also be considered.

The participant was then given a set of written instructions outlining what the task involved. Instructions for the cognitive interviewing task (adapted from best practice instructions; Gilhooly & Green, 1996) were read out to the participant. They were told that the study required them to 'say out loud your thought processes as you answer each question.' The instructions included a Likert scale ranging from 1 'never true' to 7 'always true'. Participants were asked to describe everything they were thinking from when they read the question item to arriving at a response on the scale- i.e. how much they agreed with each ACT questionnaire 'statement' on the individual cards. They were asked to think their answers out loud without trying to plan or explain what they were saying, acting like they were in the room by themselves and to remember to keep talking. They were asked to speak as loudly and clearly as possible and to refer to the summary instructions as needed. An example statement and answer to an item from the PSEQ was also provided: *'I'm thinking that when I'm with my family I'm able to cope with the pain, however when I'm at work it becomes much more difficult-therefore I would rate this question a 3.'* (Example answer to PSEQ item 4 'I can cope with my pain in most situations'). The participant was then provided with another example item from the PSEQ to practice and warm-up to the 'think aloud' technique required of them in the main task. This cognitive interviewing training ensured that participants were clear what to do, in acknowledgment that 'thinking out loud' responses may not come naturally. When it was apparent that participants were confident in their ability to perform the task they had the opportunity to ask questions before the study began. A verbal check was given following the third item to again ensure participants felt comfortable in completing the task. This cognitive interviewing technique identified the processes used by the participant to arrive at their answer, checking for answers that are incongruent to the question. This methodology can be used to explore possible problems with questionnaire items when used in this way (Van Someren, Barnard & Sandbery, 1994). As little interjection was used whilst participants generated this verbal information this method was aligned to a 'pure' think-aloud methodology as opposed to the 'probing' technique (both described by Beatty & Willis, 2007). This methodology has the strength of a standardised procedure and thus reduces the likelihood of interviewers biasing the collection of data (Beatty & Willis, 2007). The think-aloud technique also minimises the disruption of flow and possible content change that a probing technique may interrupt (Conrad, Blair & Tracey, 2000; in Beatty & Willis, 2007). Therefore, prompts were kept to a minimum during data collection and only provided if the participant did not speak for a while or more detail was required from their answer. These included: ‘can you say a bit more…?’; ‘tell me what you’re thinking…’ and ‘just describe everything you’re thinking from when you read the card…’ If answers appeared to be going off-topic and not related to answering the item in front of them, they were thanked for the level of detail given and then prompted, ‘but it might work a bit better if (would you be OK to just) keep trying to describe your thought processes.’ Questionnaires were administered according to a randomly generated sequence, allocated to each participant. Following the provision of all 31 test items a debrief statement was provided regarding their participation in the study. All speech was subsequently transcribed verbatim for analysis and data were anonymised through the use of a randomly allocated number for each participant.

# Card sort

Participants also completed a card sort task following cognitive interviewing, to further assess construct representation. This involved participants placing each of the 31 items into piles which they felt went together and providing a brief explanation as to why they organised the items in that way. Each pile was allocated a label to describe what it represented. This took around 10 minutes to complete. This data analysis is not reported here as it was beyond the scope of this thesis. It will be analysed following the completion of the thesis. Further detailed information on the card sort method is contained in Appendix C.

# Analysis

## *Development of error classification matrix*

The first stage of analysis was modelled on techniques used within previous research (Conrad & Blair, 1996; Drennan 2003; McCorry et al., 2013; Van Oort, 2011). A coding framework or ‘problem classification matrix’ was developed which formed the main body of analysis. This taxonomy of possible problems in response processes classified errors into one of three distinct categories (lexical, logical or conceptually inconsistent errors). Drennan (2003) recognised the difficulties that respondents experience in understanding and accurately completing questionnaires. The benefits of using cognitive interviewing has been recognised in samples where *'questionnaire completion may pose particular problems'* (Drennan, 2003, p.58). This was particularly relevant to the current study, where the online clinician survey had already highlighted issues surrounding the understanding of ACT questionnaires amongst the chronic pain population. However, the data collected from cognitive interviewing is qualitative and therefore subjective in nature as it relies on individual researcher interpretations of the data. Conrad and Blair (1996, 1999) developed a taxonomy of problem classification to address this fundamental flaw in cognitive interviewing data analysis. Their classification system evaluated logical problems, temporal problems, inclusion/exclusion problems, lexical problems and computational problems. Logical problems were described as difficulties in answering the question, related to complex item structure (the respondent may have difficulty in connecting concepts e.g. 'and' or 'other than'). Temporal problems were described as response issues related to time, however could be viewed as a variant of logical error. Lexical problems were described as difficulties in understanding the meaning and use of words in the context within which they are used on the item. Inclusion/exclusion problems were viewed as a type of lexical issue around the meaning of words when the respondent is unclear what constitutes a certain category e.g. interpreting the word 'doctors' as inferring other medical professionals when the item was intended to specifically ask about physicians. Finally, computational problems were described as those not fitting into the other categories. It has been well recognised throughout the research literature that these issues may lead to inaccurate completion of questionnaire items, as respondents fail to understand what questions are asking of them, or could lead to incomplete responses as they skip the question entirely (Conrad & Blair, 1996; Drennan, 2003; Van Oort, Schroder & French 2011). Drennan (2003) purported that the problem classification matrix data is a useful first step to standardising the analysis of cognitive interviewing data, however it remains not fully objective. The present study utilised Conrad and Blair's (1996) taxonomy of possible problems in understanding questionnaire items. In doing so this questionnaire validation study aimed to explore the content validity of ACT questionnaires in a systematic way.

The process of analysing the interview transcripts for errors involved reading through each response for the 31 questionnaire items. Each item was systematically coded as: no error present, logical error identified or lexical error identified. A third error category was added specifically for the purposes of this study and in response to the research hypothesis: respondents do not understand the items, therefore responses are not likely to reflect the ACT concept being measured. This was categorised as an error related to item precision: ‘conceptually inconsistent error’. (During the final analysis an additional fourth category also emerged, ‘response error’* and has been included in Table 5.1). Each item could be coded within multiple categories. The classification of errors was guided using Table 5.1. In deciding whether there was an issue with item precision (conceptually inconsistent error) the researcher looked for responses which suggested an error was made in relation to the ACT concept being measured. Irrespective of the score given by respondents, the item was considered to be ‘on-target’ if the question was answered in the way presumed – based on the ACT process under study – to be intended by the original questionnaire author. Items coded into any of the error categories were considered ‘problematic’ for the interpretation of that item.

**Table 5.1 Problem classification taxonomy (from Conrad & Blair, 1996; 1999)**

| **Logical error** | **Lexical error** | **Conceptually Inconsistent error** | **Response error*** |
|---|---|---|---|
| Respondent has difficulties in answering the question. They understand the question but have trouble performing the task. Issues are related to complex item structure. Respondent may have difficulty in connecting concepts e.g. ‘and’ or ‘other than’. When deciding if a logical error is present consider: does the respondent hold in mind **all** aspects of an item e.g. if the item consists of two parts? | Respondent struggles to understand the meaning behind the words and the context in which they are used in the question. | Respondent *technically* answers the question correctly (no lexical/logical errors) BUT the answer differs from the way that was intended by the questionnaire author. The answer suggests they do not understand the question. Therefore, the response is not an accurate reflection of the ACT concept being measured. The question is interpreted in such a way that it becomes the opposite of the ACT process under study. | Respondent appears to misread the direction of the Likert scale, providing a score which contradicts their cognitive interviewing answer. Although technically a variant of logical error, this new error type provides further examples of respondent misinterpretation of items, whereby the scores provided do not accurately represent the level of psychological flexibility indicated in their detailed answer. |

## Check of Interrater Reliability

Before analysing the cognitive interviewing responses to see where problems occurred in participant understanding of each item, it was necessary to first assess interrater reliability of the problem classification framework. Two coders; rater One (lead researcher) and rater Two (supervisor A) assessed the same sample of responses. Rater Two fulfilled the criteria to be recognized as an 'expert' in ACT with over 8 years of clinical and research experience using ACT. Agreement between the raters was explored for: 1) identifying the presence of an error within an item response; 2) the type of error classified (lexical/ logical/conceptually inconsistent). Reliability checks measured how the raters classified individual items within equivalent error categories; it indicated the extent to which the classification system led to reliable measurement of errors.

Interrater agreement was checked over three separate time points (see Table 5.2). At time 1, the score for interrater reliability was low (as described in the results section to follow). Therefore an additional check of reliability was performed following subsequent discussions between the raters around areas of discrepancy and a repeat assessment of the sample responses (time 2). Interrater reliability improved slightly (see Results). A third independent rater; rater Three (supervisor B) was then also involved in the third analysis of sample responses (time 3) (see Appendix for raw data with reflections). This gave further rigour to the reliability check and provided additional contextual information around our decision making for error identification/classification. The raters met following each independent analysis to discuss any areas of discrepancy and thus further train themselves on error measurement for the final analysis. After the final interrater reliability check (time 3) any remaining discrepancies between the raters were discussed so that any changes to error classification were reflected in the final analysis. Table 5.2 outlines which researcher was involved in each check of interrater reliability over the various time points. Three separate data sets (from participants 1, 3 and 4) were used consistently during the check of interrater reliability and at all time points.

**Table 5.2 Raters conducting the interrater reliability checks over three occasions**

| **Time point** | **1** | **2** | **3** |
|---|---|---|---|
| Rater 1 (lead researcher) | ✓ | ✓ | ✓ |
| Rater 2 (supervisor A) | ✓ | ✓ | ✓ |
| Rater 3 (supervisor B) | | | ✓ |

Cohen’s kappa coefficients (κ, Cohen, 1960) were calculated to determine if there was agreement between the initial two raters for identifying and classifying errors. They represent the proportion of agreement *over and above* chance and range from -1 to +1. Cohen’s kappa coefficients were calculated through entering the raw data into SPSS, Version 23 (IBM Corp, 2017) for statistical analysis.

As this statistical method was not designed to measure agreement between more than two raters, Fleiss’ Kappa (κ, Fleiss, 1971) were also calculated as an additional confirmation of interrater reliability for the third analysis (using three independent raters) and for time points one and two (see Results). This statistical method is used for nominal level of measurement. A decision was made to calculate Fleiss’ Kappa’s for error agreement (i.e. was an error made) and error classification (i.e. if so, what sort of error was it?) individually as opposed to a combined overall result. They were viewed as two different decisions made when assessing for error within the items, therefore collapsing into a combined score risked losing information about correlation and agreement between raters. These reliability checks and the resulting discussion ensured the raters learnt from each other and consolidated the information required to consistently categorise errors amongst the ACT questionnaire items. It allowed the raters to have confidence in the error classification matrix as without such reliability checks, validated conclusions cannot be made from subsequent analysis. Fleiss’ Kappa coefficients were calculated through an online statistical utility ‘ReCal’ (Reliability Calculator) which is used for computing interrater reliability coefficients for nominal-level data using two or more coders (Freelon, 2010).

# RESULTS

## Sample Characteristics

Participant demographics for all 20 individuals who took part in the study are summarised in Table 5.3.

**Table 5.3: Participant demographics**

| | | **n** |
|---|---|---|
| **Gender** | Male | 7 |
| | Female | 13 |
| **Age** | 18-25 | 0 |
| | 26-35 | 1 |
| | 36-45 | 4 |
| | 46-55 | 9 |
| | 56-65 | 5 |
| | 65+ | 1 |
| **Method of therapy** | ACT group | 16 |
| | 1:1 therapy | 5 |
| **Stage of therapy –number of sessions attended**<br>(Group; 1-4 = early, 5-8 = late)<br>(1:1; 1-6 = early, 7 ≥ = late) | Early | 6 |
| | Late | 14 |
| **TOPF standard score (qualitative description)** | <69 (Extremely low) | 0 |
| | 70-79 (Borderline) | 2 |
| | 80-89 (Low Average) | 1 |
| | 90-109 (Average) | 10 |
| | 110-119 (High Average) | 5 |
| | 120-129 (Superior) | 2 |
| | >130 (Very Superior) | 0 |

65% of the sample were female and 45% were within the 46-55 age category. Seventy-six percent were recruited from the ACT group for chronic pain (one participant had completed the ACT group and was also receiving 1:1 therapy) and 70% were recruited from the ‘late’ stage of therapy (categorised by number of sessions attended). The mean TOPF score across all participants fell within the ‘Average’ range (*M* = 102, *SD* = 13.6). Additional information on medication taken for pain was gathered; all 20 participants reported use of pain relief medication. These included: adalimumab, amitriptyline, baclofen, dihydrocodeine, naproxen, morphine, tramadol, codeine, paracetamol, gabapentin, co-codamol, ibuprofen, pregabalin, duloxetine, fentanyl patches and matrifen transdermal patches. Information on the type of chronic pain experienced and origin of pain was also gathered and included: pilonidal disease, degenerative disk disease, post-surgery pain (spine, kidney, jaw), osteoarthritis, fibromyalgia, spinal cord nerve damage, functional neurological disorder (FND), cluster headaches, chronic migraines, complex regional pain syndrome (CRPS), lower back pain and sciatica.

# Examples of error classification

620 individual items in total were systematically coded for logical, lexical or conceptually inconsistent errors (some coded within multiple categories). During the analysis a new 'type' of error emerged which had not been accounted for in the original error classification taxonomy. These 'response errors' occurred when a respondent appeared to have misread the direction of the Likert scale, providing a score which contradicted their cognitive interviewing answer. Although technically a variant of logical error, this new error type provided further examples of respondent misinterpretation of items, whereby the score provided did not accurately represent the level of psychological flexibility indicated by their detailed answer. Table 5.4 below illustrates each error type with supporting quotes extracted from the raw data. Examples are also provided for 'on-target' responses; i.e. the respondent answered the question in the way presumed to be intended by the original questionnaire author – i.e. to be consistent with the psychological flexibility process understudy.

**Table 5.4: Example raw data for each error classification**

| **Error Classification** | **Item** | **Error classification rationale** | **Supporting Quotes (place of error is highlighted for ease of reading)** |
|---|---|---|---|
| Logical error | AAQ3- 'I worry about not being able to control my worries and feelings.' | The respondent does not refer to the part of the question regarding '*being able to control'* their worries. Instead, they focus only on the amount /content of their worrying. | *'I'd say three for that, because - this is the same for previous answer – it's the same answer for – it's a similar answer to the previous question, because I'd say ninety percent of the time I'm fine, it's just on a night if I've had a bad, painful headache that it's worry about– you're worrying about what's going to be and what's going to happen, but then obviously once you've calmed down you can get perspective quicker and realise that it's me being silly. Not silly but blowing it all out of* |

|  |  |  |  |
|---|---|---|---|
|  |  |  | *proportion, so I'd say it's 'Seldom true' really.' (Participant 7)* |
|  | CPAQ2- 'Keeping my pain level under control takes first priority whenever I am doing something.' | The respondent only answers whether they keep *'my pain level under control'* (evident in: 'unless I forget about my back-up'). They do not refer to whether it takes *'first priority'* during activities, i.e. the second part of question. | *'So I'm trying to think: "What do I- what does the question mean: 'I keep my pain under control'?" Yeah I think if I know I've got to do something I'm going to take as many tablets as I can to make sure that the pain doesn't break through but hopefully I'll get the something done so- yes, I think that's almost always true, unless I forget about my back-up, I've not got the tablets yet.' (Participant 4)* |
|  | CPAQ8- 'My worries and fears about what pain will do to me are true.' | The respondent does not include the concept of 'truth' of thoughts. They interpret as worry about pain in the short term, therefore do not answer the question. | '*A bit of a middle of the road one, sometimes it's true sometimes not and so I'd go for a 4. I do worry about overdoing it because I do tend to again…take more painkillers than I probably should because I know I'm about to go and do something and then I go beyond what I feel comfortable with because the pain's not telling me to slow down. I do sort of have to plan where my limit is and stop before I get to where I think my limit is rather than going beyond it.'* (Participant 3) |
| Lexical error | AAQ5- 'Emotions cause problems in my life.' | The respondent appears unclear about what is meant by the term *'emotions'*. They refer to whether or not they are 'over emotional'. | *'Emotions' um, I don't think I'm over emotional really. I don't know what it means by 'emotion', does it mean: Do I think I'm over emotional? I don't know. I think, um, I think I'm a worry – worrier, but is that emotion? I don't know. I think, uh - I think 'worry' perhaps causes problems but – in my life, I don't know if –* |

| | | | |
|---|---|---|---|
| | | | *is 'emotion'- does that mean- it means 'sadness, doesn't it, 'emotion'? ((whispered)) Sadness ((whispered)). No, I think– sadness then makes you feel a bit depressed though, doesn't it? Causes problems. Um, I'm not sure about that question. (Participant 4)* |
| | PIPS14-'I postpone things on account of my pain.' | The respondent refers to 'planning' and not postponing. The score is therefore indicative of planning behaviour. | *'I'll go for a 6 on that I do occasionally put things off, but again I plan around what the event is or whatever I plan the things around because of how I'm going to feel if I do a certain thing. I don't necessarily postpone them, but I do plan very carefully.' (Participant 3)* |
| Conceptually Inconsistent errors (CI) | PIPS13-'Because of my pain I no longer plan for the future.' | The answer does not tap into acceptance of pain. It focuses on whether or not the respondent uses planning, (the function appearing to account for pain). | *'Um I think – no, I don't think that's true so I'm going to say 'Never true' but I'm trying to think of, uh, planning for the future. I feel as though I've got to do more planning, really 'cos, uh, I've got no, I've got to do more planning. (Participant 4)* |
| | CPAQ4-'Before I can make any serious plans, I have to get some control over my pain.' | The answer does not reflect acceptance/ willingness constructs as the respondent already has 'control over my pain'- through managing their medication. The item is interpreted as implying 'you should have control over pain'. Therefore, low score is an inaccurate reflection of psychological flexibility. | *'I feel I already have some control over my pain. I've been doing this, I've been monitoring and maintaining my medication levels for a long time… so if I make any serious plans I already know where I'm at with my pain level and how to control it- so I would say... I have control over my pain so this is very seldom true.' (Participant 1)* |

|  |  |  |  |
|---|---|---|---|
|  | PIPS15-'I cancel planned activities when I am in pain.' | The respondent is so good at experiential avoidance (via planning and pills) they do not cancel plans. Therefore, the low score does not accurately represent the actual level of psychological flexibility. | *'I seldom cancel things and certainly not at the last minute. Most of the time I'm that well planned and up early and pain-killered up or whatever that planned activities usually happen. So very seldom, I'll go for a 2 on that.' (Participant 3)* |
| Response error | PIPS2-'I don't do things that are important to me to avoid feeling my pain.' | The respondent does not agree with the item and states that they 'try and do things that are important'. However, their final score (suggesting low flexibility) contradicts their detailed answer (suggesting high flexibility). | *'I don't do things that are important to me to avoid feeling my pain.' No, it's not right, I always try and do things that are important to me at least now, in this time, in this period, I try because, uh, I'm very perfectionist person and when something is important to me I do it, despite my pain and anything. 'Almost always true', number six.' (Participant 5)* |
|  | PIPS11-'It's not me that controls my life, it's my pain.' | The full answer reveals the respondent does feel in control of their life and not the pain, yet they appear to have mistakenly scored the opposite end of the scale. | *'It's not me that controls my life it's my pain. No I still try to control my own life, you just get used to living with the pain, and you know through medication trying to control it, but I wouldn't say the pain controls my life, I still like to think I'm in control, and that is the reason why I still try and plan things, although they might not go ahead. Frequently true.' (Participant 15)* |
| No error present | PIPS 8- 'I avoid doing things when there is a risk it will hurt or make |  | *'I avoid doing things when there is a risk it will hurt or make things worse.' Um I think –except in those special circumstances that I've just talked about- I am very, very cautious, um, about making the pain* |

| | | | |
|---|---|---|---|
| | things worse.’ | | *worse, um, certainly when there’s a risk of falling, which the pain and weakness cause me to fall quite a lot, so, um, I would say that was um, number five, ‘Frequently true’. (Participant 8)* |
| | AAQ 2- ‘I’m afraid of my feelings.’ | | *‘I’m afraid of my feelings.’ Oh, no! Never (h) No, I’ve learned that (h) If you’d have asked me that at 16 I probably would have cried at the question, but no. I wear my heart on my sleeve, I – do you know what I mean? It’s like I’ve been saying to my mum recently: “If you feel sad, feel sad. If you feel angry, feel angry. If you feel happy, feel happy Don’t be annoyed at yourself for going through something that is going to make you sad and be feeling bad about it – You know what I mean? Just no I’m not afraid of it. I’d do a two actually, because sometimes I do let it get to me, it does break me, but no. Like, if I’m feeling it, I’m feeling it and I make sure that I do because otherwise you don’t learn from it.’ (Participant 9)* |
| | CPAQ 8- ‘My worries and fears about what pain will do to me are true.’ | | *‘My worries and fears about what pain will do to me are true.’ I don’t know. I don’t – I don’t think that’s true. I don’t think about the worry or a fear of what the pain will do to me, it’s more of what the pain - what I’m feeling, not what it’ll do to me. So I don’t– ((sigh)) I don’t think that’s- I think that might be number one, ‘cos it’s– I never worry about what it’s going to do to me because how can you worry about* |

| | | | |
|---|---|---|---|
| | | | *something you don’t know? Just – I don’t really know how to answer that one. Um, yeah, yeah, because I don’t really worry about what it’s doing- what it will do to me it’s– you haven’t got time to worry about– you’ve got other things, when the pain is at its worst, you just want to breathe. So you don’t really have time. I don’t look ahead and worry about what it’s going to do to me in the future it’s all about now. (Participant 12)* |
| Lexical and Conceptually Inconsistent errors (CI) | PIPS 10-‘I put a lot of effort into fighting my pain.’ | Lexical- *‘fighting’* is changed into ‘working’ with pain’.<br>CI- the lexical error leads to a CI error. Their answer reveals some degree of acceptance and attempts to live in line with values- yet their final score reflects high levels of psychological inflexibility. | *‘Yes I do put a lot of effort in to fighting my pain, I don’t know whether it’s fighting the pain or working with the pain, I do try and do things that will help me like exercise and things that I’ve been taught in physiotherapy, but the same again I can’t always do it, it just, it just depends on how I’m feeling that day as to whether I can do it. I mean at the moment I’m not able to get on to my exercise bicycle but there’s other things that I can do and I try to do them as regular as I can, so I would say that that’s almost always true.’ (Participant 15)* |
| | PIPS13- ‘Because of my pain I no longer plan for the future.’ | Lexical- ‘*no longer*’ is substituted for ‘constantly’ planning. Current as opposed to long-term plans are referred to.<br>CI- these plans serve the function of avoiding pain (they plan constantly). The | *‘Erm I’m not sure which way round to read that… erm because of the pain I constantly plan for the future erm and there’s a bit of a double negative there, so I think that’s never true to make it that I plan for the future all the time.’ (Participant 3)* |

| | | | |
|---|---|---|---|
| | | low score of ‘never true’ does not reflect levels of psychological flexibility indicated in this answer. | |
| Logical and Conceptually Inconsistent errors (CI) | AAQ1- ‘My painful experiences and memories make it difficult for me to live a life that I would value.’ | Logical- refers only to painful memories.<br>CI- only refers to presence or absence of an experience, with no sense of what is done when memories interfere. | *‘Frequently true, I had PTSD post-hospital, and I still have memories of that, and certainly when I have heightened pain those memories come to the fore.’ (Participant 18)* |
| | AAQ3-‘I worry about not being able to control my worries and feelings.’ | Logical- answer refers to the extent of worry and does not include the concept of control. Therefore, an important part of the question is not answered.<br>CI- answer suggests they consider the extent of their worry, (does not want family to worry about her). Ability to ‘*control my worry’*, is not considered in answering the question. | *‘This is probably true I do worry about things, but I don’t like to let anybody else know how I’m worrying or how I’m feeling erm I try to cover up a lot of things and I lie to my family and friends so they’re nowhere near aware of what my worries and feelings are so I would probably say this is true-number 6, almost always true.’ (Participant 1).* |
| | PIPS 1- ‘I would do almost anything to get rid of my pain.’ | Logical- Answer comments on whether or not the respondent has done anything that has ‘released the pain’.<br>CI- Score indicates a degree of flexibility, yet answer shows a degree of experiential avoidance. | *“I would do anything to get rid of my pain.’ I don’t know how to answer that. ‘Sometimes true’, four. Yeah, I don’t really - I haven’t really done that has actually released the pain I’ve not experienced that because it’s continuous, does that make sense? It’s like if I could do something I would have done it if that makes sense. So* |

| | | | |
|---|---|---|---|
| | CPAQ 2- ‘Keeping my pain level under control takes first priority whenever I am doing something.’ | Logical- the respondent answers a different question, whether they are managing to keep pain levels under control.<br>CI- they mark seldom true, yet the answer reveals they would try and control pain if they could. Therefore, the score not an accurate depiction of level of psychological flexibility. | *I would say that that’s probably ‘Seldom true’, three, yeah.” (Participant 6)*<br>*‘Well sometimes I haven’t really got a control over the pain level, I take my medication every day, but depending how bad it is it doesn’t always go away, it might help me but it certainly doesn’t go away, so that’s hard to say that, that, keep, I can, I don’t feel like I’ve got control over keeping my pain level under control, that’s, you know, that’s my body that’s not, I haven’t got no say in it I don’t feel, no, so. Seldom true.’ (Participant 15).* |
| Lexical, logical and Conceptually Inconsistent errors | CPAQ4- ‘Before I can make any serious plans, I have to get some control over my pain.’ | Lexical- the meaning of *‘serious plans’* is interpreted as those which must be in the future.<br>Logical- the respondent appears to struggle to answer the item.<br>CI –the final score suggests high flexibility, yet the answer reveals low levels of acceptance. | *‘No, not now. I can make a serious plan because a serious plan is something that’s going to be in the future which gives you time to prepare, so you don’t- before making the plan you don’t- that- that to me is a silly question, because why would you have to do that when- if you’re making a plan it’s in the future, so you’ve got time to control your pain without- I’d say that were ‘Never true’, number one.’ (Participant 12)* |
| Lexical and Logical error | AAQ2- ‘I’m afraid of my feelings.’ | Lexical- the respondent refers to ‘blame’ as opposed to whether they are *‘afraid’* of their feelings.<br>Logical –the item is not answered in the way it was intended. | *‘I’d say, seldom, seldom true. Feelings of, I don’t* blame anyone *it’s just one of those things so I’d say it’s seldom true.’ (Participant 16)* |

# Interrater reliability check

Cohen's kappa's (κ) were calculated using SPSS and interpreted using guidelines provided by Cohen (1960). These are summarised in Table 5.5 below.

**Table 5.5 Assessing agreement with Cohen's kappa**

| Value of Kappa (κ) | Strength of agreement |
|---|---|
| ≤ 0 | None |
| 0.01–0.20 | None to slight |
| 0.21–0.40 | Fair |
| 0.41– 0.60 | Moderate |
| 0.61–0.80 | Substantial |
| 0.81–1.00 | Almost perfect |

κ = *Cohen's kappa coefficient*

The interrater reliability check determined if there was agreement between rater 1 (lead researcher) and rater 2's (supervisor A) judgement on whether: 1) errors existed in respondent's answers to the individual items and 2) whether this error was classified as logical, lexical or conceptually inconsistent error (Refer to Appendix E for SPSS output).

**Table 5.6 Results of Cohen's kappa calculation in SPSS** [a, b]

| | Time one analysis | Time two analysis |
|---|---|---|
| **Error agreement** (κ)= | .460 | .561 |
| **Error classification** [c] (κ)= | .282 | .366 |

a. *Values reported for participant data sets one, three and four.*

b. *One missing data point- one item not answered to a level sufficient enough to code for error (participant 4 – AAQ 1).*

c. *When more than one error identified for any given item each separate classification was compared across raters.*

Table 5.6 demonstrates that following the first analysis there was moderate agreement between the two raters' judgements of whether an error was present, $\kappa$ = .460 (95% CI), $p$ < .0005. Following the second analysis the level of agreement increased but remained in the moderate range, $\kappa$ = .561 (95% CI), $p$ < .0005. Following the first analysis there was fair agreement between the two raters' judgements of which error type was classified, κ = .282 (95% CI), $p$ < .0005. Following the second analysis the level of agreement improved slightly however was still classified as fair, $\kappa$ = .366 (95% CI), $p$ < .0005.

As there remained a fair-moderate level of agreement around the presence of errors amongst items and the error (type) classified it was felt that more work was needed to improve interrater reliability. A third analysis with the additional rater 3 (supervisor B) provided the opportunity to further discuss areas of discrepancy and consolidate the rationale for categorising errors. As described previously, because Cohen's kappa calculations are limited to measuring agreement between two raters only, Fleiss' Kappa (κ, Fleiss, 1971) coefficients and their 95% CIs were also calculated using the statistical program ReCal and interpreted using guidelines provided by Landis and Koch (1977) (see Table 5.7 below).

**Table 5.7 Assessing agreement with Fleiss' Kappa (k) (from, Landis & Koch, 1977)**

| Value of k | Strength of agreement |
|---|---|
| <0.00 | Poor |
| 0.00 to 0.20 | Slight |
| 0.21 to 0.40 | Fair |
| 0.41 to 0.60 | Moderate |
| 0.61 to 0.80 | Substantial |
| 0.81 to 1.00 | Almost perfect |

*k = Fleiss' Kappa coefficient*

This second interrater reliability check was performed to determine the level of agreement between raters on the presence of errors and their classification. The results are shown in Table 5.8 for the three separate analysis (time one, two and three).

**Table 5.8 Results of Fleiss' Kappa calculation in SPSS**[d]

| | Time one analysis | Time two analysis | Time three analysis |
|---|---|---|---|
| **Error agreement** | .453 | .689 | .765 |
| **Error classification** [e] | .271 | .297 | .676 |

*d. Values reported for participant data sets one and three.*
*e. Agreement across error classification calculated where all three raters reported error present.*

Table 5.8 demonstrates that following the first analysis there was moderate agreement between the rater 1 and 2's judgements of whether an error was present; κ = .453 (95% CI). Following the second analysis the level of agreement improved to substantial; κ = .689 (95% CI). Following the third analysis (involving an additional rater) interrater reliability improved further and the level of agreement remained at substantial, κ = .765 (95% CI). Following the first analysis there was fair agreement between rater 1 and 2's judgement of which error was classified; κ = .271 (95% CI). Following the second analysis the level of agreement remained within the fair agreement range; κ = .297 (95% CI). Following the third (three rater) analysis interrater reliability improved and the level of agreement was substantial; κ = .676 (95% CI). These checks reached the level required (Landis and Koch, 1977) for us to then confidently use the error classification matrix during the main analysis.

# Main Error Classification Analysis

## 1) What errors are made and how frequently?

Firstly, we explored what type of errors were made and how frequently these occurred. This analysis specifically related to the original hypothesis which suggests that errors are made as individuals answer the questionnaire items, thus intended processes are not captured. This section focused on identifying the difficulties experienced by participants in completing the items.

### *Frequency of errors made per category of error classification*

Table 5.9 demonstrates the error frequencies as the proportion of respondents making at least one error on each measure (*total error*). The table also shows the proportion of respondents making errors within each separate error category for each measure (CPAQ-8, PIPS and AAQ-II).

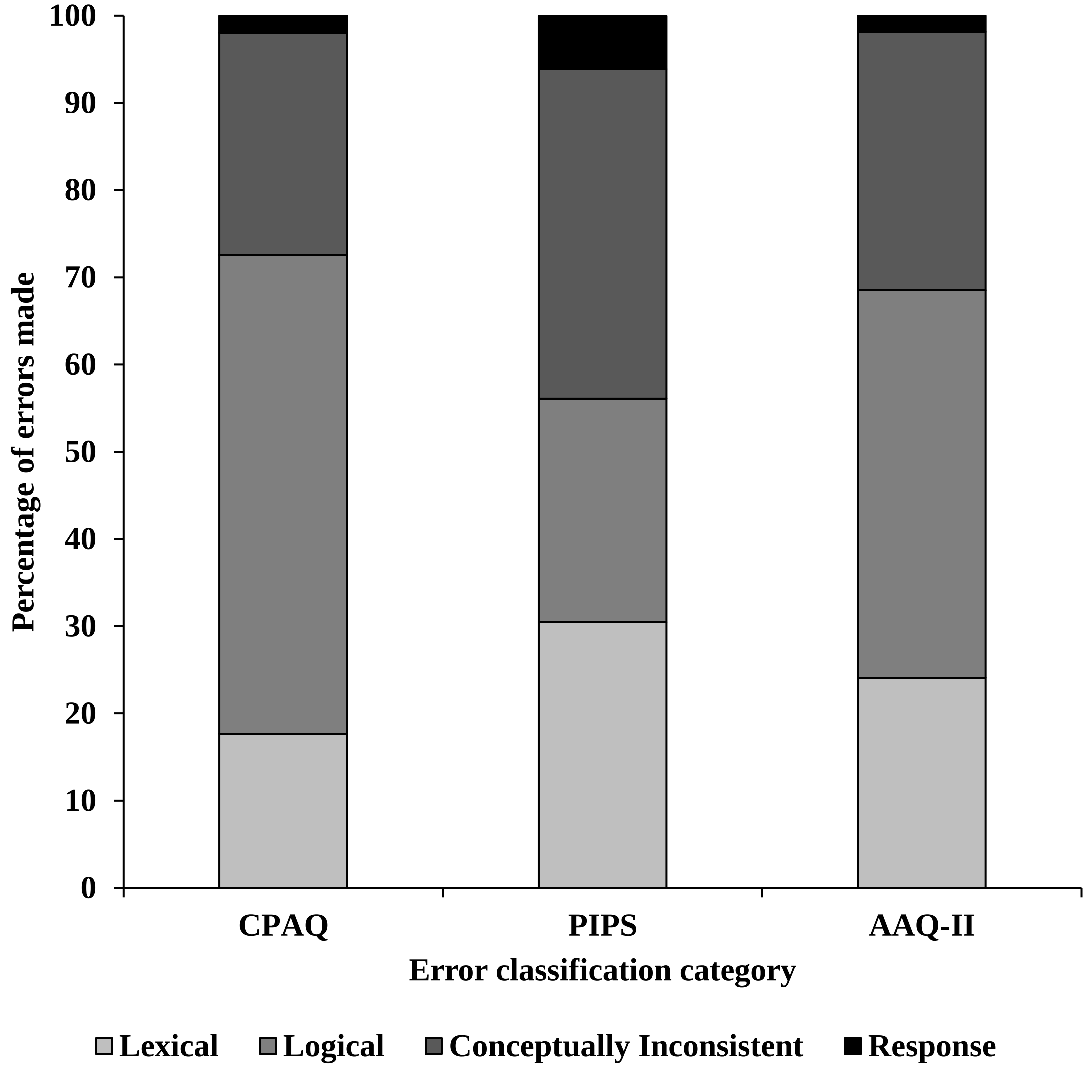
**Table 5.9 Percentage of respondents making errors (any error across each measure and at error-type level)**

| Measure | Error classification | Proportion of respondents making errors per measure (%) |
|---|---|---|
| CPAQ-8 | Lexical | *17.65* |
| | Logical | *54.9* |
| | Conceptually Inconsistent | *25.49* |
| | Response error | *1.96* |
| | *Total error* | *95* |
| PIPS | Lexical | *30.48* |
| | Logical | *25.60* |
| | Conceptually Inconsistent | *37.80* |
| | Response error | *6.1* |
| | *Total error* | *95* |
| AAQ-II | Lexical | *24.07* |
| | Logical | *44.44* |
| | Conceptually Inconsistent | *29.63* |
| | Response error | *1.85* |
| | *Total error* | *90* |

It is notable that across all measures, response errors were infrequent (1.85-6.1%) indicating that most participants did not make this type of error. The most commonly occurring errors made within the CPAQ-8 and AAQ-II measures were of a logical type, (54.9%) and (44.44%), respectively. Within the PIPS measure, conceptually inconsistent type errors were most commonly made (37.80%). It should also be noted that certain items accounted for a larger proportion of errors made and this will be further explored in the analysis.

The above data has also been presented in bar chart form (see figure 5.2) to clearly illustrate the frequency of each individual error type, comparatively across the three measures.

**Figure 5.2 Bar chart to illustrate frequency of error occurrence**



This visual representation of the data shows that across the measures, logical errors were most prevalent in the CPAQ-8 and AAQ-II and accounted for the largest proportion of errors made. Conceptually inconsistent errors were most prevalent in the PIPS. Again, the data shows that response errors were infrequent across the questionnaires.

## 2) Where are the errors made?

The second part of the analysis also related to the occurrence of errors across the ACT questionnaires. However, the focus here was to establish specifically where the errors occurred. In doing so this will tell us for each questionnaire and its individual items, whether the intended processes were understood by the participants.

### *Examination of individual items and error types per item*

The percentage of total error (out of all errors made) for each item, was calculated for each measure individually so that we could establish whether errors tended to be made on particular items. Tables 5.10, 5.11 and 5.12 demonstrate the frequency of error occurrence for all participants combined, across each item and each error category. Refer to appendix F for tables containing detailed data for type and frequency of errors made by each participant.

**Table 5.10 Frequency of error classification across error types and items (CPAQ-8)**

| ***CPAQ-8 Item:***<br>***Error Classification:*** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **Total error (% total error)** |
|---|---|---|---|---|---|---|---|---|---|
| **Lexical error** | 1 | 0 | 0 | 3 | 2 | 1 | 1 | 1 | **9 (18)** |
| **Logical error** | 2 | 2 | 2 | 9 | 0 | 0 | 0 | 13 | **28 (55)** |
| **Conceptually Inconsistent** | 1 | 2 | 1 | 8 | 0 | 1 | 0 | 0 | **13 (25)** |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Response error** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | **1 (2)** |
| **Total error (% of Total error)** | **4 (8)** | **5 (10)** | **3 (6)** | **20 (39)** | **2 (4)** | **2 (4)** | **1 (2)** | **14 (27)** | **51** |

**Table 5.11 Frequency of error classification across error types and items (PIPS)**

| ***PIPS Item:***<br>***Error Classification:*** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **Total error (% total error)** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Lexical error** | 0 | 0 | 0 | 1 | 1 | 2 | 2 | 0 | 4 | 8 | 0 | 2 | 1 | 1 | 1 | 2 | **25 (30)** |
| **Logical error** | 2 | 1 | 0 | 4 | 2 | 2 | 2 | 1 | 1 | 1 | 0 | 2 | 1 | 0 | 1 | 1 | **21 (26)** |
| **Conceptually Inconsistent** | 2 | 3 | 1 | 3 | 1 | 1 | 5 | 0 | 3 | 6 | 0 | 1 | 3 | 1 | 1 | 0 | **31 (38)** |
| **Response error** | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | **5 (6)** |
| **Total error (% Total error)** | **4 (5)** | **7 (9)** | **1 (1)** | **8 (10)** | **4 (5)** | **5 (6)** | **9 (11)** | **1 (1)** | **8 (10)** | **15 (18)** | **1 (1)** | **5 (6)** | **6 (7)** | **2 (2)** | **3 (4)** | **3 (4)** | **82** |

**Table 5.12 Frequency of error classification across error types and items (AAQ-II)**

| ***AAQ-II Item:***<br>***Error Classification:*** | **1*** | **2** | **3** | **4** | **5** | **6** | **7** | **Total error (% total error)** |
|---|---|---|---|---|---|---|---|---|
| **Lexical error** | 3 | 2 | 0 | 2 | 3 | 1 | 2 | **13 (24)** |
| **Logical error** | 2 | 2 | 7 | 3 | 5 | 2 | 3 | **24 (44)** |
| **Conceptually Inconsistent** | 1 | 1 | 5 | 0 | 6 | 1 | 2 | **16 (30)** |
| **Response error** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | **1 (2)** |
| **Total error (% Total error)** | **7 (13)** | **5 (9)** | **12 (22)** | **5 (9)** | **14 (26)** | **4 (7)** | **7 (13)** | **54** |

**One missing data value*

Findings showed that participants struggled with and made more errors on certain items. For example, CPAQ-8 items 4 and 8 were most problematic and accounted for 38% and 28% of the total errors made for this measure, respectively. PIPS item 7 (11% of total error) and item 10 (18% of total error) were most problematic. AAQ-II item 3 (22% of total error) and item 5 (26% of total error) produced most errors.

## Post-hoc analysis

After the data had been analysed in reference to the original hypothesis, further questions were generated from the data. For example, certain participants appeared to account for larger proportions of the total errors made. Additional post-hoc comparisons were completed in response to these findings.

### *Comparison of problematic question items*

Firstly, we looked to the previously identified, most problematic questions to see if there were any patterns in error responses across the participants. Table 5.13 demonstrates example responses for the two most problematic items (for each measure) and the error type they were categorised as. Reflections are offered on similarities observed between these responses. This enabled us to see where patterns may be occurring when errors are made. For example, comments were made on errors related to particular words on the items or whether respondents struggled to hold in mind certain aspects of an item. The aim of exploring the error patterns was to offer possible solutions on how we might resolve these (refer to Discussion section).

***Table 5.13 Example responses for most problematic items and reflections on similarities/ patterns of error occurring.***

| **Measure (Item)** | **Error type** | **Example responses** | **Commonalities between particular error types made across items** |
|---|---|---|---|
| **CPAQ (4)**<br>Before I can make any serious plans, I have to get some control over my pain. | ***Logical*** | *‘Yeah – (sighs) yes and no. Um, I make plans but I kind of – because you make plans knowing what you can do and you just try to leave what you can’t do until later. So four, that’s ‘Sometimes true’.* (Participant 11)<br>*‘.. it now takes priority planning about my pain and I know my levels, and what flares, so I would say – it’s difficult because I’m waiting for surgery- three.’ (Participant 7)*<br>*‘I feel that we can’t really make plans cause we just, it’s just you never know how you’re going to be on a day to day, be nice to but that’s almost always true, yeah.’ (Participant 14)*<br>*‘Always I try, but because everything comes frequently it happens, this is life, and I can’t do this, most of the time I think. I think ‘Very seldom true’, number two.’ (Participant 5)* | Respondents struggle to hold in mind all parts of the question; with a focus on the first part. For example, their answers refer to whether they ‘make plans’, how much they prioritise ‘planning about pain’ – or how able they are to make plans. They do not hold in mind how much *‘control over my pain’* is needed to make plans. |
| | ***Lexical*** | *‘No, not now. I can make a serious plan because a serious plan is something that’s going to be in the future which gives you time to prepare, so you don’t- before making the plan you don’t- that- that to me is a silly question, because why would you have to do that when- if you’re making a plan it’s in the future, so you’ve got time to control your pain without- I’d say that were ‘Never true’, number one.’ (Participant 12)* | Confusion over interpretation of the words ‘*serious plans’* – the respondents interpret this as meaning a ‘future’ plan. |

|  |  |  |  |
|---|---|---|---|
|  |  | *'I wouldn't know what it meant because um: 'Before I can make any serious plans –'What would be my serious plans? I don't know, um I don't know, that one's a bit confusing*- that would be a bit confusing for me because, um, in the *past what I've done is taken two tablets. It wasn't serious anyway, it was just a* thing that I did, *I don't know, I wouldn't know.' (Participant 10).* |  |
|  | ***Conceptually Inconsistent*** | *'I feel I already have some control over my pain. I've been doing this, I've been monitoring and maintaining my medication levels for a long time… so if I make any serious plans I already know where I'm at with my pain level and how to* control it- so I would say... I have control over my pain so this is very seldom true.*' (Participant 1)*<br>*'Well that's quite a difficult one cause it's like one that I've just answered. Before I can make serious plan, well I'm still, I can still make plans but I don't know what you mean by serious plans, serious plans, I don't understand the* question really. *Yeah I don't really understand the question before I can make serous plans, no, it's.*I have to get some control over my pain. *I'd say seldom true, yeah.' (Participant 15)*<br>**'***No, not now. I can make a serious plan because a serious plan is something that's going to be in the future which gives you time to prepare, so you don't- before making the plan you don't-* that- *that to me is a silly question, because* why would you have to do that when- *if you're making a plan it's in the future, so you've got time to control your pain* without- *I'd say that were 'Never true', number one.' (Participant 12)* | Often because of lexical/logical errors made, the respondent's answers do not get at the 'pain willingness/ acceptance' aspect of the question as intended i.e. because they answer in relation to whether or not they **have** attained control over pain.<br>Respondents end up scoring themselves as 'seldom true' – i.e. high psychological flexibility, when their answers actually reveal this may not be the case (seemingly low acceptance of pain). |
| **CPAQ (8)** | ***Logical*** | 'Bit *of a middle of the road one, sometimes it's true*, sometimes not *and so I'd go for a 4. I do worry about overdoing it because I do tend to again, take more painkillers than I probably should because I know I'm about to go and do something and then I go beyond what I feel comfortable with because the pains* | Participants often do not consider whether the thoughts are 'true' or not within their answers- more around how often they |

| | | | |
|---|---|---|---|
| My worries and fears about what pain will do to me are true. | | *not telling me to slow down. I do sort of have to plan where my limit is and stop before I get to where I think my limit is rather than going beyond it.'(Participant 3)*<br>'Yes. Always I have fears and lots of worries about future and my pain and what's happening to me gradually getting - yeah. 'Almost always true', number six.' (Participant 5)<br>*'Again sometimes I used to get carried away with my worries and fears about the pain and where it's going but I am trying to control it so erm 4.'(Participant 2)*<br>*'I don't really understand that one, if it's what I think it is then no not really, I worry that it takes over my life, so maybe, so I'm just going to go for number four for that one, I think. Yeah I do worry that if, how long it's going to go on for and if it's ever going to go away, cause I've tried a lot of things and I've tried a lot of avenues and I'm just reaching a brick wall really with it, so yeah.' (Participant 20)* | worry. Logical errors occur because the answer needs to include the concept of 'truth' of thoughts. |
| **PIPS (10)**<br>I put a lot of effort into fighting the pain. | ***Lexical*** | *'I would say this is true I do put a lot of effort into … not necessarily fighting my pain but trying to ignore my pain as best I can or in finding ways that I can cope with my pain and still do what I want to do. It's becoming more and more difficult because things are deteriorating at a rate that's beyond my control so at the moment I am finding things very difficult.'(Participant 1)*<br>*'I do, I do try to do things a lot of the time, um, that will help me to, help myself really. Um I just- I ask people as well, you know: "How do you cope? How do you cope?" I've been told a lot of things which I've took on board myself about how other people cope and "Where could I get more help?" And things like that as well. And, um just and helping myself as well but I didn't do myself any justice* | Participants appear to choose their interpretation of 'fighting pain' with numerous examples apparent.<br>The word '*fighting*' is often recast and therefore interpreted differently to the way we presumed was intended by the question. For example, it is changed to 'coping' with |

|  |  |  |  |
|---|---|---|---|
|  |  | *taking too many tablets, that’s the only thing. Yes. So, I think that would be almost true, ‘Almost always true’. (Participant 10)*<br>*I don’t know whether it’s fighting the pain or working with the pain…’ (Participant 15)*<br>*‘Well I do try to not, that’s what I’d like- to be pain free, but it’s not going to happen so I’d say, almost always true.’ (Participant 16)* | the pain, ‘working with the pain’ or being ‘pain free’. |
|  | ***Conceptually Inconsistent*** | *‘… I don’t know whether it’s fighting the pain or working with the pain, I do try and do things that will help me like exercise and things that I’ve been taught in physiotherapy, but the same again I can’t always do it, it just, it just depends on how I’m feeling that day as to whether I can do it. I mean at the moment I’m not able to get on to my exercise bicycle but there’s other things that I can do and I try to do them as regular as I can, so I would say that that’s almost always true.’ (Participant 15)*<br>*‘Yeah, you put a lot of effort when you have enough energy and you feel – you hope it’s positive, positive effect on you but, hmm, no, unfortunately I don’t put a lot of effort, um, just try to deal with and be there, that’s it. No, ‘Very seldom true’, number two. (Participant 5)* | Some of the lexical errors also lead to conceptually inconsistent errors. For example, misinterpreting the meaning of the word ‘*fighting*’ leads to participants scoring themselves ‘always true’ (high psychological inflexibility) when their answer reveals they are living in line with values. Not understanding this concept also causes errors the other way round- e.g. participant views fighting pain as a positive thing, however score reflects high flexibility. |
| **PIPS (7)**<br>I say things like ‘‘I don’t have any energy’’, ‘‘I am not well enough’’, ‘‘I don’t have time’’, ‘‘I don’t dare’’, ‘‘I have | ***Logical*** | *‘…Okay, that’s - all those are very different things Uhhh now I don’t know how to answer that one. I have said some of those but - ‘Sometimes true’ on that one? I don’t – like, that’s a lot of things to - and I feel like those are very- some of those are not conflicting but they’re bit some are linked to physical, some are linked to mental I think? So I – uh, sometimes’ (Participant 9)* | The structure of the item often causes confusion. There are too many components, leading respondents to go for a non-committal ‘middle answer’. |

|  |  |  |  |
|---|---|---|---|
| too much pain'', ''I feel too bad'' or ''I don't feel like it''. |  |  |  |
|  | ***Lexical*** | *‘No I never, I never say those, I don’t, not out-loud, but I might say it in my head, so that’s never true, I don’t say those things, out loud, right, I might think them but I don’t say them.’* (Participant 14) | The word ‘*say*’ is causing confusion- the intention of the question is about their ‘thinking/acting/doing/behaving’ as opposed to actually speaking the thoughts out-loud to other people. |
|  | ***Conceptually Inconsistent*** | *‘…definitely not me I try to keep what I’m feeling to myself with regards to my energy levels and the way I feel. I certainly don’t say I don’t care or I don’t have time. I’m very careful as to what I say especially to my family as they’re super sensitive to how I am that I wouldn’t really need to say it anyway and I desperately try not to be negative with them. So I would say this is number 1- never true.’ (Participant 1)*<br>*‘…I don’t say those things, out loud, right, I might think them but I don’t say them.’(Participant 14)* | Because of the lexical error (caused by the word *‘say’*) this item does not tap into the construct of experiential avoidance/ flexibility as intended. |
| ***AAQ-11 (3)***<br>I worry about not being able to control my worries and feelings. | ***Logical*** | *‘Yeah, I am worried a lot of the time. Um, pain and- and other stuff, you know, it messes up your brain so it’s difficult sometimes to distinguish between what’s painful or what’s mentally painful six.’* (Participant 11)<br>*This is probably true I do worry about things, but I don’t like to let anybody else know how I’m worrying or how I’m feeling erm I try to cover up a lot of things and I lie to my family and friends so they’re nowhere near aware of what my worries and feelings are so I would probably say this is true-number 6, almost always true.’(Participant 1)* | The logical errors tend to be made because they refer to the extent of their worry and do not include the idea of ‘control’. |

|  |  |  |  |
|---|---|---|---|
|  |  | *'I don't know whether I do or not. I don't think I worry that much really. So 'Seldom true' for that one.' (Participant 4)*<br>*'… I'd say ninety percent of the time I'm fine, it's just on a night if I've had a bad, painful headache that it's worry about – you're worrying about what's going to be and what's going to happen, but then obviously once you've calmed down you can get perspective quicker and realise that it's me being silly. Not silly but blowing it all out of proportion, so I'd say it's 'Seldom true'. (Participant 7)* |  |
|  | ***Conceptually Inconsistent*** | *'No, I'd say never true that one, I, I'm usually pretty good at controlling stuff like worries and control, and so I'd say never true.' (Participant 16)*<br>*'This is probably true I do worry about things, but I don't like to let anybody else know how I'm worrying or how I'm feeling erm I try to cover up a lot of things and I lie to my family and friends so they're nowhere near aware of what my worries and feelings are so I would probably say this is true-number 6, almost always true.'(Participant 1)*<br>*'I don't know whether I do or not. I don't think I worry that much really. So 'Seldom true' for that one.' (Participant 4)* | Many of these logical errors lead to conceptually inconsistent errors also. The respondents scores suggest high flexibility- yet answers reveal otherwise- e.g. they are 'good' at controlling worries (so actually high inflexibility- opposite of score given) or the concept of 'controlling worries' is not considered in their answer. |
| ***AAQ-II (5)***<br>Emotions cause problems in my life. | ***Logical*** | *'…if I've got something that's niggling I'll deal with it so, um – or I get prompted from my partner to deal with it, you know, um - so I'd say 'Seldom true' . (Participant 6)* | The concept of '*emotions*' is missed out often- focus is on how they 'deal' with problems. |

| | | | |
|---|---|---|---|
| | | '*Learning to control them and understand them I think is more to the point. When I let my emotions run - it doesn't cause problems all the time, I don't suppose but I'd say that was 'Sometimes true' (Participant 12)* | |
| | ***Lexical*** | *'Emotions' I don't think I'm over emotional really. I don't know what it means by 'emotion', does it mean: Do I think I'm over emotional? I don't know. I think, um, I think I'm a worry – worrier, but is that emotion? I don't know. I think, uh - I think 'worry' perhaps causes* problems *but – in my life, I don't know if – is 'emotion'* - does that mean- *it means 'sadness, doesn't it, 'emotion'?* Sadness. *No, I think – sadness then makes you feel a bit depressed though, doesn't it? Causes problems. Um, I'm not sure about that question.'* (Participant 4)<br>'*…no,* my issues *don't cause it, they just – again it's a difficult one because is it the emotions from the pain or the emotions from the effects of the pain? So it's difficult to put your thumb on. Yeah, they cause problems, but not ones that I hope I can't deal with. So, um, four.' (Participant 11)* | The word *'emotions'* is causing difficulties- not sure what is meant by this term. |
| | ***Conceptually Inconsistent*** | '*I am emotional erm, but generally in a positive way. I'm constantly telling my kids how much I love them and my husband. I try and be very positive with my emotions although I do blubber a lot if I see anything sad or I'm aware of anything sad and so I try and avoid those sort of scenarios. So I wouldn't say they cause problems in my life because I try and share the positive ones and cover up any negative ones and so would say this is not true, number 1-never true.'(Participant 1)*<br>*'…emotions, sometimes, because you talk all you want and say well I'm in pain but, and people say oh yeah I know, but they don't really know, so I'd say sometimes true.' (Participant 16)* | Actual levels of experiential avoidance are often not reflected in scores. Respondents think about 'emotions' and experiencing them in relation to what other people think/see of them. Others do not 'see' them express emotion or understand it – therefore an answer reflecting level of acceptance of difficult emotions is not given. |

### *Comparison of total errors made by each participant*

Secondly, we explored the data for any emerging patterns which might explain the frequency of errors made between the participants, through a comparison of 'TOPF scores' against total errors made. Table 5.14 shows the percentage of total error that each participant contributed and their TOPF scores.

**Table 5.14 Percentage of total errors made by each participant and TOPF scores.**

| Participant | Total errors made | % of total errors made | TOPF score |
|---|---|---|---|
| 16 | 21 | 11.1 | 78 |
| 3 | 16 | 8.5 | 117 |
| 15 | 15 | 7.9 | 96 |
| 1 | 14 | 7.4 | 109 |
| 11 | 13 | 6.9 | 89 |
| 5 | 11 | 5.8 | 94 |
| 10 | 10 | 5.3 | 97 |
| 6 | 10 | 5.3 | 77 |
| 19 | 10 | 5.3 | 95 |
| 12 | 9 | 4.8 | 110 |
| 9 | 9 | 4.8 | 106 |
| 14 | 9 | 4.8 | 93 |
| 18 | 7 | 3.7 | 123 |
| 4 | 7 | 3.7 | 113 |
| 7 | 7 | 3.7 | 94 |
| 20 | 7 | 3.7 | 97 |
| 8 | 5 | 2.6 | 123 |
| 13 | 5 | 2.6 | 116 |
| 2 | 2 | 1.1 | 117 |
| 17 | 2 | 1.1 | 97 |

The data here suggests that certain participants accounted for a larger percentage of the errors made. It was notable that the participant with the lowest TOPF score (78) contributed towards the highest percentage of total errors made on the ACT measures (11.1%). Correlation coefficients were calculated to explore the relationship between errors made and TOPF scores. Results of the Pearson correlations indicated that there was a significant negative association between errors made by the participant and TOPF score, *r* = - .44, *p* = .05.  This negative relationship indicates that as TOPF scores increased (indicating a higher intelligence) the number of errors made by the participant decreased, which is what we would expect. There was a non-significant correlation between TOPF scores and lexical errors made by each participant (*r* =-.08, *p* = .75). This indicated there was not a relationship here. There was a significant negative association between logical errors made by the participant and TOPF score, (*r* = -.56, *p* = .01.), suggesting lower TOPF scores were linked to that person making more logical errors.There was a non-significant correlation between TOPF scores and conceptually inconsistent errors made by each participant (*r* = -.32, *p* = .16) indicating no relationship here. The ‘response error’ category was not explored as so few errors were made here it would not have been a meaningful calculation.

### *Number of sessions*

Lastly, we explored the data for differences between the stage of ACT therapy participants were in (early or late) and the total errors made. The aim here was to see if the number of sessions (therefore exposure to ACT therapy and its concepts) impacted on the participant’s ability to understand the items. A Spearmans' correlation was calculated as the stage of therapy variable was ordinal (either early or late stage of therapy). Results showed that there was no significant correlation between total errors made and stage of therapy (*r* = -.08, *p* = n.s.). This would suggest that increased exposure to the ACT model did not impact on the amount of errors made.

# CHAPTER SIX

## Chapter overview

Chapter 6 forms the discussion for the main study (Study Part Two). This chapter will firstly highlight the main findings and theoretical implications in the context of the wider literature. It will then move onto strengths and limitations, clinical implications of important findings and suggestions for further research will be offered.

## GENERAL DISCUSSION

### Review of background and aims

The present study provided a much-needed investigation into the content validity of ACT questionnaires for people experiencing chronic pain. Participants' understanding of the questionnaire items was examined using cognitive interviewing methods and a number of items were found to lack aspects of content validity. The findings are important for future refinement of questionnaires and help to establish where there may be patterns in misunderstanding the questionnaire items. The significance of these findings in relation to the existing literature will now be discussed, beginning with a summary of the research and clinical problems which stem from the validity issues in ACT questionnaires.

As previously mentioned, research evidence supports the efficacy of ACT for the highly prevalent condition of chronic pain (Vowles & McCracken, 2008; Vowles, McCracken, & O'Brien, 2011; Hughes, Clark, Colclough, Dale & McMillan, 2017). ACT aims to improve psychological flexibility through engendering improvement in the six constituent sub-processes (Hayes et al., 2006). Change in these processes in clinical and research settings is measured using self-report questionnaires such as the CPAQ, PIPS and AAQ-II. If respondents do not understand the questionnaire items this directly impacts what can be inferred from their responses and this has widespread implications for research and clinical practice. ACT clinical practice and research relies on our understanding of the route by which changes are made; a result of reliably and validly assessing ACT constructs to determine meaningful change.

As previously referred to, establishing the validity of process measures is crucial for continued ACT research and is an important part of the research agenda for ACT. It has been suggested that the functional contextualist foundations upon which ACT arose are *"fundamentally oriented toward the development of analytically adequate processes of change"* (Hayes, et al., 2012, p.13). In order to 'predict-and-influence' or target the processes of interest we must understand which aspects of clinical intervention work and *how*. The advancement and refinement of ACT theory itself through process research depends on process measures which can accurately assess these mechanisms of change. ACT measures must represent change in psychological flexibility and the determining factors. As a result, clinical psychologists working within chronic pain services may be better equipped to effectively capture change in clients undergoing ACT for chronic pain. Clinicians need to understand more about the process of change in a psychological intervention to enable them to adapt treatment methods during therapy and to develop more effective approaches with improved outcomes for individuals experiencing chronic pain. We need to understand how ACT constructs contribute towards specific strengths and deficits in psychological inflexibility. This depends on people being able to interpret and understand the questions in the way intended by the questionnaire design.

The Hayes et al. (2012) review also made some important recommendations for future research on processes of change and it is suggested that qualitative research is as equally necessary as correlational based studies. Their concerns regarding the lack of alternative, contextually-informed experimental methods to supplement the existing psychometric studies of process measures echoed those of the present study. Previous research has relied on these correlational methods and although these indicate possible issues with the content validity of ACT process measures this has never before been extensively tested using the empirical methods of the current study (Francis et al., 2016; Reneman et al., 2010; Wolfgast, 2014). Research *directly* investigating how individuals interpret and then respond to ACT questionnaires is needed. This will tell us *how* individuals understand the items in the questionnaires and what might influence their response. It will also help us to understand where there may be patterns of errors, which will be useful for tailoring and improving existing questionnaires to improve their content validity.

The present study attempted to address these concerns. It aimed to investigate the content validity of the commonly used ACT questionnaires in individuals experiencing chronic pain. More specifically, it aimed to establish whether respondents understand the question items (without making errors) therefore capturing ACT processes as intended by the original questionnaire author.

Findings from a preliminary study (Study Part One) aimed to establish the most frequently used ACT process measures in clinical research and practice. These findings informed the measures selected for this main study.

This research was necessary because at present we rely on self-report questionnaires to assess whether an individual has made changes during ACT therapy. This is problematic for four main reasons previously summarised from the research literature:

1) Respondent understanding and interpretation of ACT process measures has not been explored in a comprehensive and empirical study. The evidence base to date has focused on quantitative, correlational analysis.
2) If individuals do not understand the questions on process measure questionnaires and hence do not respond in the manner intended then what can be inferred from the results has limited utility.
3) This has implications for ACT process research- as we cannot be certain of the route by which changes are made or the contribution of specific strengths and deficits in key subcomponents. This is needed to refine the theory itself.
4) Lastly, without understanding how a treatment works through measuring changes in process variables, clinicians will not be able to effectively adapt treatment methods during therapy (knowing which components to target and in what sequence).

This ACT questionnaire validation study hopefully contributed to a more comprehensive understanding of where problems lie in the use of ACT questionnaires and in turn will influence the ability for future research to evaluate the processes occurring during treatment.

## Summary of findings

As previously discussed in Chapter 4, findings from Study Part One revealed the ACT process measures to be of most regular use (clinically and in research) included the CPAQ PIPS and AAQ-I/II. Findings from the main study (Study Part two) are now summarised.

### *Frequency of error occurrence*

In relation to the original research hypotheses, results showed that errors were consistently made as individuals answered the questionnaire items. Problems in response processes occurred amongst the three distinct error categories: lexical errors, logical errors and conceptually inconsistent errors. During the main analysis a fourth error 'type' also emerged: 'response errors'. The most frequently occurring errors were those of a logical type. For example, within the CPAQ-8, logical errors were most prevalent and accounted for *54.9%* of total errors made. Within the AAQ-II logical errors accounted for *44.44%* of the total errors made. Within the PIPS measure, conceptually inconsistent type errors were most commonly made, accounting for *37.8%* of the total errors made. Errors from the 'response error' category were infrequent, accounting for no more than 6.1% of total errors made. However, this was to be expected as they are technically a variant of logical error. When the three measures were compared, although we observed a trend for most frequently occurring error types, the differences between the three measures was marginal. Calculations for the proportion of respondents making any errors across each of the measures varied from 90-95%

### *Identifying particularly problematic items*

The number of errors made varied greatly between the individual questionnaire *items*, with certain items accounting for a larger proportion of errors made. Results showed that participants struggled to understand some of the items more than others and subsequently made more errors. A range of problematic questions were compared across individual participant responses to investigate whether any patterns emerged in the error classification categories (Table 5.13). Patterns in error making across each category were revealed and are summarised below:

-*Lexical errors*: Were often due to misinterpretation of a word or several words on the item. This led to the word being recast, thus interpretations and responses were different to the way intended by the questionnaire author.

-*Logical errors:* Respondents had difficulty answering these questions. Whilst they understood the words, they still had trouble holding in mind all parts of the question, especially if it was comprised of multiple components which required them to connect different concepts.

-*Conceptually Inconsistent errors:* Respondents did not answer the item in relation to the key ACT constructs as intended. Although *technically* answering the question correctly (no lexical/logical errors) their answer was not an accurate reflection of the ACT concept being measured. Often, these items were interpreted in the opposite direction of the ACT process under study. For example, a participant might have scored themselves as ‘seldom true’ (to indicate high psychological flexibility) when their answer actually revealed this may not be the case (e.g. seemingly low acceptance of pain, therefore reduced psychological flexibility). Although these errors occurred in isolation they were also often linked to the lexical/logical errors also made on that item. Importantly, these findings imply that there are issues related to aspects of content validity on all three of the ACT measures investigated in this study, specifically around the understandability/ interpretability of certain items. If individuals do not understand the questions on process measure questionnaires and hence do not respond in the manner intended, then what can be inferred from the responses is limited. Thus, the questionnaires may not accurately determine meaningful change.

During the main analysis (see chapter 5) we also looked for any specific patterns in error occurrence for each of the most problematic items to understand what the errors were attributed to. These included the CPAQ-8, items 4 and 8; PIPS, items 7 and 10 and AAQ-II, items 3 and 5. Potential causes of these errors were discussed in terms of reoccurring patterns between individual participant answers. These findings were used to develop and make suggestions regarding the future refinement of the questionnaires and will later be discussed in relation to future questionnaire development (refer to Table 6.1).

### *Impact of individual differences on error rates*

Lower TOPF scores were significantly associated with more frequent response errors. These findings have important clinical implications as they suggest that individual differences in a respondent’s intelligence could impact on how that person understands and interprets the items. Therefore, we cannot always be certain that variability on scores yielded for the psychological flexibility process measures is accounted for by clinical change or these individual differences in ability to interpret questions. It should be noted that this relationship was not significant when comparing TOPF scores against lexical or conceptually inconsistent errors. There was however, a significant relationship between logical errors made by the participant and their TOPF score, suggesting lower intelligence did influence the rate at which individuals made this type of error. Again, if logical errors are strongly correlated with intelligence then this has implications for what we can infer from the measures. The findings suggest that individuals with lower intelligence were more susceptible to making errors on certain items which were logically more difficult to understand. The TOPF is an assessment of vocabulary and therefore suggests that individuals with low scores may also have difficulty understanding and interpreting the meaning of certain process measure items. These findings also resonate with a quote about ACT process measures, gathered in Study Part One (Clinician survey):

> *‘They are too wordy and they are well above the average adult national literacy level. It's bizarre that for such an experientially based therapy, the questions are so wordy! We need simple tools which people can easily make sense of.’* (Example ACT clinician survey response, from Study Part One)

This calls into question the content validity of these items. If they cannot be understood by individuals completing them we cannot be certain they are assessing the mechanisms of change in the way intended. These issues may stem from the questionnaires being composed by researchers who may design the wording, concepts and structure of items in a way only understandable to individuals of a certain level of intellect.

Additional post-hoc analysis revealed that the stage of ACT therapy participants were in (early or late) did not appear to impact on error rate. This might indicate that increased exposure to the ACT model (and thus its concepts) did not increase their ability to understand the items. However, these results should be interpreted with caution for two reasons: 1) the overall sample size may be too small to detect a correlation; 2) the proportion of participants recruited from early and late stages of therapy was unevenly distributed with the majority (70%) coming from later stages of therapy and thus more exposed to the model.

### *Reliability of the problem classification matrix*

In conjunction with cognitive interviewing methodology (Beatty & Willis, 2007; Willis 2004) a problem classification matrix, based on the work of Conrad and Blair (1996), Drennan (2003), Van Oort (2011) and McCorry et al. (2013) was developed to identify the difficulties respondents experience in understanding and accurately completing questionnaires. It was this framework which was used to systematically investigate *how often* and *where* the errors occurred in respondent interpretation of items. Preliminary analysis through a series of interrater reliability checks confirmed some initial issues with the frameworks reliability. However, several attempts were then made to develop acceptable reliability in using this framework to rate errors.

## Findings in the context of the wider literature

### *Overview*

The study findings have revealed issues with ACT process measure content validity, mainly issues surrounding the understandability of the items to respondents. This has also raised questions with regards to the ability of the measures to capture the psychological flexibility processes they intend to. Designing a questionnaire to capture dynamic psychological processes is a complex task and it is an equally multifaceted task for clients to respond in the way the questionnaire author intended. It entails many levels of understanding to answer in a way which corresponds with the intended constructs. It requires respondents to understand the language of the item itself (lexical understanding), and the structure of the item, whilst being able to hold in mind all of its parts (logical understanding) and finally, being able to answer the question in way which reflects the ACT concept itself, e.g. not in relation to other distress variables.

The preliminary study (Study Part One) provided rationale for the selection of ACT measures to be included in the main study. It also gathered qualitative data from ACT clinicians around their experiences of clients completing the questionnaires during therapy, for example:

> *"… Patients really struggle with the language [on the ACT measures], especially the double negatives. Many of them only partially complete it and ask for help on a number of the questions."* (Example ACT clinician survey response, from Study Part One)

Interestingly, the main positive themes identified by clinicians around the usefulness of questionnaires were around their '*ease of administration*' and being '*reasonably sensitive to change.*' The more negative themes corresponded with the research literature surrounding the validity of these questionnaires. Clinicians suggested that patients did not understand the item concepts, language and wording and this led to completion errors. This idea echoed the take home message for the main study findings: that we cannot directly infer (despite the item appearing to measure clinical change) that respondents have in fact understood and therefore answered in relation to the processes the measure was originally designed to capture. These findings will now be discussed in the context of the existing research literature.

## *Reliance on self-report measures to assess change*

As mentioned in the introduction, a fundamental issue common to all ACT process measures is their reliance on self-report. The problem revolves around the fact that we are attempting to measure a *'dynamic and shifting psychological process with a static and global self-report measure'* (Wolfgast, 2014, p. 838). Therefore the questionnaires are susceptible to demand characteristics and crucially for the present study they rely on respondents having a certain level of orientation to the ACT model (Levin, Herbert & Forman, 2017). Despite having being exposed to the ACT model through individual or group therapy, all participants in the current study made errors in completing the items.

### *Main study findings in relation to: Content validity of the ACT measures*

A predominant focus of the research literature to date has been on issues with certain aspects of ACT process measure validity. For example, the discriminant validity of the AAQ-II process of experiential avoidance was found not to be distinct from distress variables such as negative effect (Gámez et al., 2011). Reneman et al. (2010) also found issues with the reliability and validity of various ACT measures when internal consistency, criterion validity and construct validity were all psychometrically assessed. More promising findings have attempted to purport the validity of ACT process measures. For example, the construct validity of the (CPVI) has been supported with its significant correlation with other measures of avoidance and acceptance of pain (McCracken & Yang, 2006). The CPAQ-8 has been found to repeat the pattern of correlations and predictive validity found in the CPAQ full version – the two subscales were found to predict pain-related disability and emotional distress (Fish et al., 2010). Yet, it was suggested that further studies are required to investigate the extent to which this measure reflects the ACT construct/process variable of acceptance of pain, in a clinical setting.

The common difficulty with existing literature assessing the validity of these questionnaires is that none have specifically and comprehensively addressed the issue of content validity. The present study was the first empirical study of these ideas. The focus of validity research has been on correlational studies which do not tell us how individuals interpret and respond to ACT questionnaires. This is essential to investigate potential content validity issues surrounding the understandability of items. Although think-aloud methods have been used to explore this criteria of content validity in other health-related outcome measures (McCorry et al., 2013; Van Oort, 2011) this methodology has not been used to test ACT measure validity. Furthermore, no research on content validity of ACT measures has utilised a standardised tool (such as the problem classification matrix; Conrad & Blair, 1996) to objectively classify errors made in completing items.

The Acceptance and Action Questionnaire has been extensively researched with many findings demonstrating poor validity (AAQ-II; Bond et al., 2011) (Gámez et al., 2011; Rochefort et al. 2018; Tyndall et al. 2019; Wolfgast, 2014). Interestingly, although error rate differences were marginal between the measures, the present study also found the AAQ-II to be the most problematic questionnaire, with logical errors most prevalent here.

A key critique of this measure has been over its lack of discriminant validity; a key finding throughout the literature is that it appears to measure general distress outcome variables e.g. difficult emotions, rather than specifically target psychological inflexibility (Wolfgast, 2014; Tyndall et al., 2019). However, perhaps more relevant to the findings of the present study was the suggestion made by Wolfgast (2014) that these issues may be specifically related to the wording of AAQ-II items. Whilst the correlational study conducted by Wolfgast (2014) could only speculate regarding the understandability of these items, the present study's findings for the presence of lexical, logical and conceptually inconsistent errors confirmed there are issues with this measure's content validity. Our findings suggested that some of the items on the ACT measures failed to meet an important criterion for content validity previously described by Terwee et al. (2007); the interpretability and readability of items. For the present study this was most evident in the logical and lexical errors made by respondents and was further examined through additional analysis of error making patterns amongst the individual items. Interestingly, conceptually inconsistent errors were often linked to the presence of logical and lexical errors (see Table 5.4 for examples).

However, the present analysis revealed that the classification of 'conceptually inconsistent' errors was more complicated. Lexical and logical errors were relatively straightforward to categorise whilst conceptually inconsistent errors required more thought. The initial interrater reliability tests gave us some indication of this; during the first two checks overall agreement between raters for error agreement *(was an error present?)* was relatively high compared to agreement for error classification *(which error category is that item in?).* An influencing factor in these higher rates of disagreement between 'type' of error category may have been the lack of agreement when classifying 'conceptually inconsistent' errors. This was certainly evident during the main analysis and will be demonstrated using the following example response for AAQ -item 3 ('I worry about not being able to control my worries and feelings'):

> *"…What I tend to do is if I'm worried about something um, if it's something that keeps biting me on the arse, say I would, um, acknowledge the fact that that's, um – what that's doing and I'll know that it's something I need to deal with, um normally I would set time aside – like, <u>I'll give it ten minutes worrying time and then I'll just say: "Right –" Literally, I used to be a mental health nurse, so I used to set an alarm, you know, after ten minutes:</u> 'That's it, you've done the worrying for today, now move on.' That's*

> *what I would do, so – so I don't really worry about not being able to control them because, um – so I would say that's 'Very seldom true.'(Participant 6)*

Here an argument could be made for the presence of a conceptually inconsistent error. The respondent is suggesting they use their 'worrying time' as a strategy to make the difficult thoughts and feelings go away. This would indicate a degree of experiential avoidance/ control, however the actual score of 'very seldom true' given would suggest the opposite (high psychological flexibility). Therefore, a conceptually inconsistent error could exist because the final score is an inaccurate representation of the construct intended. *However,* understanding this answer as the respondent using 'worrying time' as a control strategy may be an assumption made by the rater as there is only this data to go on. An alternative interpretation may be that the respondent 'acknowledges' the worry and they make ten minutes to explicitly focus on unpleasant thoughts and emotions. After this they say to themselves 'now move on', which is more in line with the concept of acceptance/ defusion from worrying thoughts and 'moving on' to live in line with their values. We cannot be certain from this answer what the function of the 'worrying time' is. In other words, we rarely know the function of the behaviour. This pattern was evident throughout the analysis when items did not elicit the functions of behaviours, they could not be coded as conceptually inconsistent. For example, when respondents revealed what they *did* yet it was more difficult to elucidate a *reason* for this unless it was explicitly stated in their answer. Although there were many occasions when it was clear than the participant made conceptually inconsistent errors independently of logical and lexical errors (see Table 5.4 for examples) and this had clear implications for the content validity of those items, we have to take results for the conceptually inconsistent errors tentatively as they were difficult to categorise and potentially unreliable.

The process of deciding upon the presence (and type of) error revealed the complexity of the decision-making being undertaken by the participant. As mentioned, this was mirrored by the discussions in the research team when agreeing on error type. When categorising 'conceptually inconsistent errors' it became apparent that on many of the ACT questionnaire items only the *traits* of behaviours could be inferred and not always the actual *function* of the behaviour. These findings echoed the literature regarding this dilemma for ACT self-report questionnaires; we can rarely know the function of the respondent's behaviour. For example, Tyndall et al. (2019) suggest that issues occur when processes are confounded with traits, i.e. it is difficult to determine whether psychological flexibility is a trait or an outcome measure. This means that we cannot always determine whether an individual's response reflects their degree of psychological flexibility or something else such as unwanted emotions or worries. This was mirrored in the present study when we also experienced difficulties in assessing whether a response indicated level of psychological flexibility or indeed another non-related construct.

Although not the main focus of this thesis, an additional aspect of content validity, the ability of the measures to capture psychological flexibility processes, was also partially explored. This criterion of content validity matches that proposed by Terwee et al. (2007), i.e. the extent to which the measures intended concepts are being measured. Similar findings are mirrored in the research literature, for example findings suggesting the AAQ-II lacks discriminant validity as factor analyses reveals it fails to distinguish between distress items and ACT construct (acceptance) items; hence the constructs did not target what they intended (Tyndall et al., 2019; Wolfgast, 2014). This resonated with results from the current study, where there were many examples of logical errors leading to that item actually becoming a measure of psychological distress. This can be seen in the following example on the CPAQ- item 8 ('My worries and fears about what pain will do to me are true'):

> *'Yes. Always I have fears and lots of worries about future and my pain and what's happening to me gradually getting - yeah. Uh 'Almost always true', number six. (Participant 5)*

Here, the participant only refers to the extent of their 'worry' and appears to not consider the 'truth' aspect of the question. This is akin to issues with the content validity of all ACT questionnaires in the present study where distinct ACT processes were potentially merged with other outcome variables, such as undesirable feelings and thoughts. This could be problematic as ACT treatments are designed to make changes with regards to specific ACT constructs, which although related to, are theoretically separate from distress constructs. Merging the distinct ACT processes with other outcome variables and potentially not capturing all constructs within psychological flexibility may create difficulties in interpreting ACT measures. A valid measure must reflect individual changes in psychological flexibility as a whole, whereby any improvements made amongst the core ACT constructs are captured.

Studies which have specifically focused on content validity (thus the items themselves) have been limited to determining whether the intended concepts conflate with other non-intended variables. For example, the content validity of the ACT questionnaires has been contested in relation to their emphasis on single sub-processes whilst neglecting others (Francis et al., 2016). Although purported as a unidimensional measure of psychological inflexibility (Bond et al., 2011), Francis et al. (2016) suggest that the AAQ-II focuses on experiential avoidance and the fusion process and does not capture changes in contact with the present moment, values or committed action. They acknowledge that items sometimes look as if they assess, for example, present moment awareness processes; ‘My painful memories prevent me from having a fulfilling life’. However, it is suggested that although we can make assumptions from this question about a respondent’s ability to maintain awareness in the moment, this process is not a direct focus of the item. As previously discussed, these kinds of studies do not assess the interpretability/ readability of items within the target populations and this has been largely neglected in the literature to date.

Findings from the present study also raise important questions about the means by which questionnaire validity is established. As Strauss and Smith (2009) noted, an established knowledge base is required in order to validate scores on measures, i.e. to assess construct validity we test whether the measure of a psychological construct relates to measures of other constructs defined by the theory. To assess the validity of new measures’ content or predictive validity, the correlational studies prevalent in the ACT process measure evidence base have also done so by comparing scores on newly developed measures with those of measures already established. Importantly however, the present study has shown that this process of assessing validity is problematic. Using other ACT measures which assess similar concepts may have similar problems, i.e. if they are shown to correlate, it might just mean they are as equally imprecise. As has been investigated in the current study we need to explore other methods of assessing validity such as coding the language used by respondents.

To summarise, the difficulties we observed through error rates within each of the ACT measures are related to shortcomings in their content validity. The intended processes were not always captured due to the respondent’s experience of completing items, which ultimately results in items not being interpreted in the way the author intended. This has far reaching implications for research and clinical decision making as noted by Levin, et al. (2017). They highlighted a lack of research to date on what *"pathological sub-processes of inflexibility contribute to what specific presentations" (p.415).* It is suggested that this is direct consequence of the validity issues with ACT measures. Without being certain of what they are actually measuring, we cannot know if psychological flexibility is the key pathological process occurring or if other variables account for changes made in therapy in empirical investigations. Therefore, we remain unclear which ACT components are most useful for treating specific psychological problems.

### *Study findings in relation to: Functional Contextualism and RFT*

As previously mentioned, ACT is underpinned by the behavioural theory of human language and cognition (RFT; Hayes, Barnes-Holmes, & Roche, 2001) which in turn rests on the philosophical foundations of functional contextualism (Hayes, Hayes, Reese, & Sarbin, 1993; Biglan & Hayes, 1996). We previously discussed the dilemma related to our reliance on measures which confound processes with traits. Psychological inflexibility is a trait-like construct rooted in functional contextualism; it is dynamic and contextually dependent (Hayes, 2004). This causes difficulties in assessing whether an individual's response reflects their degree of psychological flexibility (i.e. their behaviours or something else such as unwanted emotions or worries). Our results showed that on many of the ACT questionnaire items we were unable to elicit the functions of behaviours, only the traits. For example, in the example provided above (p.99, Participant 6, AAQ-II item 3), 'I worry about not being able to control my worries and feelings', their response appeared to target traits (this person spoke about using a control strategy to alleviate worry). However, this did not assess whether what was done in this situation actually worked for the person given their own goals and values (their response also alluded to the use of this strategy as helping them to 'move on' from worrying). Functional Contextualism is about behaving in ways which serve one's values, therefore understanding the functions and effectiveness of behaviours is crucial. For example, strategies such as suppression or problem-solving to avoid unpleasant thoughts become problematic only when they are ineffective or limit progress on values. Findings from the present study also indicate issues with the ability of process measures to consider actions independent of the context in which they occur. Although we can comment on these potential validation issues, the focus of this thesis has been on the extent to which people understand the questions as intended so that they can answer the question asked. Whilst it is relevant to consider issues such as their consistency with functional contextualism, the purpose of this study was to ascertain whether participants can respond appropriately. Answers to self-report process measures will always be subjective and based on the respondent’s own reference points. Thus, a challenge for the development of future questionnaires is for the designer to anticipate the mind of the respondent. In addition to the solutions offered in the latter part of the discussion we can also turn to RFT to help in future questionnaire design. This theory may help us to develop more consistent and precise measures if we are able to think in terms of basic behavioural processes.

## Strengths and limitations

The limitations specific to the problem classification methodology have already been described in detail above. The strengths and limitations for the whole study will now be discussed.

### *Issues with categorising conceptually inconsistent errors*

To conclude, there were certain limitations to determining content validity of ACT questionnaires (due to difficulties in categorising conceptually inconsistent errors). In ACT the function of behaviour is important in assessing what is and is not psychologically flexible. However, as respondents often did not discuss the purpose, or consequences of their behaviours when responding to process measure items it often was not possible to be certain whether responses were conceptually inconsistent. An additional point to consider was raised by Francis et al. (2016). They highlighted difficulties with directly assuming the presence of psychological flexibility for negatively-valenced questions. For example, in the current study conceptually inconsistent errors were categorised when it appeared that the item had been interpreted in such a way that it became the opposite of the ACT process under study. Kashdan and Rottenberg (2010) (as cited in Francis et al., 2016) purported that we cannot take for granted that psychological flexibility and inflexibility are exact opposites i.e. a bipolar construct. In other words, we cannot assume that a response indicating a lack of psychological inflexibility automatically means that individual is psychologically flexible. This represents a possible confounding factor for the present study and has implications for what we can infer from our error classification results. This has already been taken into consideration however, in our interpretation of the analysis.

### *Lack of generalisability to ACT- naïve respondents*

Although the stage of therapy at which participants were at when they completed the study was considered, the sample focused only on participants who had experienced at least one therapy session. It could be inferred that to some level all participants had been socialised to the ACT model. Therefore, we did not investigate whether the questionnaires were comprehensible to ACT-naïve respondents. This was beyond the scope of the thesis and not attainable due to the recruitment process involved in identifying potential participants. Although no significant differences were found between those participants who were at the early stages of therapy (thus less exposed to ACT concepts) and those later in therapy, this might be down to the small sample size used and lack of statistical power to determine an effect. Furthermore, the sample was disproportionate with more participants recruited from later in therapy. One might also argue that if the respondents were more aware of what the questionnaires were targeting this may give rise to more response biases as they attempt to get the answers ‘right’.

Additionally, the participants recruited to the study were representative of a wide range of chronic pain ‘types’. There was also a reasonable spread across gender and age group amongst the sample.

### *Limited to testing certain ACT measures*

Although Study Part One findings demonstrated that the CPVI is used regularly across research and clinical practice this questionnaire was omitted from inclusion in the main study. The rationale for this was mainly due to differences in formatting compared to the CPAQ-8, PIPS and AAQ-II. The CPVI asks respondents to rate items on a scale ranging from 0 (not at all important / successful) to 5 (extremely important / successful). Two primary scores are then obtained: a mean success rating, and a mean discrepancy rating, which is the mean of the differences between importance and success. The selection of questionnaires also needed to also consider the demand on participants. The study design took into consideration the process of cognitive interviewing being time consuming and the potential for fatigue and discomfort, particularly in the population to be recruited.

### *Counterbalancing error*

During the cognitive interviewing task the ordering of individual questionnaire items had been counterbalanced by questions instead of by the full questionnaire as intended. This will not however impact on the data but did mean we could not comment on errors based on repetition (linked to the ordering of the items). For example, one participant attributed their ‘middle score’ on the Likert subscales being due to feeling that they had provided too many high scores for previous answers. They suggested that their middle score of 4 was ‘for a change’. This was also an indication of another possible confounding factor mentioned previously in relation to self-report questionnaires; that of demand characteristics.

### *Novel methodology*

A significant strength of the present study was that it was the first comprehensive empirical study of these ideas. Previous literature has identified a lack of ACT questionnaire construct validity (Wolfgast, 2014). Attempts have been made to determine construct validity as a whole through tests on discriminant, internal consistency and criterion validity (Reneman et al., 2010). Studies specifically focusing on content validity (thus the items themselves) have been limited to determining whether the intended concepts conflate with other non-intended variables, whilst the interpretability/ readability of items within the target populations has been largely neglected. Think-aloud methods have been used to explore this criteria of content validity in other health-related outcome measures (Van Oort, 2011; McCorry et al., 2013). However, this methodology has not been used to test ACT measure validity. Nor has it been tested specifically with the population the measure was intended for, in this case for people with chronic pain. The present study utilised Conrad and Blair’s (1996) taxonomy of possible problems in understanding questionnaire items as a standardised tool to objectively classify errors in the think-aloud responses. As the content validity of ACT measures has not been evaluated in such a comprehensive, empirical manner before, this is a significant strength of this study, with far reaching implications for the design of process measures in general as well as those specific to ACT. Furthermore, as we rely heavily on self-report questionnaires to assess change these results also have important implications for the profession of clinical psychology.

The use of the error ‘problem classification matrix’ allowed us to systematically and objectively analyse qualitative responses from the cognitive interviews. This allowed us to explore the content validity of the ACT questionnaires. We explored the extent to which the respondents responded in the manner intended for items assessing the constructs of interest (psychological flexibility variables). In other words, we explored whether participants understood the items and therefore answered commensurate with the processes under study. The cognitive interviewing ‘think aloud’ method was used as it allowed us to assess the underlying cognitive processes forming an individual’s response. We gained insight into which elements of the question they did and did not consider/understand as they read and then provided an answer for the item. McCorry, Scullion, McMurray, Houghton and Dempster (2013) noted some limitations with using this method to elucidate validity as it is reliant on individual ability to verbalise thoughts and could potentially underestimate problems in answering the questions. The present study however, attempted to minimise the impact of this potential confounding factor through providing ‘thinking out-loud’ training to participants prior to the task, with checks of their understanding part-way through. Furthermore, should a reduced reading ability have impacted on participants’ ability to verbalise, these answers could still potentially be categorised as containing an error. For example, there were numerous occasions when errors were located in very brief verbalised responses as it was still apparent a logical/lexical error had occurred in their understanding. As described in the findings, intelligence was an influencing factor in the number of errors made overall and would suggest that reading ability could be linked to respondents’ understanding of the items. However, the overall TOPF score for the sample ($M = 103$) suggested that our data from the cognitive interviews responses was representative, and within general population norms for ‘average’ intellectual functioning.

# Clinical implications and recommendations for further research

This thesis has focused on testing the content validity of ACT questionnaires. Implications for ACT process research have also been noted throughout. The aim here is to summarise the various research and clinical implications.

## *Implications for ACT process research*

Assessing ACT constructs to determine meaningful change and therefore improved outcomes for individuals experiencing chronic pain relies on our understanding of the route by which these changes are made. In order to refine ACT theory through process research we must understand how ACT constructs contribute towards specific strengths and deficits in psychological inflexibility, hence understand the mechanisms of change. In order to do this, the process measures used to assess change need to be reliable and valid. Valid process measures which can accurately assess these mechanisms of change are crucial for continued ACT construct research. Ultimately this will tell us which processes are targeted by which individual ACT constructs and how these might interact.

As mentioned in the limitations of our study, we did not assess individuals who had no prior exposure to the ACT model. An interesting avenue for future research would be to use a larger sample size to determine whether socialisation to the model has an impact of how people understand the questions. Another consideration could be the possibility that psychological flexibility scores may actually decrease from initial baseline scores. This may occur should respondents not understand the measures, thus original scores are inaccurate to begin with. This has important clinical implications for assessing change during therapy.

Future research could also expand on other aspects of content validity which have been neglected throughout ACT process research. There is currently a lack of research involving *"target populations for item generation and evaluation*" (McAndrews, Richardson & Stopa, 2019, p. 273). According to the criteria proposed by Terwee et al. (2007) this is an important component of content validity if we are to comprehend how non 'ACT experts' interpret items. McAndrews et al. (2019) suggest that qualitative techniques are required to investigate the understanding of ACT process items. Although attempts have been made to use qualitative methodology to develop scales measuring the acceptance of chronic pain, this research is limited (Risdon, Eccleston, Crombez & McCracken, 2003). The present study has important implications for future research as it has developed on the original problem classification matrix proposed by Drennan (2003). We have provided a strategic method for assessing content validity that can supplement insightful qualitative data from cognitive interviewing 'think-aloud' techniques. This way of categorising errors is an important first step to standardising the analysis of cognitive interviewing data.

### *Implications for clinical practice*

*'They are too wordy and they are well above the average adult national literacy level. It's bizarre that for such an experientially based therapy, the questions are so wordy! We need simple tools which people can easily make sense of.'* (Example ACT clinician survey response, from Study Part One)

There are also consequences for clinical practice related to our ability to accurately measure changes in ACT process variables. Clinical decision making depends on knowing which components are needed for treatment and in what order they should be delivered (Levin et al., 2017). Without understanding how a treatment works through measuring changes in process variables, clinicians will not be able to effectively adapt treatment methods during therapy. This is especially relevant for the area of chronic pain, a condition which is known to be highly prevalent and have huge economic, social and emotional burden for individuals and wider society. If clinicians know which components to target and in what sequence this may have beneficial effects for treatment outcomes. The present study and its findings around process measure validity also have wider reaching implications for the profession of clinical psychology. In completing this research we have considered the importance of developing valid questionnaires which can tell us *"what components* of *treatment to use, in what sequence and ratios, that would most efficiently and effectively provide quality care'* (Levin et al., p416). As suggested here, knowing whether certain components of the intervention are more effective for certain conditions is of huge benefit to an already overstretched NHS in which therapy sessions are often time limited.

### *Developing future measures*

In this last section, suggestions for the future development of ACT process measures will be discussed. Some suggestions for improving upon the most problematic items, identified by the current study, will also be offered. We will then move onto final conclusions for the thesis.

Future research could endorse the combined methodologies from the present study for future questionnaire construction and assessment of content validity. Another potential avenue for research would be to further develop this comprehensive testing of validity. In addition to using the ‘think-aloud’ method to elucidate validity, other empirical methods should be explored to establish other components of construct validity. Similarly, to the predominant focus of research suggesting potential issues with ACT questionnaires’ content validity, tests of construct validity as a whole concept have also relied on correlational, factor analytic methods of assessment. Whilst the present study investigated understanding of individual questionnaire items, further research on *construct representation* may help to establish whether responses are reflective of the intended constructs measured by the questionnaire (Strauss & Smith, 2009). As mentioned previously, our data collected through the card-sorting task could be utilised to elucidate respondents’ understanding of the questionnaire *constructs*, i.e. whether these map onto those intended by the questionnaire designer. Although this analysis was beyond the scope of this thesis and not in line with the research question and aims of the present study, a potential methodology for utilising this data in the future is described in the appendices (refer to appendix C).

During the main analysis (see chapter 5) we also looked for any patterns in error occurrence for each of the most problematic items. Potential causes of these errors were discussed in terms of reoccurring patterns between individual participant answers (See Table 6.1).

**Table 6.1: Suggestions for change of the most problematic ACT measure items (table adapted from Van Oort, 2011)**

| Original Item | Suggestion for improvement | Reason for suggestion |
|---|---|---|
| **CPAQ (4)**<br>Before I can make any serious plans, I have to get some control over my pain. | I need to have control over my pain before I plan to take part in activities. | The word *‘serious’* is subject to individual definition and distracts away from the main aim of the item- to assess the pursuit of activities despite the pain. The suggestion uses an alternative word. Although the item still contains two part (which had been identified as problematic), they have been swapped round to avoid respondents focusing only on the ‘making plan’ aspect of the question. The suggestion aims to keep the focus of responses on pain willingness. |
| **CPAQ (8)**<br>My worries and fears about what pain will do to me are true. | The worries I have about the impact of my pain are accurate. | Some participants did not consider whether their thoughts were *‘true’* within their answers- more around how often they worry. The suggestion uses slightly different wording to elicit thoughts about how the pain might prevent them engaging in activities they enjoy. |
| **PIPS (10)**<br>I put a lot of effort into fighting the pain. | A lot of my effort goes into trying to get rid of my pain. | The word ‘*fighting’* was recast by participants to mean something positive and was interpreted as equivalent to them ‘coping’ with the pain, or ’working with the pain’. The suggestion uses an alternative word. |
| **PIPS (7)**<br>I say things like ‘‘I don’t have any energy’’, ‘‘I am not well enough’’, ‘‘I don’t have time’’, ‘‘I don’t dare’’, ‘‘I have too much pain’’, ‘‘I feel too bad’’ or ‘‘I don’t feel like it’’. | I think things like ‘‘I don’t have any energy’’, ‘‘I am not well enough’’ ‘‘I have too much pain’’ or ‘‘I feel too bad to do this activity”. | The structure often caused confusion, with too many components. The suggestion was to reduce the number of parts to the item.<br>Additionally, the word ‘*say’* causes confusion- the intention of the question is about their ‘thinking/acting/doing/behaving’ as opposed to actually speaking the thoughts out-loud to other people, which is often how it is interpreted. A suggestion was made to change this to ‘I think’. |

| | | |
|---|---|---|
| ***AAQ-II (3)***<br>I worry about not being able to control my worries and feelings. | Not being able to control my worries is something that bothers me. | Logical errors were made because respondents referred to the extent of their worry and did not include the idea of *'control'*. Many of these logical errors lead to conceptually inconsistent errors also- their answer revealed they are 'good' at controlling worries (so actually high inflexibility- opposite of score given)- the changes hope to correct this. |
| ***AAQ-II (5)***<br>Emotions cause problems in my life. | I cannot tolerate strong feelings | The word *'emotions'* caused difficulties- respondents were not sure what is meant by this term. Respondents thought about 'emotions' and experiencing them in relation to what other people think/see of them. Others don't 'see' them express emotion or understand it. Therefore, the respondents did not give an answer reflecting level of acceptance of difficult emotions. In this new version of the item the word 'emotions' has been replaced with 'strong feelings'. |

The findings of this study provided insight into where patterns of error making occurred for particular items. It has been suggested that future ACT construct research relies on the use of process measures which are consistent with an evolving ACT theory (McAndrews, Richardson, & Stopa, 2019). As ACT develops, its' definitions and language change and this must also be reflected in the very measures used to test the theory. However, as has been investigated in the present study we must also assess the understandability of the questionnaire items in order to know whether they truly measure what they intend to. In the final table above (Table 6.1) some suggestions are offered as to how we might develop the most problematic items. The identified patterns in error making were used to inform these suggestions and are hopefully the beginnings of refining these questionnaires.

## Conclusions

The aim of this thesis was to investigate the content validity of commonly used ACT questionnaires in individuals experiencing chronic pain. The refinement of ACT theory through process research relies on our understanding of the mechanisms of change, i.e. how ACT constructs contribute towards specific strengths and deficits in psychological inflexibility. Questionnaires assessing meaningful change need to be robust and work across multiple 'terrains' however ACT process research has been hindered because of the way in which items have been structured. We currently rely on these self-report measures however there is too much ‘noise’ in the system due to respondents not being able to understand the items, leading to actual change being over or underestimated. Valid process measures which can accurately assess the mechanisms of change are crucial for continued ACT construct research. If individuals do not understand the questions on process measures and hence do not respond in the manner intended then what can be inferred from the results has limited utility; we will not understand which processes are targeted by which individual ACT constructs.

The present study was novel and recognised that this assessment of content validity has not been done before. Although previous research has assessed aspects of construct validity this has relied on the use of quantitative, correlative methodology (e.g. factor analysis, multiple regression). The present study comprehensively explored the content validity of ACT measures to assess *how* individuals understand the questionnaires and what impact this had on their response. The aims were achieved by the completion of two studies: 1) Part One- a research review and clinician survey to establish frequency of use for ACT process questionnaires; 2) Part Two- a questionnaire validation study. In the main study, cognitive interviewing methods were used to empirically evaluate the understandability of individual questionnaire items. A taxonomy of problem classification was used to objectively classify errors made in completing items, through analysing respondent’s ability to avoid making errors in arriving at their response. This enabled us to see where patterns of error making occurred amongst the items and which items were most problematic in accurately reflecting the intended process variables.

As well as the implications for research, these findings are also important for clinical practice. These findings could inform how clinical psychologists working within chronic pain services are still able to use the existing measures, prior to new measures being developed. The findings have implications for the administration of existing measures, and suggest that clinicians should be checking for understanding and prompting respondents as they complete measures. As the study has highlighted the most problematic items for each of the three measures, this provides clinicians with an awareness of which items in particular to provide support to respondents in completing. Thus, an important contribution of these findings is enabling clinicians to be better equipped to effectively capture change in clients undergoing ACT. Understanding how a treatment works through measuring changes in process variables will enable clinicians to tailor treatment targeting specific key sub-processes, with improved outcomes for their clients.

This research was important for the area of chronic pain and for the wider profession of clinical psychology. In addition to furthering research and improving clinical practice we must also understand exactly which ACT components are being measured in order to distinguish it from other therapies and thus provide evidence to support its efficacy (Levin et al., 2017). Although the ‘think-aloud’ data identified problems in understanding which may be unique to ACT concepts, there were also issues which are common to all questionnaires, i.e. those with complex or imprecise wording. Therefore, this study also has more universal implications for the design of process measures in general as well as those specific to ACT for chronic pain. In establishing where there may be patterns in misunderstanding across the questionnaire items it is hoped that this study will help towards the future development and refinement of existing ACT questionnaires.

# REFERENCES

Alonso, M. A., López, A., Losada, A., & González, J. L. (2013). Acceptance and commitment therapy and selective optimization with compensation for older people with chronic pain: A pilot study. *Behavioural Psychology, 21*(1), 59.

Alonso-Fernández, M., López-López, A., Losada, A., González, J. L., & Wetherell, J. L. (2016). Acceptance and commitment therapy and selective optimization with compensation for institutionalized older people with chronic pain. *Pain Medicine*, *17*(2), 264-277.

Amtmann, D., Askew, R. L., Kim, J., Chung, H., Ehde, D. M., Bombardier, C. H., Charles, H., Kraft, G.H., Jones, S.M., & Johnson, K. L. (2015). Pain affects depression through anxiety, fatigue, and sleep in multiple sclerosis. *Rehabilitation Psychology, 60*(1), 81-90.

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports*, *19*, 3-11.

Beatty, P.C., & Willis, G.B. (2007). Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quarterly*, *71*, 287-311.

Bedson, J., & Croft, P. R. (2008). The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskeletal Disorders*, *9*(1), 116.

Beerlage-de Jong, N., Kulyk, O. A., Wentzel, M. J., Oinas-Kukkonen, H., & van Gemert-Pijnen, J. E. W. C. (2015). 'Sorting out' the PPQ: A mixed-methods approach to evaluate Perceived Persuasiveness Questionnaire constructs. In 10th INTERNATIONAL PERSUASIVE TECHNOLOGY 2015 Conference, Chicago, United States. Retrieved from https://research.utwente.nl/en/publications/sorting-out-the-ppq-a-mixed-methods-approach-to-evaluate-perceive.

Bergström, G., Jensen, I. B., Bodin, L., Linton, S. J., Nygren, Å. L., & Carlsson, S. G. (1998). Reliability and factor structure of the Multidimensional Pain Inventory–Swedish Language version (MPI-S). *Pain*, *75*(1), 101-110.

Biglan, A., & Hayes, S. C. (1996). Should the behavioral sciences become more pragmatic? The case for functional contextualism in research on human behavior. *Applied and Preventive Psychology*, *5*(1), 47-57.

Black, M. B. (1963). On formal ethnographic procedures. *American Anthropologist*, *65*, 1347-1351.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T., & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire - II: A revised measure of psychological flexibility and experiential avoidance. *Behavior Therapy*, *42*, 676-688.

Breivik, H., Collett, B., Ventafrida, V., Cohen, R., & Gallagher, D. (2006). Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *European Journal of Pain*, *10*, 287–333.

British Pain Society (BPS). (c2014). People living with pain: useful definitions and glossary. Retrieved 3 July 2019, from https://www.britishpainsociety.org/people-with-pain/useful-definitions-and-glossary/#pain

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*, 822–848.

Buhrman, M., Skoglund, A., Husell, J., Bergström, K., Gordh, T., Hursti, T., … Andersson, G. (2013). Guided internet-delivered acceptance and commitment therapy for chronic pain patients: a randomized controlled trial. *Behaviour Research and Therapy*, *51*(6), 307-315.

Burke, A. L., Mathias, J. L., & Denson, L. A. (2015). Psychological functioning of people living with chronic pain: A meta-analytic review. *British Journal of Clinical Psychology*, *54*(3), 345-360.

Casey, M. B., Smart, K., Segurado, R., Hearty, C., Gopal, H., Lowry, D., Flanagan, D., McCracken, L. & Doody, C. (2018). Exercise combined with Acceptance and Commitment Therapy (ExACT) compared to a supervised exercise programme for adults with chronic pain: study protocol for a randomised controlled trial. *Trials*, 19, 194.

Cederberg, J. T., Cernvall, M., Dahl, J., von Essen, L., & Ljungman, G. (2016). Acceptance as a mediator for change in acceptance and commitment therapy for persons with chronic pain? *International Journal of Behavioral Medicine*, *23*(1), 21-29.

Cleeland, C. S., & Ryan, K. M. (1994). Pain assessment: global use of the Brief Pain Inventory. *Annals, Academy of Medicine, Singapore*, *23*(2), 129-138.

Cochrane, A., Barnes-Holmes, D., & Barnes-Holmes, Y. (2008). The perceived-threat behavioral approach test (PT-BAT): Measuring avoidance in high-, mid-, and low-spider-fearful participants. *The Psychological Record*, *58*(4), 585-596.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20,* 37–46.

Cohen, R., & Swerdlik, M. (1999). *Psychological testing and assessment : an introduction to tests and measurement* (Fourth edition.). (pp 146- 214). Mountain View, Calif: Mayfield Publishing Co.

Conrad, F., & Blair, J. (1996). From impressions to data: Increasing the objectivity of cognitive interviews. In Conrad, F., & Blair, J. *Proceedings of the Section on Survey Research Methods* (pp. 1-9). Alexandria, VA: American Statistical Association.

Conrad, F., Blair, J., & Tracy, E. (2000). Verbal reports are data! A theoretical approach to cognitive interviews. In Conrad, F., Blair, J., & Tracy, E. *FEDERAL COMMITTEE ON STATISTICAL METHODOLOGY RESEARCH 1999 Conference* (pp. 11-20). Washington, DC: Office of Management and Budget. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.3306&rep=rep1&type=pdf.

Coxon, A.M. (1999). Collecting Free-Sorting Data. In Coxon, A.M., *Sorting data: Collection and analysis*: *Sage University Paper Series on Quantitative Applications in the Social Sciences* (pp 5-27). Thousand Oaks, CA: Sage.

Crombez, G., Vlaeyen, J. W., Heuts, P. H., & Lysens, R. (1999). Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*, *80*, 329-339.

Dahl, J., Wilson, K. G., & Nilsson, A. (2004). Acceptance and commitment therapy and the treatment of persons at risk for long-term disability resulting from stress and pain symptoms: A preliminary randomized trial. *Behavior Therapy*, *35*, 785-801.

Dersh, J., Polatin, P. B., & Gatchel, R. J. (2002). Chronic pain and psychopathology: research findings and theoretical considerations. *Psychosomatic Medicine*, *64*(5), 773-786.

Dindo, L., Zimmerman, M. B., Hadlandsmyth, K., StMarie, B., Embree, J., Marchman, J., Tripp-Reimer, T. & Rakel, B. (2018). Acceptance and Commitment Therapy for Prevention of Chronic Postsurgical Pain and Opioid Use in At-Risk Veterans: A Pilot Randomized Controlled Study. *Journal of Pain, 19*, 1211-1221.

Drennan, J. (2003). Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing*, *42*, 57-63.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture and Activity*, *5*(3), 178-186.

Fairbank, J. C., Couper, J., Davies, J. B., & O'brien, J. P. (1980). The Oswestry low back pain disability questionnaire. *Physiotherapy*, *66*(8), 271-273.

Fayaz, A., Croft, P., Langford, R. M., Donaldson, L. J., & Jones, G. T. (2016). Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. *British Medical Journal open*, *6*, 1-12.

Feldman, G., Hayes, A., Kumar, S., Greeson, J., & Laurenceau, J. P. (2007). Mindfulness and emotion regulation: The development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). *Journal of psychopathology and Behavioral Assessment*, *29*(3), 177-190

Fingleton, C., Smart, K., Moloney, N., Fullen, B. M., & Doody, C. (2015). Pain sensitization in people with knee osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis and Cartilage*, *23*(7), 1043-1056.

Fish, R. A., McGuire, B., Hogan, M., Morrison, T. G., & Stewart, I. (2010). Validation of the Chronic Pain Acceptance Questionnaire (CPAQ) in an Internet sample and development and preliminary validation of the CPAQ-8. *Pain*, *149*(3), 435-443.

Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*. *76*(5), 378.

Francis, A. W., Dawson, D. L., & Golijani-Moghaddam, N. (2016). The development and validation of the Comprehensive assessment of Acceptance and Commitment Therapy processes (CompACT). *Journal of Contextual Behavioral Science, 5*, 134-145.

Freelon, D. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science*, 5(1), 20-33.

Fresco, D. M., Moore, M. T., van Dulmen, M. H., Segal, Z. V., Ma, S. H., Teasdale, J. D., & Williams, J. M. G. (2007). Initial psychometric properties of the experiences questionnaire: validation of a self-report measure of decentering. *Behavior Therapy*, *38*(3), 234-246.

Gámez, W., Chmielewski, M., Kotov, R., Ruggero, C., & Watson, D. (2011). Development of a measure of experiential avoidance: The multidimensional experiential avoidance Questionnaire. Psychological Assessment, *23*, 692–713.

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., Kerr, S., Tansey, L., Noel, P., Ferenbach, C., & Masley, S. (2014). The development and initial validation of the cognitive fusion questionnaire. *Behavior Therapy, 45*(1), 83-101.

Gilhooly, K., & Green, C. (1996). Protocol analysis: theoretical background. *Handbook of qualitative research methods for psychology and the social sciences*, *14*, 43-54.

Gloster, A. T., Klotsche, J., Chaker, S., Hummel, K. V., & Hoyer, J. (2011). Assessing psychological flexibility: What does it add above and beyond existing constructs? *Psychological Assessment*, *23*, 970-982.

Hann, K. E., & McCracken, L. M. (2014). A systematic review of randomized controlled trials of Acceptance and Commitment Therapy for adults with chronic pain: Outcome domains, design quality, and efficacy. *Journal of Contextual Behavioral Science*, *3,* 217-227.

Harloff, J., & Coxon, A. P. M. (2005). *How to Sort. A Short Guide on Sorting Investigations.* Retrieved 28 February, 2016 from https://www.researchgate.net/profile/SQ_Zheng/publication/220294595_How_to_Sort/links/544e67b80cf2bca5ce90b25c.pdf

Harris, R. (2009). ACT in a Nutshell. In Harris, R., *ACT Made Simple: An Easy-To-Read Primer on Acceptance and Commitment Therapy.* (pp 6-18). Oakland: New Harbinger Publications.

Hayes, S. C. (1995). Why cognitions are not causes. *The Behavior Therapist*, *18*, 59-64.

Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy, 35*, 639–665.

Hayes, S.C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Springer Science & Business Media.

Hayes, S. C., Barnes-Holmes, D., & Wilson, K. G. (2012). Contextual behavioral science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science, 1*(1-2), 1-16.

Hayes, S.C., Hayes, L.J., Reese, H.W & Sarbin, T.R. (1993). *Varieties of scientific contextualism.* Reno, NV: Context Press.

Hayes, S., Hogan, M., Dowd, H., Doherty, E., O'Higgens, S., Gabhain, S.N., Macneela, P., Murphy, A. W., Kropmans, T., O'Neill, C., Newell, J. & McGuire, B. E. (2014). Comparing the clinical-effectiveness and cost-effectiveness of an internet delivered Acceptance and Commitment Therapy (ACT) intervention with a waiting list control among adults with chronic pain: Study protocol for a randomised controlled trial. *BMJ Open, 4* (7) (no pagination).

Hayes, A., Hope, D. A., & Hayes, S. (2007). Towards an understanding of the process and mechanisms of change in cognitive and behavioral therapy: Linking innovative methodology with fundamental questions. *Clinical Psychology Review*, *27*, 679-681.

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, *44*, 1-25.

Hayes, S.C., Strosahl, K.D., Wilson, K.G., Bisset, R.T., Pistorello, J., Toarmino, D., Polusny, M.A., Dykstra, T, A., Batten, S.V., Bergan, J., Stewart, S.H., Zvolensky, M.J., Eifert, G.H., Bond, F.W., Forsyth, J.P., Karekla, M., & McCurry, S.M. (2004). Measuring experiential avoidance: A preliminary test of a working model. *The Psychological Record*, *54*, 553-578.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An experiential approach to behavior change.* New York: The Guilford Press.

Herbert, M. S., Afari, N., Liu, L., Heppner, P., Rutledge, T., Williams, K., Eraly, S., Vanbuskirk, K., Nguyen, C., Bondi, M., & Atkinson, J. H. (2017). Telehealth versus in-person acceptance and commitment therapy for chronic pain: a randomized noninferiority trial. *The Journal of Pain*, *18*(2), 200-211.

Holdnack, J. A., & Whipple Drozdick, L. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Clinical and interpretive manual*. San Antonio, TX: Pearson.

Hong, J., Reed, C., Novick, D., & Happich, M. (2013). Costs associated with treatment of chronic low back pain: an analysis of the UK General Practice Research Database. *Spine*, *38*(1), 75-82.

Hughes, L.S., Clark, J., Colclough, J.A., Dale, E., & McMillan, D. (2017). Acceptance and Commitment Therapy (ACT) for Chronic Pain: A Systematic Review and Meta-Analyses. *Clinical Journal of Pain, 33,* 552-568.

International Association for the Study of Pain (c2011). IASP Terminology. Retrieved 3 July 2019, from https://www.iasp-pain.org/terminology

International Association for the Study of Pain. (1986). Classification of chronic pain. *Pain*, *3*, 135–138.

Johnston, M., Foster, M., Shennan, J., Starkey, N. J., & Johnson, A. (2010). The effectiveness of an acceptance and commitment therapy self-help intervention for chronic pain. *The Clinical Journal of Pain*, *26*(5), 393-402.

Jones, G. T., Atzeni, F., Beasley, M., Flüß, E., Sarzi-Puttini, P., & Macfarlane, G. J. (2015). The prevalence of fibromyalgia in the general population: a comparison of the American College of Rheumatology 1990, 2010 and modified 2010 classification criteria. *Arthritis & Rheumatology*, *67*(2), 568-575.

Katz, J., & Rosenbloom, B. N. (2015). The golden anniversary of Melzack and Wall’s gate control theory of pain: Celebrating 50 years of pain research and management. *Pain Research and Management*, *20*(6), 285-286.

Katz, J., Rosenbloom, B. N., & Fashler, S. (2015). Chronic Pain, Psychopathology, and DSM-5 Somatic Symptom Disorder. *Canadian Journal of Psychiatry, 60*(4), 160–167.

Katz, J., & Seltzer, Z. E. (2009). Transition from acute to chronic postsurgical pain: risk factors and protective factors. *Expert Review of Neurotherapeutics, 9*(5), 723-744.

Kemani, M. K., Hesser, H., Olsson, G. L., Lekander, M., & Wicksell, R. K. (2016). Processes of change in Acceptance and Commitment Therapy and Applied Relaxation for long-standing pain. *European Journal of Pain, 20*(4), 521-531.

Kemani, M.K., Olsson, G.L., Lekander, M. Hesser, H., Anderson, E. & Wicksell, R. K.  (2015). Efficacy and cost-effectiveness of acceptance and commitment therapy and applied relaxation for long standing pain: A randomised controlled trial. *Clinical Journal of Pain*, *31*, 1004-1016.

Krahé, C., Springer, A., Weinman, J. A., & Fotopoulou, A. K. (2013). The social modulation of pain: others as predictive signals of salience–a systematic review. *Frontiers in Human Neuroscience*, *7*, 386.

Kroenke, K., Spitzer, R., & Williams, J. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16* (9), 606-613.

Landis, J. R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*. 33(1), 159–74.

Lavand'homme, P. (2017). Transition from acute to chronic pain after surgery. *Pain*, *158*, 50-54.

LeResche, L., Mancl, L. A., Drangsholt, M. T., Saunders, K., & Von Korff, M. (2005). Relationship of pain and symptoms to pubertal development in adolescents. *Pain*, *118*(1-2), 201-209.

Lerman, S. F., Rudich, Z., Brill, S., Shalev, H., & Shahar, G. (2015). Longitudinal associations between depression, anxiety, pain, and pain-related disability in chronic pain patients. *Psychosomatic Medicine*, *77*(3), 333-341.

Levin, M. E., Herbert, J.D., & Forman, (2017). Acceptance and commitment therapy: a critical review to guide clinical decision making. In McKay, D., Abramowitz, J.S., & Storch, E. A., (1st Ed), *Treatments for psychological problems and syndromes*, (pp 413-432). Chichester: John Wiley & Sons Ltd.

Lin, J., Klatt, L. I., McCracken, L. M. & Baumeister, H. (2018). Psychological flexibility mediates the effect of an online-based acceptance and commitment therapy for chronic pain: An investigation of change processes. *Pain, 159*, 663-672.

Lin, J., Lüking, M., Ebert, D. D., Buhrman, M., Andersson, G., & Baumeister, H. (2015). Effectiveness and cost-effectiveness of a guided and unguided internet-based Acceptance and Commitment Therapy for chronic pain: Study protocol for a three-armed randomised controlled trial. *Internet Interventions*, *2*(1), 7-16.

Lin, J., Paganini, S., Sander, L., Lüking, M., Ebert, D. D., Buhrman, M., Anderson, G., & Baumeister, H. (2017). An internet-based intervention for chronic pain: a three-arm randomized controlled study of the effectiveness of guided and unguided acceptance and commitment therapy. *Deutsches Ärzteblatt International*, *114*(41), 681.

Luciano, J. V., Guallar, J. A., Aguado, J., López-del-Hoyo, Y., Olivan, B., Magallón, R., ... & Garcia-Campayo, J. (2014). Effectiveness of group acceptance and commitment therapy for fibromyalgia: a 6-month randomized controlled trial (EFFIGACT study). *Pain*, *155*(4), 693-702.

Magni G, Caldieron C, Rigatti-Luchini S, Merskey H. (1990). Chronic musculoskeletal pain and depressive symptoms in the general population: an analysis of the first national and nutrition examination survey data. *Pain*, *43*, 299-307.

Maniadakis, N., & Gray, A. (2000). The economic burden of back pain in the UK. *Pain*, *84*(1), 95-103.

McAndrews, Z., Richardson, J., & Stopa, L. (2019). Psychometric properties of acceptance measures: A systematic review. *Journal of Contextual Behavioral Science*, *12,* 261-277.

McCorry, N. K., Scullion, L., McMurray, C. M., Houghton, R., & Dempster, M. (2013). Content validity of the illness perceptions questionnaire–revised among people with type 2 diabetes: A think-aloud study. *Psychology & Health, 28*, 675-685.

McCracken, L. M. (2013). Committed action: an application of the psychological flexibility model to activity patterns in chronic pain. *The Journal of Pain*, *14*(8), 828-835.

McCracken, L.M., & Eccleston, C. (2005). A prospective study of acceptance of pain and patient functioning with chronic pain. *Pain*, *118*, 164-169.

McCracken, L. M., Gauntlett-Gilbert, J., & Vowles, K. E. (2007). The role of mindfulness in a contextual cognitive-behavioral analysis of chronic pain-related suffering and disability. *Pain*, *131*, 63-69.

McCracken, L. M., Gutiérrez-Martínez, O., & Smyth, C. (2013). "Decentering" reflects psychological flexibility in people with chronic pain and correlates with their quality of functioning. *Health Psychology*, *32*, 820.

McCracken, L. M., & Morley, S. (2014). The psychological flexibility model: a basis for integration and progress in psychological approaches to chronic pain management. *The Journal of Pain*, *15*(3), 221-234.

McCracken, L. M., Sato, A. & Taylor, G. J. (2013a). A trial of a brief group-based form of acceptance and commitment therapy (ACT) for chronic pain in general practice: pilot outcome and process results. *Journal of Pain, 14*, 1398-406.

McCracken, L. M., Sato, A., Wainwright, D., House, W. & Taylor, G. J. (2014a). A feasibility study of brief group-based acceptance and commitment therapy for chronic pain in general practice: recruitment, attendance, and patient views. *Primary Health Care Research & Development*, *15*, 312-23.

McCracken, L. M., & Turk, D. C. (2002). Behavioral and cognitive–behavioral treatment for chronic pain: outcome, predictors of outcome, and treatment process. *Spine*, *27*(22), 2564-2573.

McCracken, L. M., & Vowles, K. E. (2007). Psychological flexibility and traditional pain management strategies in relation to patient functioning with chronic pain: an examination of a revised instrument. *The Journal of Pain, 8*(9), 700-707.

McCracken, L. M., & Vowles, K. E. (2014). Acceptance and commitment therapy and mindfulness for chronic pain: model, process, and progress. *American Psychologist*, *69*, 178.

McCracken, L.M., Vowles, K.E., & Eccleston, C. (2004). Acceptance of chronic pain: component analysis and a revised assessment method. *Pain, 107,* 159-166.

McCracken, L. M., Vowles, K. E., & Eccleston, C. (2005). Acceptance-based treatment for persons with complex, long standing chronic pain: a preliminary analysis of treatment outcome in comparison to a waiting phase. *Behaviour Research and Therapy*, *43*(10), 1335-1346.

McCracken, L.M., & Yang, S. (2006). The role of values in a contextual cognitive-behavioural approach to chronic pain. *Pain, 123,* 137-145.

McWilliams, L. A., Goodwin, R. D., & Cox, B. J. (2004). Depression and anxiety associated with three pain conditions: results from a nationally representative sample. *Pain, 111*(1-2), 77-83.

Mehling, W. E., Price, C., Daubenmier, J. J., Acree, M., Bartmess, E., & Stewart, A. (2012). The multidimensional assessment of interoceptive awareness (MAIA). *PloS one*, *7*(11), Retrieved from https://doi.org/10.1371/journal.pone.0048230

Melzack, R., & Wall, P. D. (1965). Pain mechanisms: a new theory. *Science*, *150*(3699), 971-979.

Merskey, H., & Bogduk, N. (1994). Part III: Pain Terms, A Current List with Definitions and Notes on Usage. In Merskey, H & Bogduk, N. (2nd Ed.) *Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms* (pp 209-214). Seattle: IASP Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, *50,* 741-749.

Moayedi, M., & Davis, K. D. (2012). Theories of pain: from specificity to gate control. *Journal of Neurophysiology*, *109*(1), 5-12.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & the PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine, 151,* 264-269.

Morley, S. (2004). Process and change in cognitive behaviour therapy for chronic pain. *Pain, 109,* 205-206.

Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain*, *80*(1-2), 1-13.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 223-250.

Nes, A. A., van Dulmen, S., Wicksell, R., Fors, E. A., & Eide, H. (2017). Analyzing change processes resulting from a smartphone maintenance intervention based on acceptance and commitment therapy for women with chronic widespread pain. *International Journal of Behavioral Medicine*, *24*(2), 215-229.

National Institute for Health and Care Excellence (NICE). (c2018). Guideline scope, Chronic pain: assessment and management. Retrieved 18 July 2019 from https://www.nice.org.uk/guidance

Nicholas, M.K. (1989). Self-efficacy and chronic pain. In BRITISH PSYCHOLOGICAL SOCIETY 1989 Conference, St Andrews: Scotland.

Nicholas, M. K. (2007). The pain self-efficacy questionnaire: Taking pain into account. *European Journal of Pain*, *11*(2), 153-163.

Notebaert, L., Crombez, G., Vogt J., De Houwer J., Van Damme S., & Theeuwes J. (2011). Attempts to control pain prioritize attention towards signals of pain: An experimental study. *Pain. 152*, 1068–73.

Online Surveys. (c2019). Online Surveys (formerly BOS). Retrieved 19 April 2019, from https://www.onlinesurveys.ac.uk.

Öst, L. G. (2014). The efficacy of acceptance and commitment therapy: an updated systematic review and meta-analysis. *Behaviour Research and Therapy*, *61*, 105-121.

Parkitny, L., & McAuley, J. (2010). The depression anxiety stress scale (DASS). *Journal of Physiotherapy*, *56*(3), 204.

Plumb Vilardaga, J. C. (2012). *Acceptance and commitment therapy for longstanding chronic pain in a community-based outpatient group setting* (Unpublished doctoral dissertation). University of Nevada, Reno.

Probst, T., Baumeister, H., McCracken, L., & Lin, J. (2019). Baseline psychological inflexibility moderates the outcome pain interference in a randomized controlled trial on internet-based acceptance and commitment therapy for chronic pain. *Journal of Clinical Medicine*, *8*(1), 24.

Rayner, L., Hotopf, M., Petkova, H., Matcham, F., Simpson, A., & McCracken, L. M. (2016). Depression in patients with chronic pain attending a specialised pain treatment centre: prevalence and impact on health care costs. *Pain*, *157*(7), 1472.

Reid, K. J., Harker, J., Bala, M. M., Truyers, C., Kellen, E., Bekkering, G. E., & Kleijnen, J. (2011). Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. *Current Medical Research and Opinion*, *27*, 449-462.

Reneman, M. F., Dijkstra, A., Geertzen, J. H., & Dijkstra, P. U. (2010). Psychometric properties of chronic pain acceptance questionnaires: a systematic review. *European Journal of Pain, 14,* 457-465.

Risdon, A., Eccleston, C., Crombez, G., & McCracken, L. (2003). How can we learn to live with pain? A Q-methodological analysis of the diverse understandings of acceptance of chronic pain. *Social science & medicine*, *56*(2), 375-386.

Riley, J. F., Ahern, D. K., & Follick, M. J. (1988). Chronic pain and functional impairment: assessing beliefs about their relationship. *Archives of Physical Medicine and Rehabilitation*, *69*, 579-582.

Rochefort, C., Baldwin, A. S., & Chmielewski, M. (2018). Experiential avoidance: An examination of the construct validity of the AAQ-II and MEAQ. *Behavior therapy*, *49*(3), 435-449.

Schulte, D., & Eifert, G.H. (2002). What to do when manuals fail? The dual model of psychotherapy. *Clinical Psychology: Science and Practice*, *9,* 312-328.

Schuman, H. (1966). The random probe: A technique for evaluating the validity of closed questions. *American Sociological Review, 31,* 218-222.

Scott, W., Hann, K. E., & McCracken, L. M. (2016). A comprehensive examination of changes in psychological flexibility following acceptance and commitment therapy for chronic pain. *Journal of Contemporary Psychotherapy*, *46*(3), 139-148.

Simister, H. D., Tkachuk, G. A., Shay, B. L., Vincent, N., Pear, J. J., & Skrabek, R. Q. (2018). Randomized controlled trial of online acceptance and commitment therapy for fibromyalgia. *The Journal of Pain*, *19*(7), 741-753.

Simpson, P. A., Mars, T. & Esteves, J. E. (2017). A systematic review of randomised controlled trials using Acceptance and commitment therapy as an intervention in the management of non-malignant, chronic pain in adults. *International Journal of Osteopathic Medicine, 24*, 18-31.

Smout, M., Davies, M., Burns, N., & Christie, A. (2014). Development of the valuing questionnaire (VQ). *Journal of Contextual Behavioral Science, 3*(3), 164-172.

Spissu, A., Cannas, A., Ferrigno, P., Pelaghi, A. E., & Spissu, M. (1999). Anatomic correlates of painful tonic spasms in multiple sclerosis. *Movement disorders: official journal of the Movement Disorder Society*, *14*(2), 331-335.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*(10), 1092-1097.

Steiner, J. L., Bogusch, L., & Bigatti, S. M. (2013). Values-based action in fibromyalgia: results from a randomized pilot of acceptance and commitment therapy. *Health Psychology Research*, *1*(3), 176-181.

Straub, D., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information systems*, *13*(1), 24, 380-427.

Strauss, M.E., & Smith, G.T. (2009). Construct validity: advances in theory and methodology. *Annual Review of Clinical Psychology, 5,* 1-25.

Sullivan, M. J., Bishop, S. R., & Pivik, J. (1995). The pain catastrophizing scale: development and validation. *Psychological Assessment*, *7*(4), 524-532.

Sullivan, S., Ehde, D., Turner, J. & Dillworth, T. (2012). Acceptance, catastrophizing, and depressive symptoms in persons with disability-related chronic pain. *Journal of Pain,* (1), S96.

Tabor, A., O’Daly, O., Gregory, R. W., Jacobs, C., Travers, W., Thacker, M. A., & Moseley, G. L. (2016). Perceptual inference in chronic pain: an investigation into the economy of action hypothesis. *The Clinical Journal of Pain*, *32*(7), 588-593.

Terwee, C.B., Bot, S.D., de Boer, M.D., van der Windt, D.A., Knol, D.L., Dekker, J., de Vet, H.C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60* (1), 34-42.

Thorsell, J., Finnes, A., Dahl, J., Lundgren, T., Gybrant, M., Gordh, T., & Buhrman, M. (2011). A comparative study of 2 manual-based self-help interventions, acceptance and commitment therapy and applied relaxation, for persons with chronic pain. *The Clinical Journal of Pain*, *27*(8), 716-723.

Tourangeau, R. (1984). Cognitive sciences and survey methods. *Cognitive aspects of survey methodology: Building a bridge between disciplines*, *15*, 73-100.

Trompetter, H. R., Bohlmeijer, E. T., Fox, J.-P. & Schreurs, K. M. (2015a). Psychological flexibility and catastrophizing as associated change mechanisms during online Acceptance & Commitment Therapy for chronic pain. *Behaviour Research and Therapy, 74*, 50-59.

Trompetter, H. R., Bohlmeijer, E. T., Veehof, M. M., & Schreurs, K. M. (2015b). Internet-based guided self-help intervention for chronic pain based on Acceptance and Commitment Therapy: a randomized controlled trial. *Journal of Behavioral Medicine*, *38*, 66-80.

Turk, D. C., & Okifuji, A. (2002). Psychological factors in chronic pain: evolution and revolution. *Journal of Consulting and Clinical Psychology*, *70,* 678-690.

Tyndall, I., Waldeck, D., Pancani, L., Whelan, R., Roche, B., & Dawson, D. L. (2019). The Acceptance and Action Questionnaire-II (AAQ-II) as a measure of experiential avoidance: Concerns over discriminant validity. *Journal of Contextual Behavioral Science*, *12*, 278-284.

Van Damme, S., Crombez, G., Eccleston, C., & Koster, E.H.W. (2006). Hypervigilance to learned pain signals: a componential analysis. *The Journal of Pain, 7,* 346-357.

Van der Veer, K., Hak, T., & Jansen, H. (2002). The Three-Step Test Interview (TSTI): An Observational Instrument for Pre-testing Self-Completion Questionnaires. In

*QUESTIONNAIRE DEVELOPMENT, EVALUATION AND TESTING 2004 Conference.* Charleston: South Carolina. Retrieved from https://ssrn.com/abstract=636782

Van Oort, L., Schröder, C., & French, D.P. (2011). What do people think about when they answer the Brief Illness Perception Questionnaire? A ‘think-aloud’ study. *British Journal of Health Psychology, 16,* 231-245.

Van Ryckeghem, D. M., Noel, M., Sharpe, L., Pincus, T., & Van Damme, S. (2019). Cognitive biases in pain: an integrated functional–contextual framework. *Pain, 160*(7), 1489-1493.

Van Someren, M.W., Barnard, Y.F., & Sandberg, J.A.C. (1994). *The Think Aloud Method: A practical guide to modelling cognitive processes.* London: Academic Press.

Veehof, M. M., Oskam, M. J., Schreurs, K. M., & Bohlmeijer, E. T. (2011). Acceptance-based interventions for the treatment of chronic pain: a systematic review and meta-analysis. *Pain, 152*(3), 533-542.

Veehof, M. M., Trompetter, H. R., Bohlmeijer, E. T., & Schreurs, K. M. G. (2016). Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. *Cognitive Behaviour Therapy, 45*, 5-31.

Vlaeyen, J.W.S., & Linton, S.J. (2000). Fear-avoidance and its consequences in musculoskeletal pain: A state of the art. *Pain, 85,* 317-332.

Vowles, K. E., & McCracken, L. M. (2008). Acceptance and values-based action in chronic pain: A study of treatment effectiveness and process. *Journal of Consulting and Clinical Psychology, 76,* 397-407.

Vowles, K. E., McCracken, L. M., & O’Brien, J. Z. (2011). Acceptance and values-based action in chronic pain: a three-year follow-up analysis of treatment effectiveness and process. *Behaviour Research and Therapy, 49,* 748-755.

Vowles, K. E., McCracken, L. M., Sowden, G., & Ashworth, J. (2014). Psychological flexibility in coping with chronic pain: Further examination of the Brief Pain Coping Inventory-2. *The Clinical Journal of Pain, 30*(4), 324-330.

Wagener, A. L., & Zettle, R. D. (2011). Targeting fear of spiders with control-, acceptance-, and information-based approaches. *The Psychological Record*, *61*(1), 77-91.

Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care, 34(3),* 220-233.

Wechsler, D. (2011). Test of premorbid functioning. UK version (TOPF UK). *UK: Pearson Corporation*.

Wetherell, J. L., Afari, N., Rutledge, T., Sorrell, J. T., Stoddard, J. A., Petkus, A. J., & Atkinson, J. H. (2011). A randomized, controlled trial of acceptance and commitment therapy and cognitive-behavioral therapy for chronic pain. *Pain*, *152*, 2098-2107.

Wicksell, R. K., Ahlqvist, J., Bring, A., Melin, L. & Olsson, G. L. (2008a). Can exposure and acceptance strategies improve functioning and life satisfaction in people with chronic pain and whiplash-associated disorders (WAD)? A randomized controlled trial. *Cognitive Behaviour Therapy, 37,* 169-82.

Wicksell, R., Kemani, M., Jensen, K., Kosek, E., Kadetoff, D., Sorjonen, K., Ingvar, M. & Olsson, G. (2013). Acceptance and commitment therapy for fibromyalgia: A randomized controlled trial. *European Journal of Pain, 17,* 599-611.

Wicksell, R.K., Melin, L., Lekander, M., Olsson, G.L. (2009). Evaluating the effectiveness of exposure and acceptance strategies to improve functioning and quality of life in longstanding paediatric pain – A randomized controlled trial. *Pain, 141,* 248–257.

Wicksell, R. K., Olsson, G. L. & Hayes, S. C. (2010). Psychological flexibility as a mediator of improvement in Acceptance and Commitment Therapy for patients with chronic pain following whiplash. *European Journal of Pain, 14,* 1059.e1-1059.e11.

Wicksell, R. K., Renöfält, J., Olsson, G. L., Bond, F. W., & Melin, L. (2008b). Avoidance and cognitive fusion–central components in pain related disability? Development and preliminary validation of the Psychological Inflexibility in Pain Scale (PIPS). *European Journal of Pain*, *12*(4), 491-500.

Williams, A. C. D. C., & Craig, K. D. (2016). Updating the definition of pain. *Pain*, *157*(11), 2420-2423.

Williams, A., Eccleston C., & Morley, S. (2012). *Psychological therapies for the management of chronic pain (excluding headache) in adults.* Cochrane Database Systematic Reviews. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23152245.

Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.

Willis, G.B., & Artino, A.R. (2013). What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *Journal of Graduate Medical Education*, *5*, 353–356.

Willis, G. B., DeMaio, T., & Harris-Kojetin, B. (1999). Is the bandwagon headed to the methodological promised land? Evaluating the validity of cognitive interviewing techniques. *Cognition and survey research*, *133*, 153.

Wilson, K. G., Sandoz, E. K., Kitchens, J., & Roberts, M. (2010). The Valued Living Questionnaire: Defining and measuring valued action within a behavioral framework. *The Psychological Record*, *60*(2), 249-272.

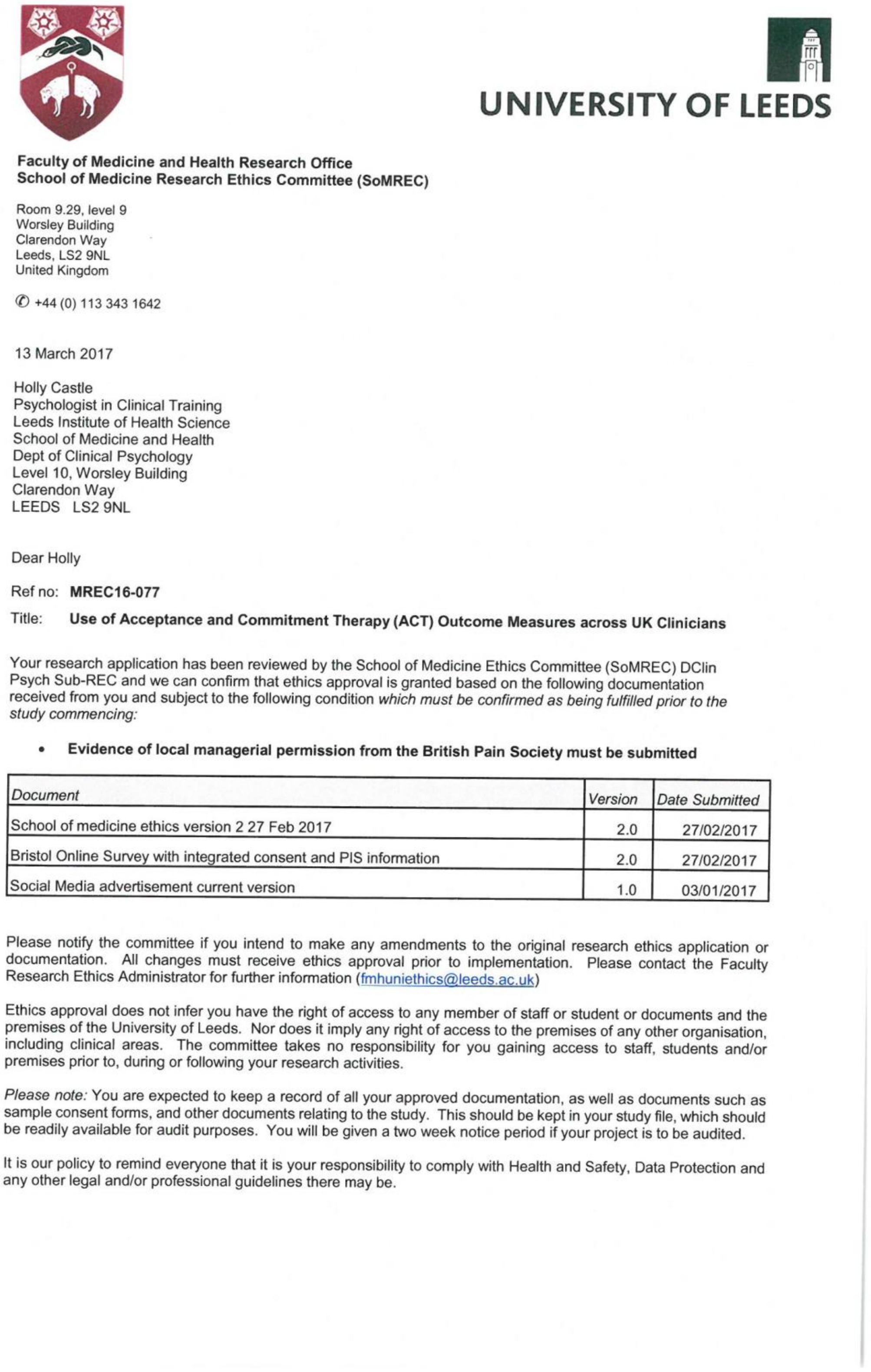
Wolfgast, M. (2014). What does the acceptance and action questionnaire (AAQ-II) really measure? *Behaviour Therapy*, *45*, 831-839.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361-370.

# APPENDICES

## Appendix A. Ethical approval and documentation

### A.1 Study Part One: (Clinician survey) ethics approval letter



UNIVERSITY OF LEEDS

**Faculty of Medicine and Health Research Office**
**School of Medicine Research Ethics Committee (SoMREC)**

Room 9.29, level 9
Worsley Building
Clarendon Way
Leeds, LS2 9NL
United Kingdom

✆ +44 (0) 113 343 1642

13 March 2017

Holly Castle
Psychologist in Clinical Training
Leeds Institute of Health Science
School of Medicine and Health
Dept of Clinical Psychology
Level 10, Worsley Building
Clarendon Way
LEEDS LS2 9NL

Dear Holly

Ref no: **MREC16-077**

Title: **Use of Acceptance and Commitment Therapy (ACT) Outcome Measures across UK Clinicians**

Your research application has been reviewed by the School of Medicine Ethics Committee (SoMREC) DClin Psych Sub-REC and we can confirm that ethics approval is granted based on the following documentation received from you and subject to the following condition *which must be confirmed as being fulfilled prior to the study commencing:*

- **Evidence of local managerial permission from the British Pain Society must be submitted**

| *Document* | *Version* | *Date Submitted* |
|---|---|---|
| School of medicine ethics version 2 27 Feb 2017 | 2.0 | 27/02/2017 |
| Bristol Online Survey with integrated consent and PIS information | 2.0 | 27/02/2017 |
| Social Media advertisement current version | 1.0 | 03/01/2017 |

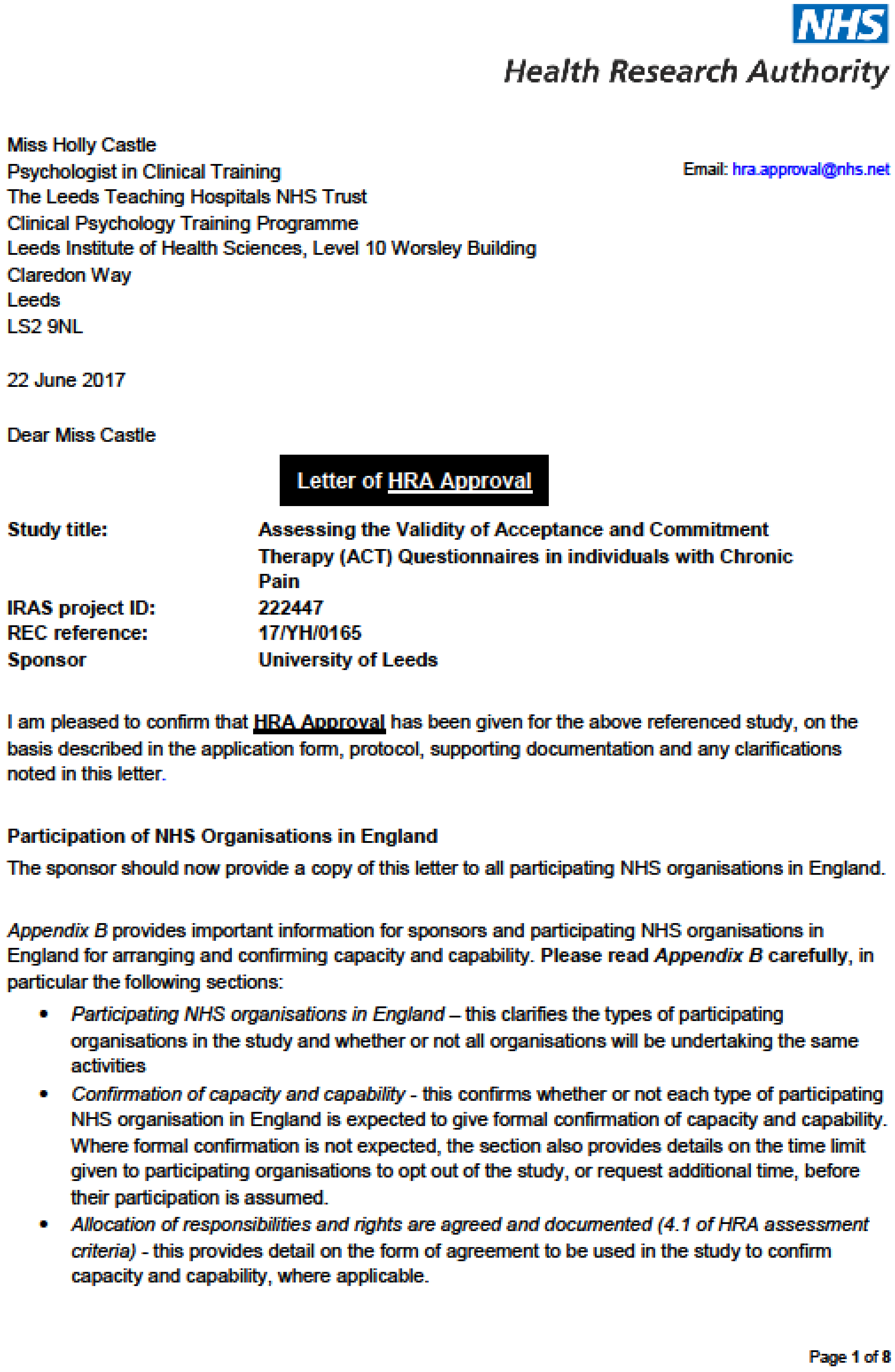
Please notify the committee if you intend to make any amendments to the original research ethics application or documentation. All changes must receive ethics approval prior to implementation. Please contact the Faculty Research Ethics Administrator for further information (fmhuniethics@leeds.ac.uk)

Ethics approval does not infer you have the right of access to any member of staff or student or documents and the premises of the University of Leeds. Nor does it imply any right of access to the premises of any other organisation, including clinical areas. The committee takes no responsibility for you gaining access to staff, students and/or premises prior to, during or following your research activities.

*Please note:* You are expected to keep a record of all your approved documentation, as well as documents such as sample consent forms, and other documents relating to the study. This should be kept in your study file, which should be readily available for audit purposes. You will be given a two week notice period if your project is to be audited.

It is our policy to remind everyone that it is your responsibility to comply with Health and Safety, Data Protection and any other legal and/or professional guidelines there may be.

## A.2 Study Part Two: (Questionnaire validation study) NHS ethics approval letter



NHS
Health Research Authority

Miss Holly Castle
Psychologist in Clinical Training
The Leeds Teaching Hospitals NHS Trust
Clinical Psychology Training Programme
Leeds Institute of Health Sciences, Level 10 Worsley Building
Claredon Way
Leeds
LS2 9NL

Email: hra.approval@nhs.net

22 June 2017

Dear Miss Castle

**Letter of HRA Approval**

| | |
|---|---|
| **Study title:** | **Assessing the Validity of Acceptance and Commitment Therapy (ACT) Questionnaires in individuals with Chronic Pain** |
| **IRAS project ID:** | **222447** |
| **REC reference:** | **17/YH/0165** |
| **Sponsor** | **University of Leeds** |

I am pleased to confirm that HRA Approval has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

**Participation of NHS Organisations in England**

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

*Appendix B* provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

## A.3 Study Part One: recruitment invitation, consent information and online survey

*You are being invited to participate in a research study titled:*

**‘Use of Acceptance and Commitment Therapy (ACT) Outcome Measures across UK Clinicians.’**

*This study is being done by Holly Castle from the University of Leeds. As a practicing ACT clinician, we would like to invite you to take part in an online survey. Before you decide whether you would like to take part we would like you to understand why the research is being done and what it would involve.*

***What is the purpose of the study?***

*Whilst the use of the questionnaires is well documented across the research literature (Hann & McCracken, 2014) their clinical usage remains unclear. The aim of this study is to explore the clinical usage and utility of ACT questionnaires. It will explore initial thoughts from clinicians on the usefulness of these questionnaires and how often they are used in practice. This data will supplement the research literature on use of ACT measures across RCT’s. Ultimately, this will inform a later study looking at the content validity of the ACT questionnaires.*

***What would taking part involve?***

*It will take you approximately 5-10 minutes to complete. Your participation in this study is entirely voluntary and you do not have to answer any questions you do not want to. If you decide not to carry on with the study any data collected up to your withdrawal will not be automatically submitted, should you exit the survey.*

*We believe there are no known risks associated with this research study; however, as with any online related activity the risk of a breach is always possible. To the best of our ability your participation in this study will remain confidential, and only anonymised data will be published. We will minimise any risks by converting the email address you give for consent to an anonymous participant number. Any personal identifying information (e.g names) you include in your answers will be removed before data is analysed or presented. We are confident that presented data will not enable identification of any one participant.*

*Please tick the boxes as they apply before proceeding with the survey:*

*I confirm that I have read and understand the information provided for the above study.* ☐

*I am a practicing ACT clinician.* ☐

*I agree to take part in the online survey.* ☐

*Please provide your email address as an indication you consent to taking part in this online survey:*

---

*(The survey will request that the participant selects each box and will only move onto the next screen if all are completed.)*

**PART ONE Background Information**

1) How long have you practiced as an ACT clinician?

______________________________________________________________

2) In which country do you practice as an ACT clinician?

______________________________________________________________

3) Do you use any other therapeutic models in your clinical practice?
Yes No

If you selected Yes, please specify:

______________________________________________________________

**PART TWO Use of ACT Outcome measures**

4) Do you regularly use ACT with clients experiencing chronic pain?
Yes No

a) Which of the following ACT questionnaires have you used with your chronic pain clients?

Chronic Pain Acceptance Questionnaire (CPAQ)
Chronic Pain Values Inventory (CPVI)
Psychological Inflexibility in Pain Scales (PIPS)
Acceptance and Action Questionnaire (AAQ/AAQ-ll)
Mindfulness Attention Awareness Scale (MAAS)
Other

If you selected Other, please specify:

______________________________________________________________

5) In your clinical practice how often do you use each questionnaire with your clients?

| | With none of my clients | With some of my clients | With about half of my clients | With most of my clients | With all of my clients |
|---|---|---|---|---|---|
| CPAQ | | | | | |
| CPVI | | | | | |
| PIPS | | | | | |
| AAQ/AAQ-II | | | | | |
| MAAS | | | | | |
| Other ACT questionnaire | | | | | |

(a) If you ticked 'Other' ACT questionnaire please specify which and how often you use it below:

____________________________________________________________

1) In working with chronic pain, which ACT questionnaires – including those listed above – do you think are most useful and why?

____________________________________________________________

2) In working with chronic pain, which ACT questionnaires do you find most problematic and why?

____________________________________________________________

3) Clinically, how useful in general do you find ACT questionnaires?

____________________________________________________________

**PART THREE- Finished! Thank you very much for your participation.**

**What happens with the results?**

If I get enough responses this will allow me to see how ACT questionnaires are used amongst clinicians. I am also interested in how clinically useful the questionnaires are in practice. I plan to give a summary of the results to the British Pain Society.

**Further information and contact details**

This study has also been reviewed and given favourable opinion by the University of Leeds Research Ethics Committee. If you would like any further information about the research, please contact either of the following:

**Twitter advertisement**

Seeking #ACT Clinicians working in Chronic #Pain for an online questionnaire study:
ACT Clinicians with people with Chronic #Pain needed for an online study of ACT outcome measures:
Do you practice #ACT for chronic #pain? We want to know which questionnaires are most useful.
*note. We may address these tweets to researchers in the area.

**Facebook advertisement**

Two alternative titles:
*Seeking ACT clinicians who work in Chronic Pain: we would be really grateful if you would help us with a quick online survey about which ACT questionnaires you use.*
*We are seeking current ACT clinicians who work with people experiencing Chronic Pain for an online questionnaire study about ACT outcome measures*.

## A.4 Study Part Two: recruitment invitation and information sheet

UNIVERSITY OF LEEDS

The Leeds Teaching Hospitals NHS
NHS Trust

### ‘Study of Acceptance and Commitment Therapy (ACT) Questionnaire Validity in Individual’s with Chronic Pain.’

*We would like to invite you to take part in an interview study to help us improve the accuracy of some questionnaires. This study is being conducted as part of my Doctorate in Clinical Psychology at the University of Leeds. We hope to better understand patient perceptions of the questionnaires in this study. This study will require one meeting with myself which can be arranged at a convenient time and location. Before you decide whether you would like to take part we would like you to understand why the research is being done and what it would involve. Talk to others about the study if you wish. Ask us if there is anything that is not clear.*

### Why have I been invited?

We have good evidence that a type of psychotherapy called Acceptance and Commitment Therapy (ACT) is helpful for many people who are living with chronic pain (Wicksell, Melin, Lekander & Olsson, 2009). In clinical trials we often look for changes in questionnaires to let us know whether the treatment is actually helpful. However, we suspect that some of the questionnaires that have been used in these studies might not accurately measure what they intend to measure, and therefore may not be useful in clinical trials. Therefore, to improve the accuracy of future clinical trial results, we want to check whether these questionnaires actually measure what they are supposed to measure. Thus, we would like to invite you to take part in a interview study to help us improve the accuracy of some questionnaires (Beerlage-de Jong et al., 2015).

### What is the purpose of the study?

The aim of this study is to explore accuracy of chronic pain questionnaires. The study will be focusing on participants understanding and perception of individual questions which make up several questionnaires. Clinicians use questionnaires to measure changes made throughout the course of therapy. Sometimes we don’t know which parts of therapy are helpful for people experiencing chronic pain. Exploring if these questionnaires measure what they are intended to measure, will tell us more about *how* changes are made during therapy. Ultimately, this may help improve the effectiveness of therapy. We hope to better understand patient perceptions of the questionnaires in this study

### Do I have to take part?

No, it is up to you to decide to join the study. We will describe the study and go through this information sheet with you, answering any questions you have. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. This would not affect the standard of care you receive.

**What would taking part involve? What will happen to me if I take part?**

This study happens in one session and we will endeavour to make sure it lasts no longer than 60-75 minutes. We will initially go through the consent procedure with you and ask you some questions on the type of pain you experience, which type of therapy you are attending and give you a brief test of reading ability prior to beginning the study. Within the study session there are two parts which will both be audio-recorded. The first part of the study will ask participants to put each card into piles which they feel go together. The second part of the study will require that each participant says out-loud what they are thinking as they answer each question on 4 different questionnaires. (The questions will be written on pieces of card). This will help us to understand possible problems with individual questionnaire items. The study will take place at Fielding House or at your home address, depending on your preferred option. The research will require that you can meet the researcher once.

**What are the possible benefits of taking part?**

It cannot be promised that this study will help you directly. However, we hope that the information we get from this study helps to improve our measurement of therapy and will thus lead to improved treatments for people experiencing chronic pain. In acknowledgement of the time taken by participants to be involved in this study a shopping gift voucher of the value of £5 will be provided as a thank-you for taking part.

**What are the possible disadvantages and risks of taking part?**

The time required in both taking part and travelling to the location of this study is a potential inconvenience. To minimise this we can be flexible over the location of the study and can complete the study at Fielding House or your home depending on where you would find more convenient. Also the session will last no longer than 75 minutes. The questionnaires included in the study are used as standard in treatment and we do not expect them to evoke any distress. However, should you experience any distress whilst completing this study we would like to remind you that you can withdraw at any time. You may choose to use any existing support for example the current support you receive from the pain clinic. You could also make contact with the **Single Point of Access (SPA) on 0300 300 1485.** The SPA is available 24 hours, seven days a week. Staff at the SPA are from our crisis teams and are experienced in supporting people with mental health problems. They will initially offer telephone advice and support but will undertake face to face assessments if required.

**What happens when the research study stops?**

After the study has finished all participants will be given the option of receiving information on the final results of the study. If you would like to receive the final results of the study please indicate in the consent form.

**What if there is a problem?**

Any complaint about the way you have been dealt with during the study or any possible harm you might have suffered will be addressed. More detailed information on this is given in Further Information and Contact details.

**What will happen if I don't want to carry on with the study?**

If this study needs to be stopped for any reason we will tell you and inform your therapist. You should also know that you could decide at any point you that you do not want to carry on with the study. If you decide not to carry on with the study we will continue to use the data collected up to your withdrawal. This data would remain in the study and be processed and analysed along with the data for all other participants.

### What will happen if I do want to carry on with the study?

All research within in the NHS is looked at by independent group of people, called the Health Research Authority (HRA), to protect your interests. This study has also been reviewed and given favourable opinion by the NHS Research Ethics Committee (REC). Before we begin the session I will go through the Information sheet and consent form with you should you have any questions. You will be given a copy of the consent form and participant information sheet to keep and the consent to take part will also be recorded in your case notes. Dr Laidler will be advised that you have taken part in this study. Your GP will also be advised that you are taking part in the study. It should also be acknowledged that indemnity for this research is covered by the University of Leeds Public liability policy.

### What will happen to the results of the research study?

It should also be noted that the findings from the report may be presented at academic conferences. Again no one will be able to identify you from the participant number allocated and at no point will your identity be divulged. The researcher is using this study for the purposes of completing an academic qualification (Doctorate in Clinical Psychology).

### Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. All information which is collected about you during the course of my research project will be kept strictly confidential, and any information about you which leaves Fielding House will have your name and address removed so that you cannot be recognised. The information provided by yourself for this study will be entered into a statistical programme on a computer to enable me to analyse it. You will remain anonymous as your name will be converted to a participant number. The only occasion when confidentiality would be broken, and information passed onto a third party, would be in the event of a disclosure that you or somebody else had been harmed or was at risk of being harmed in some way. If such an issue becomes apparent during the study then the Leeds Teaching Hospital NHS Trust safeguarding policy would be adhered to. This would involve informing the relevant authorities; the Leeds Teaching Hospitals NHS Trust Safeguarding Team (which includes a Safeguarding Vulnerable Adults at Risk named professional or the Trust's Children's Safeguarding team).

### What Happens Now?

Please take your time to read this Information Sheet and consider whether you would like to take part in this research study and ask friends and family about it if you wish. If you wish to take part please put your contact number on the tear off slip below and return it to Dr Laidler within the next 4 weeks. I will then contact you over the phone to discuss your interest and arrange for us to meet up for the research study. Hopefully we will be able to arrange a convenient time for us to meet (this could be at Fielding House or at your home address if this is more convenient for you). I will enrol you in the research study by getting you to complete a Consent Form when we meet. Please feel free to contact me using the details below if you have any questions about the research study.
Further information and contact details

*The principal investigator:*
Holly Castle, Trainee Clinical Psychologist, University of Leeds,
e-mail address: psc4h2c@leeds.ac.uk

*The primary research supervisor:*
Dr Christopher Graham, Fellow in Behavioural Medicine & Clinical Psychologist, University of Leeds,
e-mail address: C.D.Graham@leeds.ac.uk
Telephone: 0113 343 3910

*The field supervisor:*
Dr Vivienne Laidler, Senior Clinical Psychologist, St James's Hospital (Fielding House),
e-mail address: Vivienne.laidler@nhs.net
Telephone: 0113 206 5897

If you have any concerns about any aspect of this study, you should ask to speak to the researcher who will do their best to answer your questions. You can also approach:

Clare Skinner: *The Faculty Head of Research Support, University of Leeds*
e-mail address: C.E.Skinner@leeds.ac.uk
Telephone: 0113 343 4897
She will be able to advise you on the formal complaint procedure if you are unhappy with any aspect of the study.

I would like to express my gratitude to the Everybody's Voice Group who were consulted in the development of this Information sheet**.** If you would like any further information about the research please contact any of the following:

Thank you very much for taking the time to read this.  With warm regards,

……………………………………………………………………………………………………

**(If you are interested in being contacted about the study then please return this section of the invitation back to Dr Vivienne Laidler)**

**Phone number………………….**

## A.5 Study Part Two consent forms

UNIVERSITY OF LEEDS

The Leeds Teaching Hospitals NHS Trust

**'Study of Acceptance and Commitment Therapy**

**(ACT) Questionnaire Validity in Individual's with Chronic Pain.'**

Name of Researcher: Holly Castle

**Please initial box**

1. I confirm that I have read and understand the information sheet dated 14/03/17 (version 1) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐

3.I agree to take part in the above study. ☐

4. I understand that this study involves the audio recording of my interview with the researcher and that no identifying information will be associated with this recording or the transcript. ☐

5. I agree for my GP/Dr Laidler to be advised of my participation in this study. ☐

6. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from the University of Leeds, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

7. Would you like to receive information on the findings from this study at a later date? (A summary of the findings will be sent to your email address) Please indicate: Yes or No

Your name: ______________________ Date: ________________ Signature: ______________

……………………………………………………………………………………………………

When completed: 1 for participant; 1 for researcher site file; 1 original to be kept in case notes

Name of person taking consent: __________ Date: ______________ Signature: ________________

# Appendix B. Test materials (Study Part Two)

## B.1 Participant Instruction Sheet

### Participant Instruction Sheet

We are interested in how people with experience of chronic pain respond to questionnaires. To explore this, I'm going to now ask you to first do a task which involves telling us your thought processes when answering the questionnaires. This will be followed by a simple sorting task which will involve you organising individual question items into piles which you feel are similar.

**PART ONE**

'I'm going to now ask you to say out loud your thought processes as you answer each question. Each of the statements on these cards ask you to rate out of 7 how much you agree with them. 1 being never true to 7 always being true'.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| **never true** | **very seldom true** | **seldom true** | **sometimes true** | **frequently true** | **almost always true** | **always true** |

'I want you to THINK your answer out for me and say it ALOUD. By this I mean I want you to describe everything you are thinking from when you read the card.'

Try not to plan what you're saying or try and explain what you're saying. Try and act like you're in the room by yourself and most importantly just remember to try and keep talking.

Please try and speak as LOUDLY and CLEARLY as you can.

So, I would like you to:

- **THINK your answer out for me and**
- **say it ALOUD.**
- **Describe everything you are thinking from when you read the card**

**Example question**

***'I can cope with my pain in most situations'***

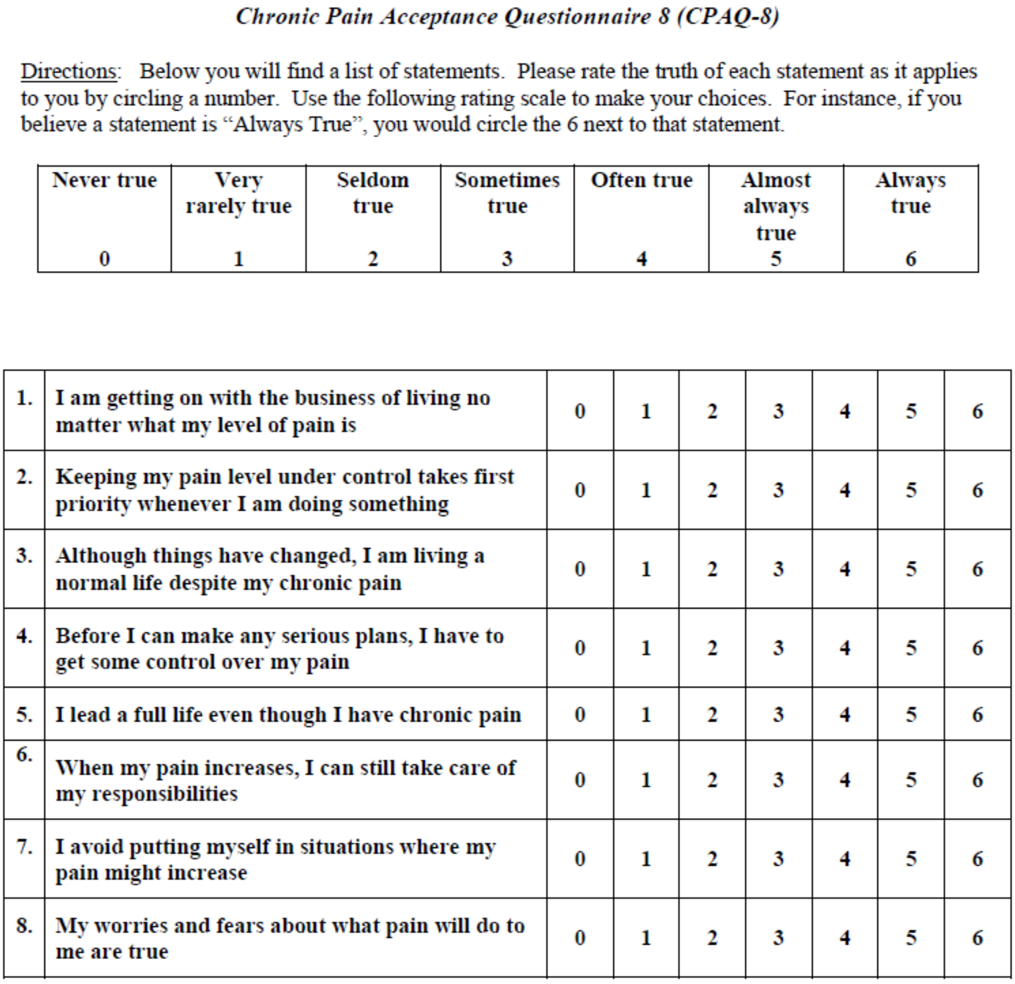
**Example answer**

***'I'm thinking that when I'm with my family I'm able to cope with the pain, however when I'm at work it becomes much more difficult- therefore I would rate this question a 3.'***

**PART TWO**

Place each card into piles which you feel are SIMILAR and would go together.

### B.2 Chronic Pain Acceptance Questionnaire 8 (CPAQ-8; Fish et al., 2010)



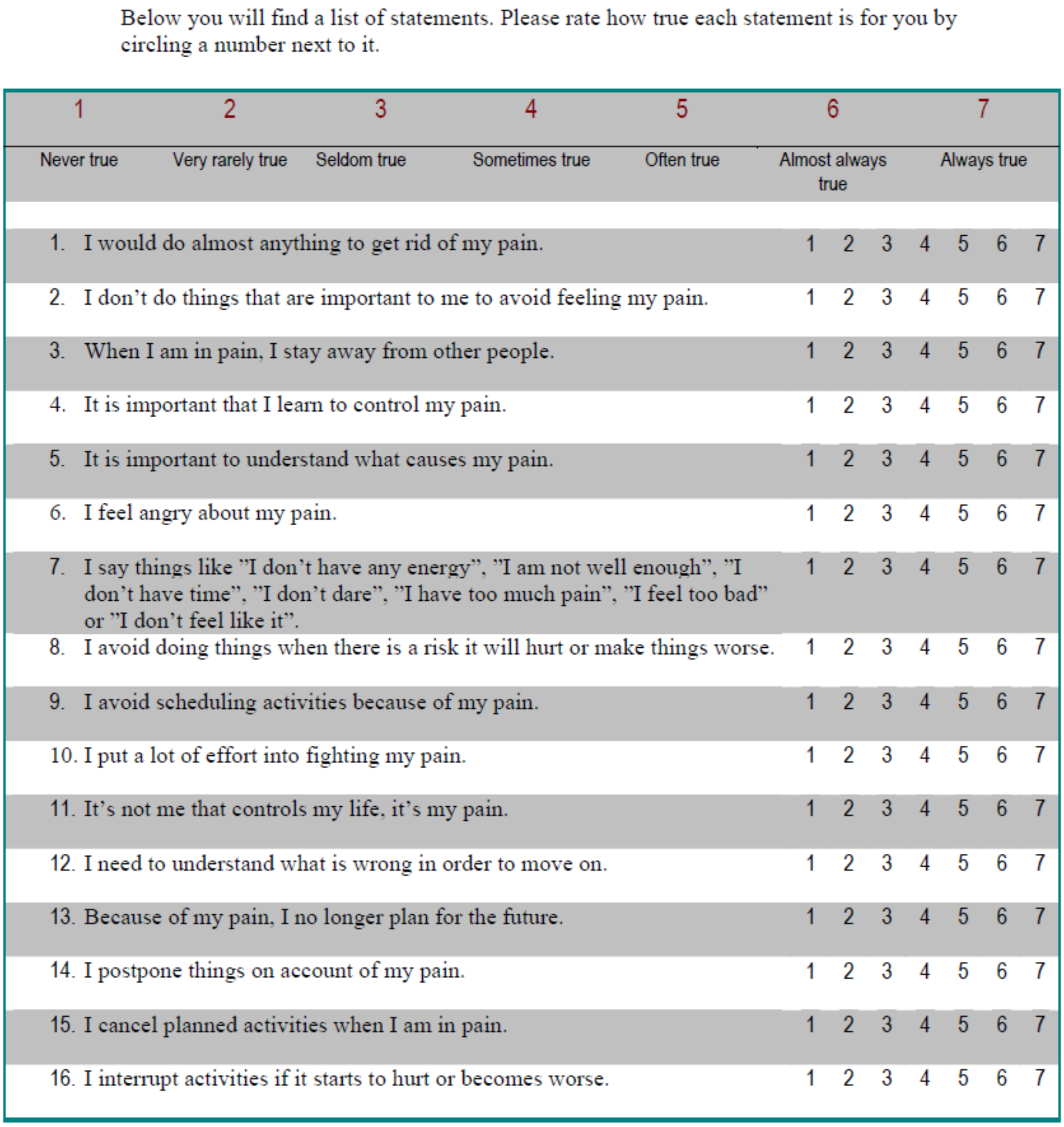
*Chronic Pain Acceptance Questionnaire 8 (CPAQ-8)*

Directions: Below you will find a list of statements. Please rate the truth of each statement as it applies to you by circling a number. Use the following rating scale to make your choices. For instance, if you believe a statement is "Always True", you would circle the 6 next to that statement.

| Never true | Very rarely true | Seldom true | Sometimes true | Often true | Almost always true | Always true |
|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. | I am getting on with the business of living no matter what my level of pain is | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2. | Keeping my pain level under control takes first priority whenever I am doing something | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 3. | Although things have changed, I am living a normal life despite my chronic pain | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 4. | Before I can make any serious plans, I have to get some control over my pain | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 5. | I lead a full life even though I have chronic pain | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 6. | When my pain increases, I can still take care of my responsibilities | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 7. | I avoid putting myself in situations where my pain might increase | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 8. | My worries and fears about what pain will do to me are true | 0 | 1 | 2 | 3 | 4 | 5 | 6 |

## B.2 Psychological Inflexibility in Pain Scale (PIPS; Wicksell et al., 2008)

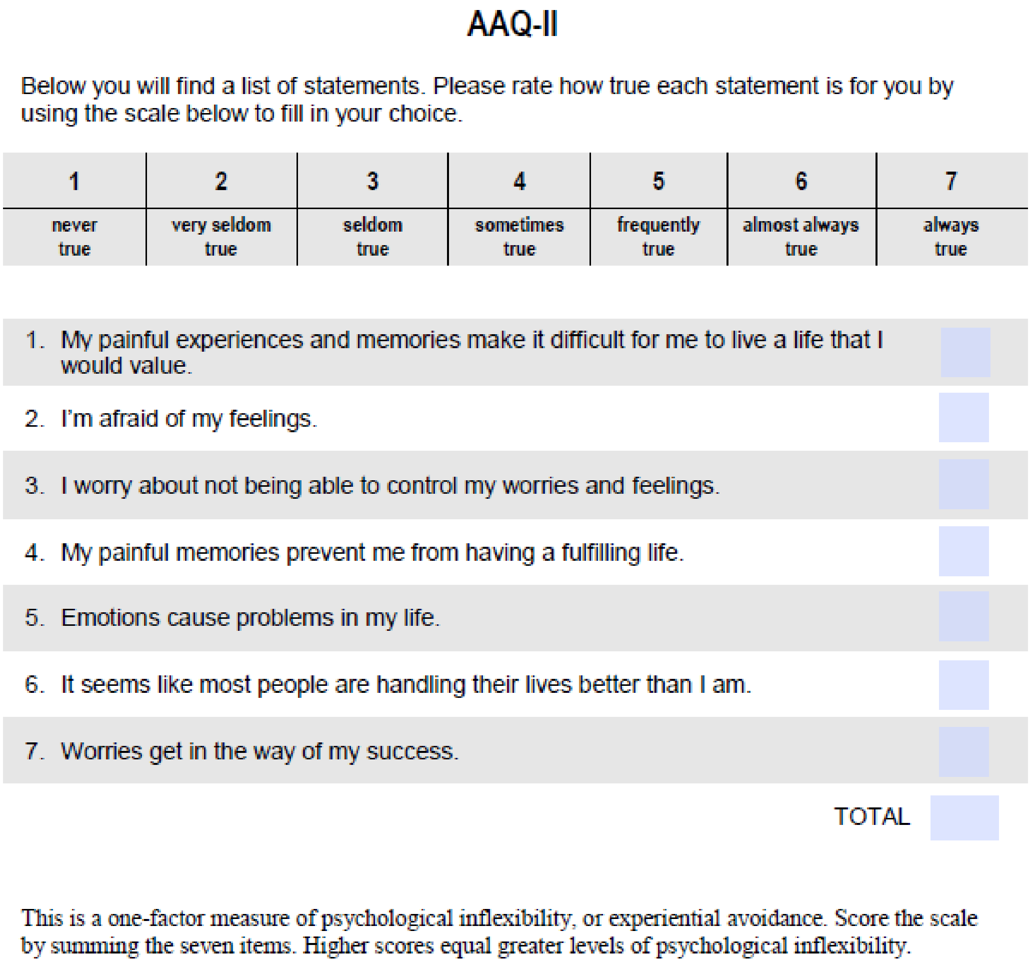


Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Never true | Very rarely true | Seldom true | Sometimes true | Often true | Almost always true | Always true |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. I would do almost anything to get rid of my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2. I don't do things that are important to me to avoid feeling my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 3. When I am in pain, I stay away from other people. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 4. It is important that I learn to control my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 5. It is important to understand what causes my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 6. I feel angry about my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 7. I say things like "I don't have any energy", "I am not well enough", "I don't have time", "I don't dare", "I have too much pain", "I feel too bad" or "I don't feel like it". | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8. I avoid doing things when there is a risk it will hurt or make things worse. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 9. I avoid scheduling activities because of my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 10. I put a lot of effort into fighting my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 11. It's not me that controls my life, it's my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 12. I need to understand what is wrong in order to move on. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 13. Because of my pain, I no longer plan for the future. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 14. I postpone things on account of my pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 15. I cancel planned activities when I am in pain. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 16. I interrupt activities if it starts to hurt or becomes worse. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

## B.3 Acceptance and Action Questionnaire - II (AAQ-II; Bond et al., 2011)



### AAQ-II

Below you will find a list of statements. Please rate how true each statement is for you by using the scale below to fill in your choice.

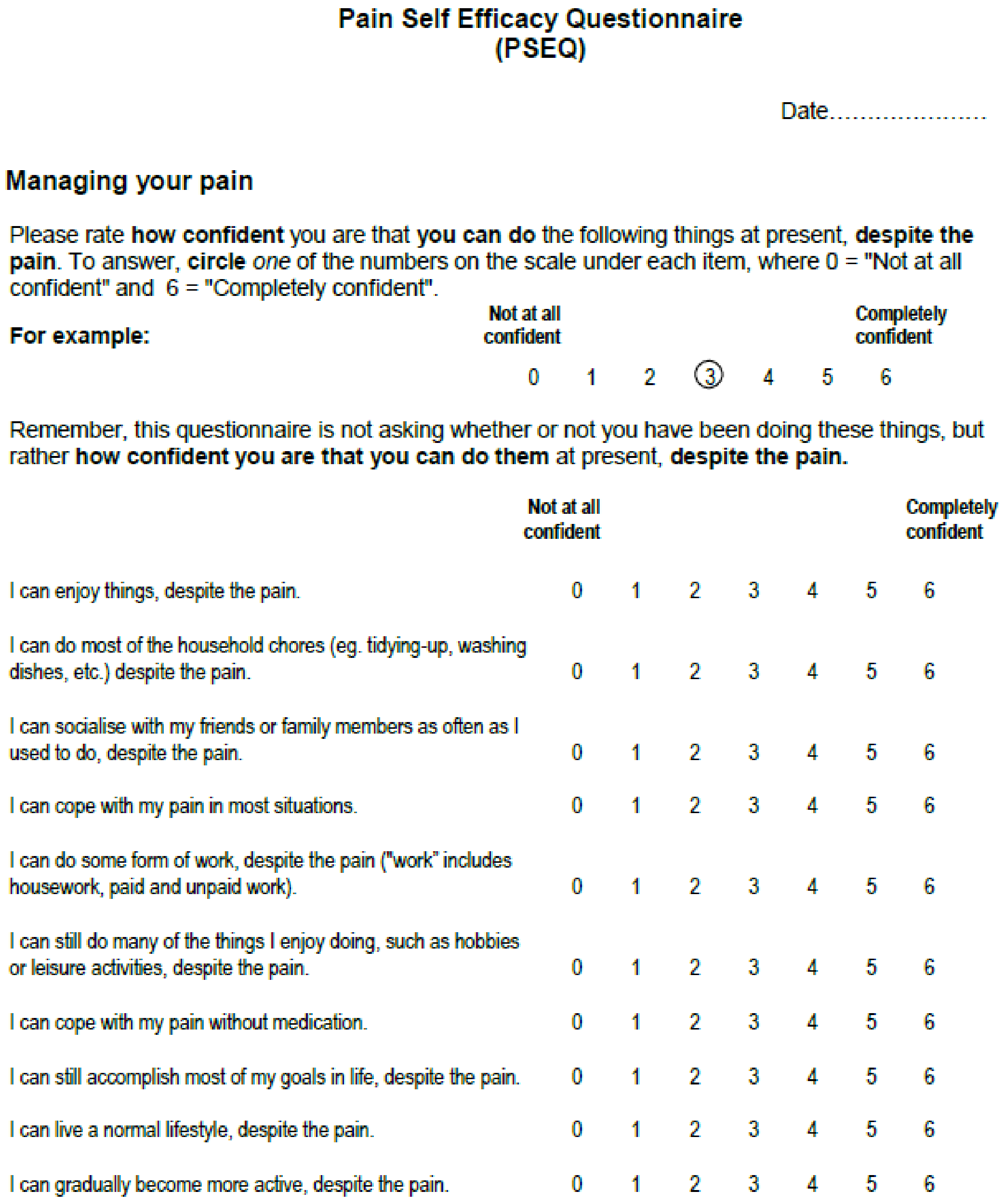
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| never true | very seldom true | seldom true | sometimes true | frequently true | almost always true | always true |

1. My painful experiences and memories make it difficult for me to live a life that I would value.
2. I'm afraid of my feelings.
3. I worry about not being able to control my worries and feelings.
4. My painful memories prevent me from having a fulfilling life.
5. Emotions cause problems in my life.
6. It seems like most people are handling their lives better than I am.
7. Worries get in the way of my success.

TOTAL

This is a one-factor measure of psychological inflexibility, or experiential avoidance. Score the scale by summing the seven items. Higher scores equal greater levels of psychological inflexibility.

## B.4 Pain Self Efficacy Questionnaire (PSEQ; Nicholas, 1989)



### Pain Self Efficacy Questionnaire
### (PSEQ)

Date....................

**Managing your pain**

Please rate **how confident** you are that **you can do** the following things at present, **despite the pain**. To answer, **circle** *one* of the numbers on the scale under each item, where 0 = "Not at all confident" and 6 = "Completely confident".

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **For example:** | **Not at all confident** | | | | | | **Completely confident** |
| | 0 | 1 | 2 | [✓ 3] | 4 | 5 | 6 |

Remember, this questionnaire is not asking whether or not you have been doing these things, but rather **how confident you are that you can do them** at present, **despite the pain.**

| | **Not at all confident** | | | | | | **Completely confident** |
|---|---|---|---|---|---|---|---|
| I can enjoy things, despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can do most of the household chores (eg. tidying-up, washing dishes, etc.) despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can socialise with my friends or family members as often as I used to do, despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can cope with my pain in most situations. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can do some form of work, despite the pain ("work" includes housework, paid and unpaid work). | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can still do many of the things I enjoy doing, such as hobbies or leisure activities, despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can cope with my pain without medication. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can still accomplish most of my goals in life, despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can live a normal lifestyle, despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| I can gradually become more active, despite the pain. | 0 | 1 | 2 | 3 | 4 | 5 | 6 |

# Appendix C. Card Sort methodology description

Card sort data was not analysed as was beyond the scope of this thesis and involved a separate aspect of construct validity; *construct representation*. The research aims and hypothesis of the present study were focused on checking content validity; i.e. assessing understanding of individual questionnaire items. Whilst construct representation may also help to establish whether responses are indeed reflective of the intended constructs measured by the questionnaire (Strauss & Smith, 2009). This aspect of validity looks a different yet important part of questionnaire overall validity and therefore will be analysed following the completion of the thesis.

Card sorting methodology has been used to evaluate the validity of questionnaires and their constructs from the perspective of the individual completing it (Beerlage-de Jong, Kulyk, Wentzel, & Oinas-Kukkonen, 2015). During a card sort participants are asked to group items into piles which 'go together', and should these piles correlate with the intended process variables, will tell us if they measure what they intend to measure. This closely parallels natural thinking patterns, is therefore not taxing for participants or time-consuming and can be used in large numbers (Coxon, 1999).

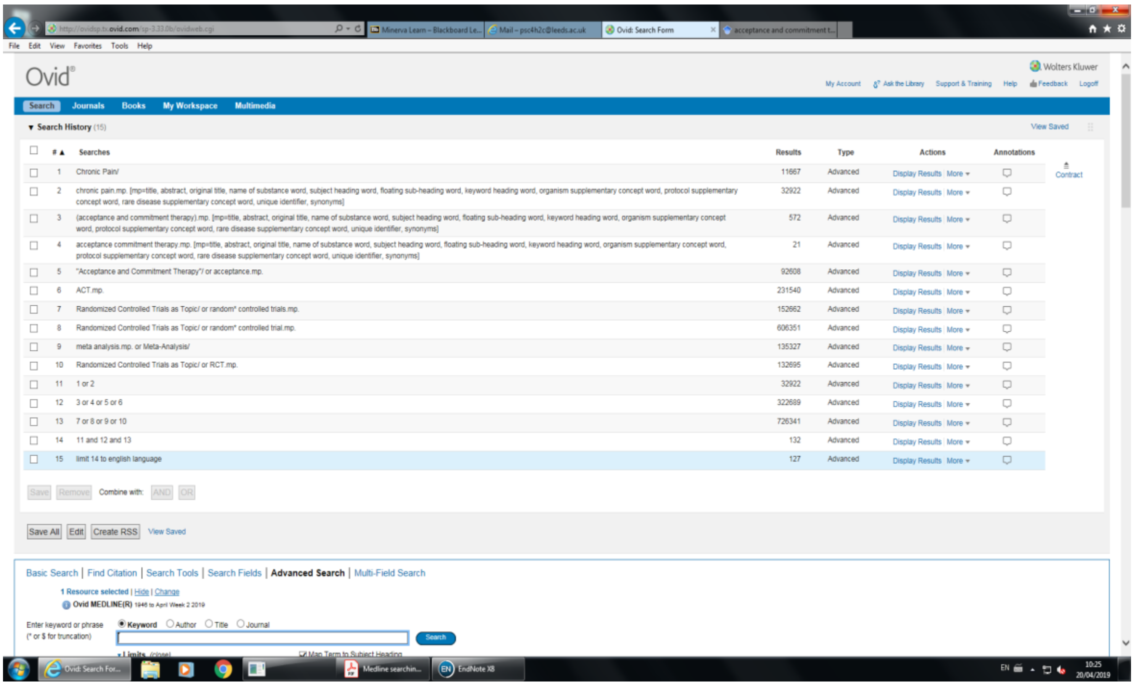
**Method:** Participants completed a card sort task following cognitive interviewing. This involved participants placing each of the 31 items into piles which they felt went together and providing a brief explanation as to why they organised the items in that way. Each pile was allocated a summary statement to describe what each pile represented. This took around 10 minutes to complete. There was an unspecified number of piles which varied between participants. As no pre-determined set of category names were provided, this constituted an 'open' card sorting methodology.

**Proposed analysis:** Resulting data can be statistically analysed for significant differences between individual's perceptions of the questionnaire and the process the questionnaire author intended to capture. Hence, the card sorting task assessing non-verbal thinking processes. This method can be used to evaluate how people organise large numbers of questionnaire items. The resulting data provides "*native classifications of phenomena, how people perceive their environment, what they attend to when making discriminations* *between classes of things and events; to get descriptive categories which are 'psychologically real,'' (Black, 1963, p. 1348).* Card sorting data comes from a respondent placing individual card or 'items' into piles which then become distinct categories. This data can give us information on patterns of difference between the items themselves and be combined with each participant's data (Coxon, 1999). Should the card sorting task reflect the target constructs of interest, this may further establish a different aspect of questionnaire validity- *construct representation* (Strauss & Smith, 2009). In addition to producing data which can be analysed quantitatively, Harloff & Coxon (2005) note the usefulness of also attaining brief qualitative information on what caused each participant to organise the piles in the way that they did.

Aggregate confusion matrixs' are used to organise the data and combine individual participant data sets. 2-way matrixes are used for each participant to explore differences between each individual. This is also known as a 'co-occurrence matrix' and can be used to illustrate how many (of 30+ participants) put the row questionnaire item and column questionnaire item into the same category during their card sorting (Coxon, 1999). The co-occurrence matrix displays binary data of 0 to indicate no co-occurrence and 1 to indicate one co-occurrence in one group, for each participant. A cluster analysis can be used to determine partitions in the data. Depending on how individuals sort the piles this gives an indication of how items correlate with those originally intended by the questionnaire authors, and thus gives an indication of the degree to which constructs are represented in the way intended by the questionnaire authors. The qualitative comments provided by participants regarding the separation of the card piles can also be helpful in aiding our understanding as to why the data has been sorted in this way.

# Appendix D. Literature search screenshots

Medline – 20/04/19



Search History (15)

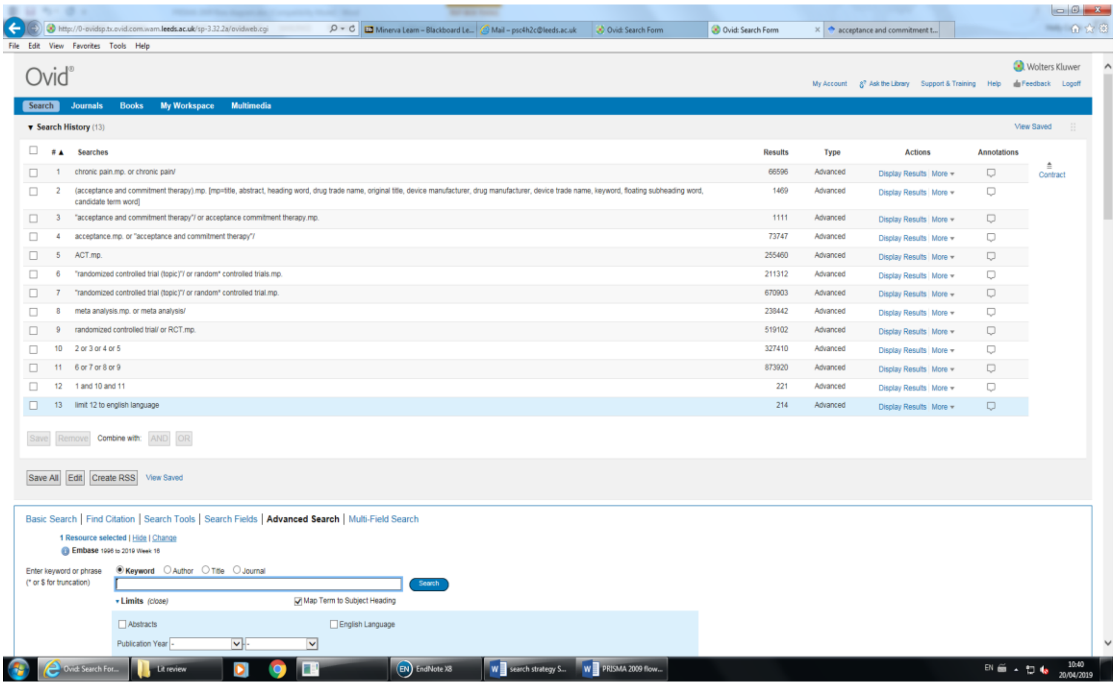
| # | Searches | Results | Type |
|---|---|---|---|
| 1 | Chronic Pain/ | 11667 | Advanced |
| 2 | chronic pain.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] | 32922 | Advanced |
| 3 | (acceptance and commitment therapy).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] | 572 | Advanced |
| 4 | acceptance commitment therapy.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] | 21 | Advanced |
| 5 | "Acceptance and Commitment Therapy"/ or acceptance.mp. | 92608 | Advanced |
| 6 | ACT.mp. | 231540 | Advanced |
| 7 | Randomized Controlled Trials as Topic/ or random* controlled trials.mp. | 152662 | Advanced |
| 8 | Randomized Controlled Trials as Topic/ or random* controlled trial.mp. | 606351 | Advanced |
| 9 | meta analysis.mp. or Meta-Analysis/ | 135327 | Advanced |
| 10 | Randomized Controlled Trials as Topic/ or RCT.mp. | 132695 | Advanced |
| 11 | 1 or 2 | 32922 | Advanced |
| 12 | 3 or 4 or 5 or 6 | 322689 | Advanced |
| 13 | 7 or 8 or 9 or 10 | 726341 | Advanced |
| 14 | 11 and 12 and 13 | 132 | Advanced |
| 15 | limit 14 to english language | 127 | Advanced |

1 Resource selected | Hide | Change
Ovid MEDLINE(R) 1946 to April Week 2 2019

Enter keyword or phrase (* or $ for truncation): ● Keyword ○ Author ○ Title ○ Journal

Limits (close) ☑ Map Term to Subject Heading

EMBASE- 20/04/19



Search History (13)

| # | Searches | Results | Type |
|---|---|---|---|
| 1 | chronic pain.mp. or chronic pain/ | 66596 | Advanced |
| 2 | (acceptance and commitment therapy).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] | 1469 | Advanced |
| 3 | "acceptance and commitment therapy"/ or acceptance commitment therapy.mp. | 1111 | Advanced |
| 4 | acceptance.mp. or "acceptance and commitment therapy"/ | 73747 | Advanced |
| 5 | ACT.mp. | 255460 | Advanced |
| 6 | "randomized controlled trial (topic)"/ or random* controlled trials.mp. | 211312 | Advanced |
| 7 | "randomized controlled trial (topic)"/ or random* controlled trial.mp. | 670903 | Advanced |
| 8 | meta analysis.mp. or meta analysis/ | 238442 | Advanced |
| 9 | randomized controlled trial/ or RCT.mp. | 519102 | Advanced |
| 10 | 2 or 3 or 4 or 5 | 327410 | Advanced |
| 11 | 6 or 7 or 8 or 9 | 873920 | Advanced |
| 12 | 1 and 10 and 11 | 221 | Advanced |
| 13 | limit 12 to english language | 214 | Advanced |

1 Resource selected | Hide | Change
Embase 1996 to 2019 Week 16

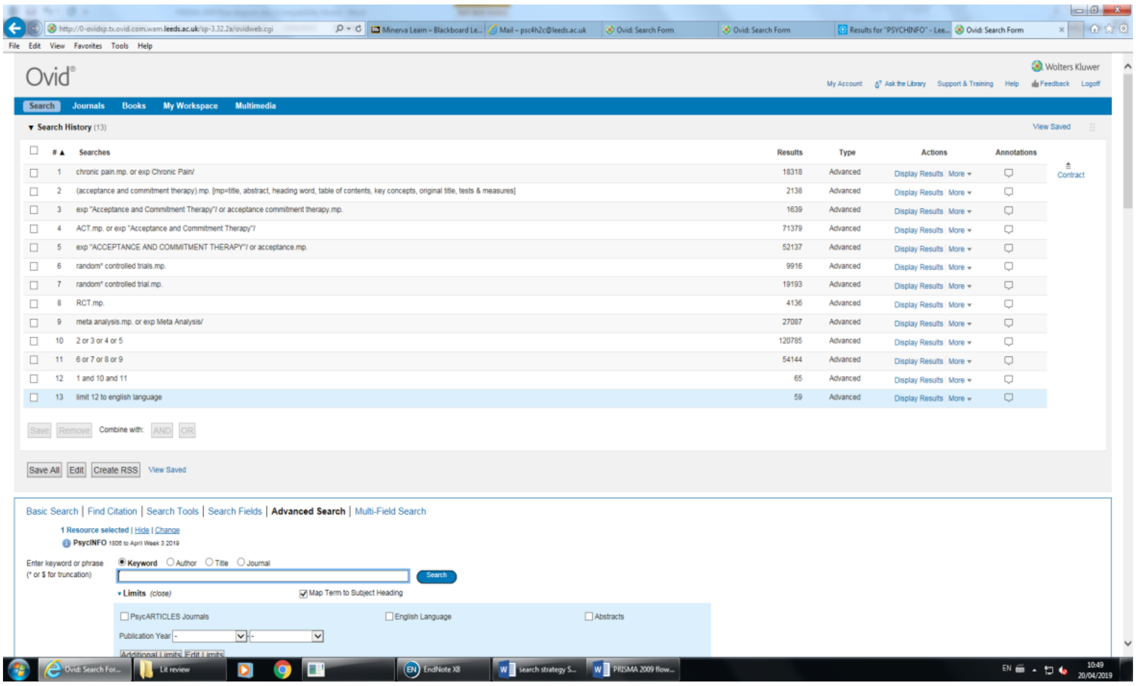
Enter keyword or phrase (* or $ for truncation): ● Keyword ○ Author ○ Title ○ Journal

Limits (close) ☑ Map Term to Subject Heading

☐ Abstracts ☐ English Language

Publication Year - - -

PschINFO- 20/04/19



▼ Search History (13)

| ☐ | # ▲ | Searches | Results | Type | Actions | Annotations |
|---|---|---|---|---|---|---|
| ☐ | 1 | chronic pain.mp. or exp Chronic Pain/ | 18318 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 2 | (acceptance and commitment therapy).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] | 2138 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 3 | exp "Acceptance and Commitment Therapy"/ or acceptance commitment therapy.mp. | 1639 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 4 | ACT.mp. or exp "Acceptance and Commitment Therapy"/ | 71379 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 5 | exp "ACCEPTANCE AND COMMITMENT THERAPY"/ or acceptance.mp. | 52137 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 6 | random* controlled trials.mp. | 9916 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 7 | random* controlled trial.mp. | 19193 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 8 | RCT.mp. | 4136 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 9 | meta analysis.mp. or exp Meta Analysis/ | 27087 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 10 | 2 or 3 or 4 or 5 | 120785 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 11 | 6 or 7 or 8 or 9 | 54144 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 12 | 1 and 10 and 11 | 65 | Advanced | Display Results More ▾ | 🗨 |
| ☐ | 13 | limit 12 to english language | 59 | Advanced | Display Results More ▾ | 🗨 |

1 Resource selected | Hide | Change

PsycINFO 1806 to April Week 3 2019

Enter keyword or phrase (* or $ for truncation) ● Keyword ○ Author ○ Title ○ Journal

▾ Limits (close) ☑ Map Term to Subject Heading

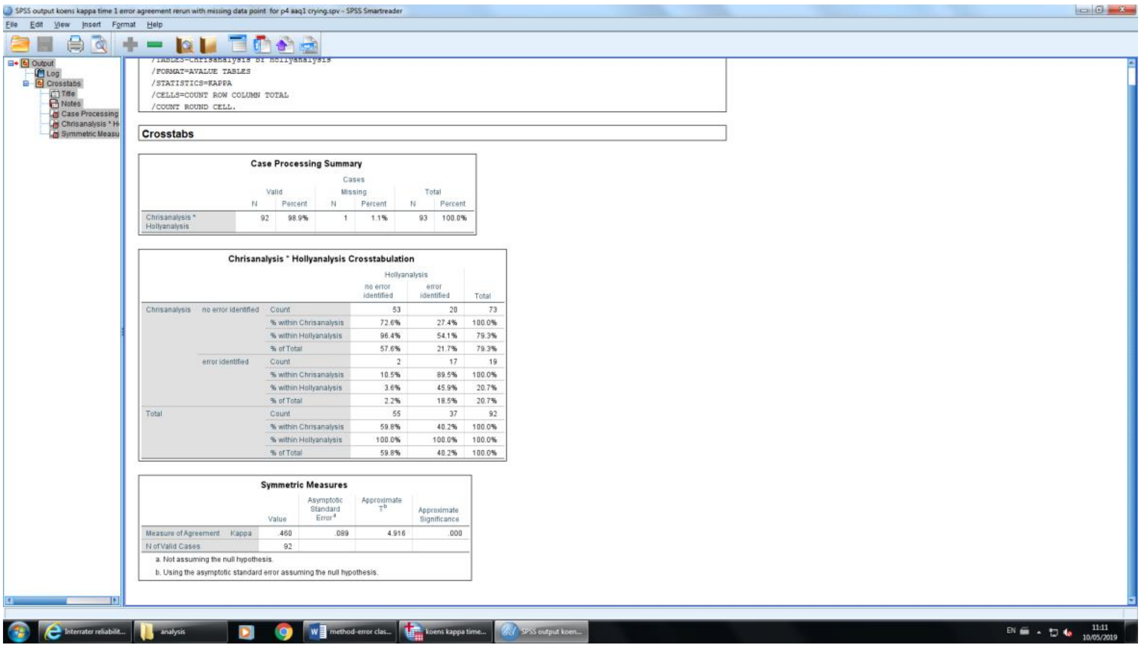
☐ PsycARTICLES Journals ☐ English Language ☐ Abstracts

Publication Year - -

# Appendix E. Interrater reliability analysis (SPSS and ReCal output screenshots)

## E.1 Time One analysis: Cohen’s kappa (κ)

1) *Cohen’s kappa* (κ) *for error agreement:*



```
  /TABLES=Chrisanalysis BY Hollyanalysis
  /FORMAT=AVALUE TABLES
  /STATISTICS=KAPPA
  /CELLS=COUNT ROW COLUMN TOTAL
  /COUNT ROUND CELL.
```

**Crosstabs**

**Case Processing Summary**

| | Cases | | | | | |
|---|---|---|---|---|---|---|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Chrisanalysis * Hollyanalysis | 92 | 98.9% | 1 | 1.1% | 93 | 100.0% |

**Chrisanalysis * Hollyanalysis Crosstabulation**

| | | | Hollyanalysis | | |
|---|---|---|---|---|---|
| | | | no error identified | error identified | Total |
| Chrisanalysis | no error identified | Count | 53 | 20 | 73 |
| | | % within Chrisanalysis | 72.6% | 27.4% | 100.0% |
| | | % within Hollyanalysis | 96.4% | 54.1% | 79.3% |
| | | % of Total | 57.6% | 21.7% | 79.3% |
| | error identified | Count | 2 | 17 | 19 |
| | | % within Chrisanalysis | 10.5% | 89.5% | 100.0% |
| | | % within Hollyanalysis | 3.6% | 45.9% | 20.7% |
| | | % of Total | 2.2% | 18.5% | 20.7% |
| Total | | Count | 55 | 37 | 92 |
| | | % within Chrisanalysis | 59.8% | 40.2% | 100.0% |
| | | % within Hollyanalysis | 100.0% | 100.0% | 100.0% |
| | | % of Total | 59.8% | 40.2% | 100.0% |

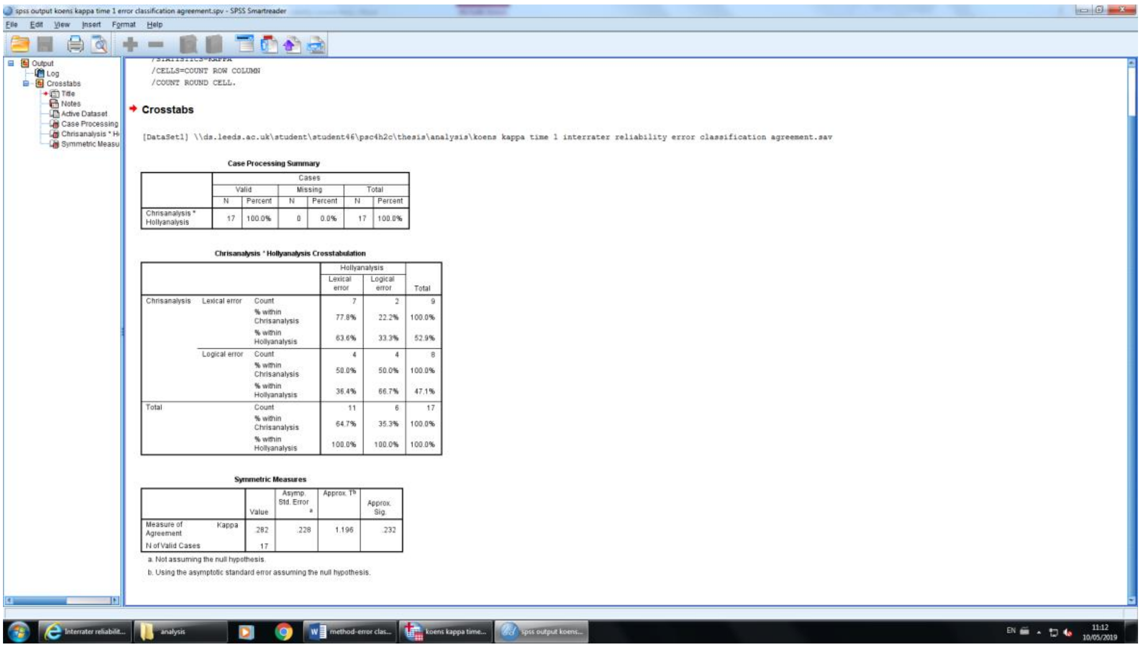
**Symmetric Measures**

| | | Value | Asymptotic Standard Error[a] | Approximate T[b] | Approximate Significance |
|---|---|---|---|---|---|
| Measure of Agreement | Kappa | .460 | .089 | 4.916 | .000 |
| N of Valid Cases | | 92 | | | |

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

2) *Cohen’s kappa* (κ) *for error classification:*



```
  /STATISTICS=KAPPA
  /CELLS=COUNT ROW COLUMN
  /COUNT ROUND CELL.
```

**Crosstabs**

```
[DataSet1] \\ds.leeds.ac.uk\student\student46\psc4h2c\thesis\analysis\koens kappa time 1 interrater reliability error classification agreement.sav
```

**Case Processing Summary**

| | Cases | | | | | |
|---|---|---|---|---|---|---|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Chrisanalysis * Hollyanalysis | 17 | 100.0% | 0 | 0.0% | 17 | 100.0% |

**Chrisanalysis * Hollyanalysis Crosstabulation**

| | | | Hollyanalysis | | |
|---|---|---|---|---|---|
| | | | Lexical error | Logical error | Total |
| Chrisanalysis | Lexical error | Count | 7 | 2 | 9 |
| | | % within Chrisanalysis | 77.8% | 22.2% | 100.0% |
| | | % within Hollyanalysis | 63.6% | 33.3% | 52.9% |
| | Logical error | Count | 4 | 4 | 8 |
| | | % within Chrisanalysis | 50.0% | 50.0% | 100.0% |
| | | % within Hollyanalysis | 36.4% | 66.7% | 47.1% |
| Total | | Count | 11 | 6 | 17 |
| | | % within Chrisanalysis | 64.7% | 35.3% | 100.0% |
| | | % within Hollyanalysis | 100.0% | 100.0% | 100.0% |

**Symmetric Measures**

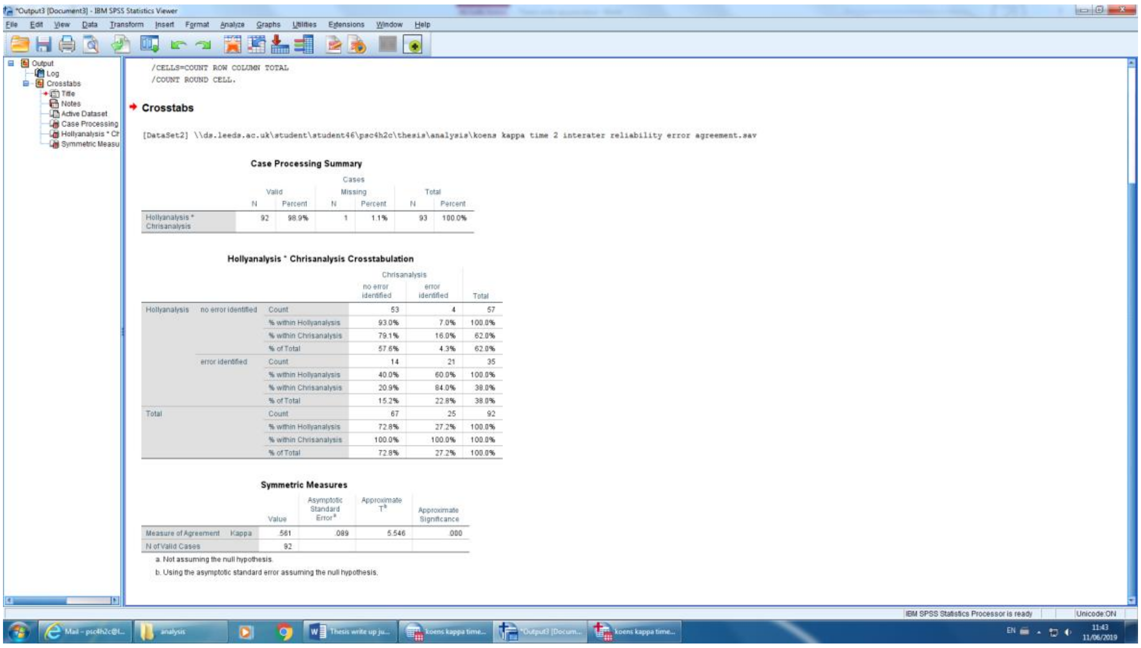
| | | Value | Asymp. Std. Error[a] | Approx. T[b] | Approx. Sig. |
|---|---|---|---|---|---|
| Measure of Agreement | Kappa | .282 | .228 | 1.196 | .232 |
| N of Valid Cases | | 17 | | | |

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

## E.2 Time Two analysis: Cohen’s kappa (κ)

1) *Cohen’s kappa (κ) for error agreement:*



```
/CELLS=COUNT ROW COLUMN TOTAL
/COUNT ROUND CELL.
```

**Crosstabs**

```
[DataSet2] \\ds.leeds.ac.uk\student\student46\psc4h2c\thesis\analysis\koens kappa time 2 interater reliability error agreement.sav
```

**Case Processing Summary**

| | Cases | | | | | |
|---|---|---|---|---|---|---|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Hollyanalysis * Chrisanalysis | 92 | 98.9% | 1 | 1.1% | 93 | 100.0% |

**Hollyanalysis * Chrisanalysis Crosstabulation**

| | | | Chrisanalysis | | |
|---|---|---|---|---|---|
| | | | no error identified | error identified | Total |
| Hollyanalysis | no error identified | Count | 53 | 4 | 57 |
| | | % within Hollyanalysis | 93.0% | 7.0% | 100.0% |
| | | % within Chrisanalysis | 79.1% | 16.0% | 62.0% |
| | | % of Total | 57.6% | 4.3% | 62.0% |
| | error identified | Count | 14 | 21 | 35 |
| | | % within Hollyanalysis | 40.0% | 60.0% | 100.0% |
| | | % within Chrisanalysis | 20.9% | 84.0% | 38.0% |
| | | % of Total | 15.2% | 22.8% | 38.0% |
| Total | | Count | 67 | 25 | 92 |
| | | % within Hollyanalysis | 72.8% | 27.2% | 100.0% |
| | | % within Chrisanalysis | 100.0% | 100.0% | 100.0% |
| | | % of Total | 72.8% | 27.2% | 100.0% |

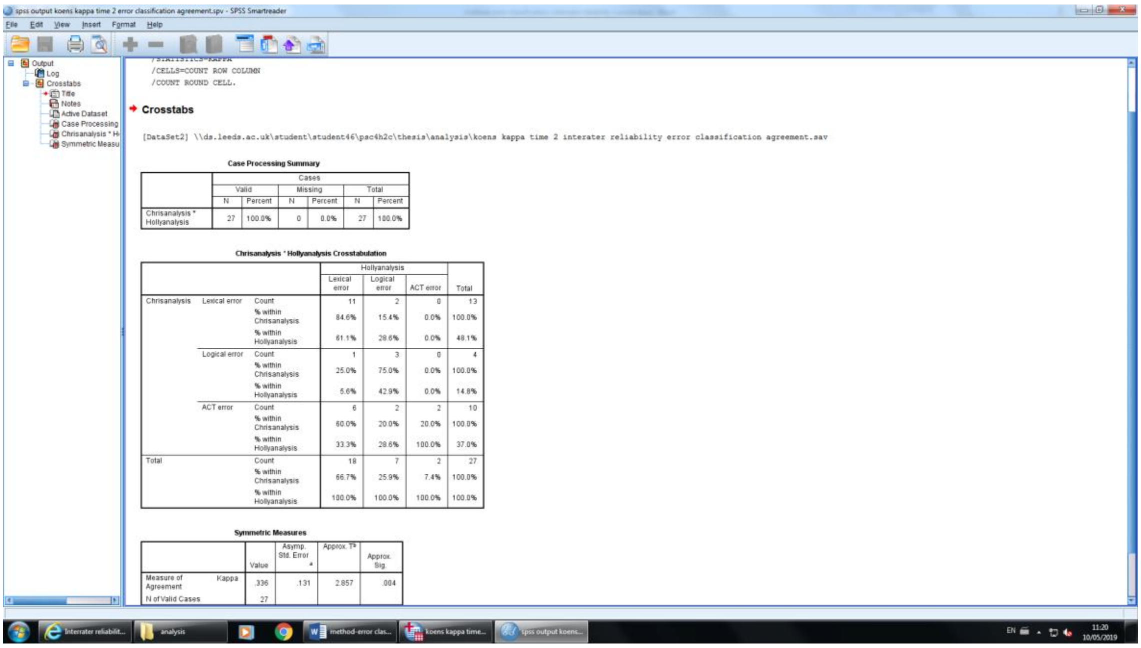
**Symmetric Measures**

| | | Value | Asymptotic Standard Error[a] | Approximate T[b] | Approximate Significance |
|---|---|---|---|---|---|
| Measure of Agreement | Kappa | .561 | .089 | 5.546 | .000 |
| N of Valid Cases | | 92 | | | |

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

2) *Cohen’s kappa* (κ) *for error classification:*



```
/CELLS=COUNT ROW COLUMN
/COUNT ROUND CELL.
```

**Crosstabs**

```
[DataSet2] \\ds.leeds.ac.uk\student\student46\psc4h2c\thesis\analysis\koens kappa time 2 interater reliability error classification agreement.sav
```

**Case Processing Summary**

| | Cases | | | | | |
|---|---|---|---|---|---|---|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Chrisanalysis * Hollyanalysis | 27 | 100.0% | 0 | 0.0% | 27 | 100.0% |

**Chrisanalysis * Hollyanalysis Crosstabulation**

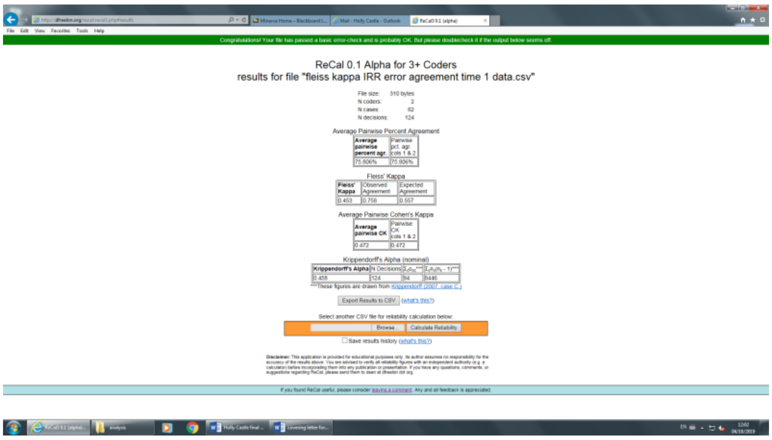
| | | | Hollyanalysis | | | |
|---|---|---|---|---|---|---|
| | | | Lexical error | Logical error | ACT error | Total |
| Chrisanalysis | Lexical error | Count | 11 | 2 | 0 | 13 |
| | | % within Chrisanalysis | 84.6% | 15.4% | 0.0% | 100.0% |
| | | % within Hollyanalysis | 61.1% | 28.6% | 0.0% | 48.1% |
| | Logical error | Count | 1 | 3 | 0 | 4 |
| | | % within Chrisanalysis | 25.0% | 75.0% | 0.0% | 100.0% |
| | | % within Hollyanalysis | 5.6% | 42.9% | 0.0% | 14.8% |
| | ACT error | Count | 6 | 2 | 2 | 10 |
| | | % within Chrisanalysis | 60.0% | 20.0% | 20.0% | 100.0% |
| | | % within Hollyanalysis | 33.3% | 28.6% | 100.0% | 37.0% |
| Total | | Count | 18 | 7 | 2 | 27 |
| | | % within Chrisanalysis | 66.7% | 25.9% | 7.4% | 100.0% |
| | | % within Hollyanalysis | 100.0% | 100.0% | 100.0% | 100.0% |

**Symmetric Measures**

| | | Value | Asymp. Std. Error[a] | Approx. T[b] | Approx. Sig. |
|---|---|---|---|---|---|
| Measure of Agreement | Kappa | .336 | .131 | 2.857 | .004 |
| N of Valid Cases | | 27 | | | |

## E.3 Time One analysis: Fleiss’ Kappa (κ)

*1) Error Agreement- (κ) moderate*



Congratulations! Your file has passed a basic error-check and is probably OK. But please doublecheck it if the output below seems off.

ReCal 0.1 Alpha for 3+ Coders
results for file "fleiss kappa IRR error agreement time 1 data.csv"

| | |
|---|---|
| File size: | 310 bytes |
| N coders: | 2 |
| N cases: | 62 |
| N decisions: | 124 |

Average Pairwise Percent Agreement

| **Average pairwise percent agr.** | Pairwise pct. agr. cols 1 & 2 |
|---|---|
| 75.806% | 75.806% |

Fleiss' Kappa

| **Fleiss' Kappa** | Observed Agreement | Expected Agreement |
|---|---|---|
| 0.453 | 0.758 | 0.557 |

Average Pairwise Cohen's Kappa

| **Average pairwise CK** | Pairwise CK cols 1 & 2 |
|---|---|
| 0.472 | 0.472 |

Krippendorff's Alpha (nominal)

| **Krippendorff's Alpha** | N Decisions | $\Sigma_c o_{cc}$*** | $\Sigma_c n_c(n_c - 1)$*** |
|---|---|---|---|
| 0.458 | 124 | 94 | 8446 |

***These figures are drawn from Krippendorff (2007, case C.)

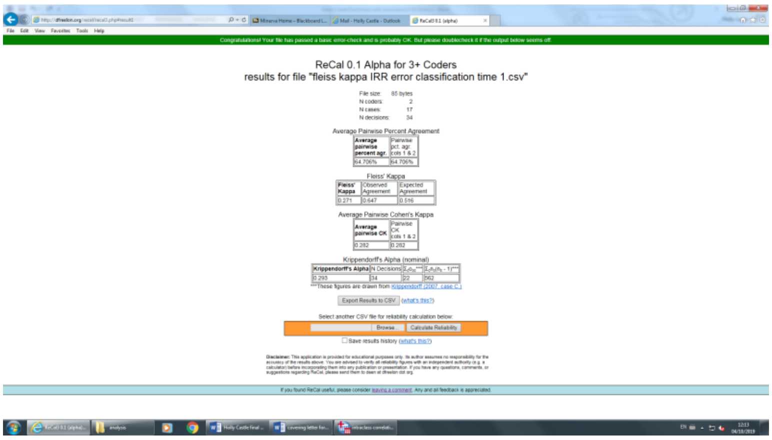
Select another CSV file for reliability calculation below:

☐ Save results history (what's this?)

**Disclaimer:** This application is provided for educational purposes only. Its author assumes no responsibility for the accuracy of the results above. You are advised to verify all reliability figures with an independent authority (e.g. a calculator) before incorporating them into any publication or presentation. If you have any questions, comments, or suggestions regarding ReCal, please send them to deen at dfreelon dot org.

If you found ReCal useful, please consider leaving a comment. Any and all feedback is appreciated.

*2) Error Classification- (κ) fair*



Congratulations! Your file has passed a basic error-check and is probably OK. But please doublecheck it if the output below seems off.

ReCal 0.1 Alpha for 3+ Coders
results for file "fleiss kappa IRR error classification time 1.csv"

| | |
|---|---|
| File size: | 85 bytes |
| N coders: | 2 |
| N cases: | 17 |
| N decisions: | 34 |

Average Pairwise Percent Agreement

| **Average pairwise percent agr.** | Pairwise pct. agr. cols 1 & 2 |
|---|---|
| 64.706% | 64.706% |

Fleiss' Kappa

| **Fleiss' Kappa** | Observed Agreement | Expected Agreement |
|---|---|---|
| 0.271 | 0.647 | 0.516 |

Average Pairwise Cohen's Kappa

| **Average pairwise CK** | Pairwise CK cols 1 & 2 |
|---|---|
| 0.282 | 0.282 |

Krippendorff's Alpha (nominal)

| **Krippendorff's Alpha** | N Decisions | $\Sigma_c o_{cc}$*** | $\Sigma_c n_c(n_c - 1)$*** |
|---|---|---|---|
| 0.293 | 34 | 22 | 562 |

***These figures are drawn from Krippendorff (2007, case C.)

Select another CSV file for reliability calculation below:

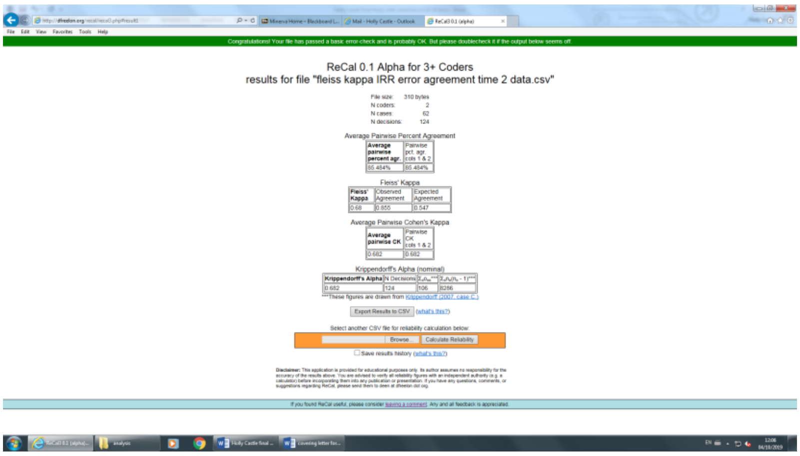
☐ Save results history (what's this?)

**Disclaimer:** This application is provided for educational purposes only. Its author assumes no responsibility for the accuracy of the results above. You are advised to verify all reliability figures with an independent authority (e.g. a calculator) before incorporating them into any publication or presentation. If you have any questions, comments, or suggestions regarding ReCal, please send them to deen at dfreelon dot org.

If you found ReCal useful, please consider leaving a comment. Any and all feedback is appreciated.

## E.4 Time Two analysis: Fleiss’ Kappa (κ)

*1) Error Agreement- (κ) substantial*



Congratulations! Your file has passed a basic error-check and is probably OK. But please doublecheck it if the output below seems off.

ReCal 0.1 Alpha for 3+ Coders
results for file "fleiss kappa IRR error agreement time 2 data.csv"

| | |
|---|---|
| File size: | 310 bytes |
| N coders: | 2 |
| N cases: | 62 |
| N decisions: | 124 |

Average Pairwise Percent Agreement

| **Average pairwise percent agr.** | Pairwise pct. agr. cols 1 & 2 |
|---|---|
| 85.484% | 85.484% |

Fleiss' Kappa

| **Fleiss' Kappa** | Observed Agreement | Expected Agreement |
|---|---|---|
| 0.68 | 0.855 | 0.547 |

Average Pairwise Cohen's Kappa

| **Average pairwise CK** | Pairwise CK cols 1 & 2 |
|---|---|
| 0.682 | 0.682 |

Krippendorff's Alpha (nominal)

| **Krippendorff's Alpha** | N Decisions | $\Sigma_c o_{cc}$*** | $\Sigma_c n_c(n_c - 1)$*** |
|---|---|---|---|
| 0.682 | 124 | 106 | 8286 |

***These figures are drawn from Krippendorff (2007, case C.)

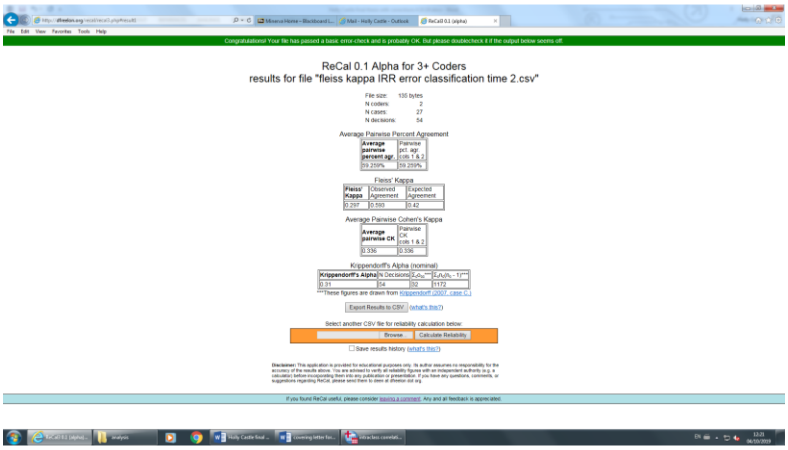
Select another CSV file for reliability calculation below:

☐ Save results history (what's this?)

**Disclaimer:** This application is provided for educational purposes only. Its author assumes no responsibility for the accuracy of the results above. You are advised to verify all reliability figures with an independent authority (e.g. a calculator) before incorporating them into any publication or presentation. If you have any questions, comments, or suggestions regarding ReCal, please send them to deen at dfreelon dot org.

If you found ReCal useful, please consider leaving a comment. Any and all feedback is appreciated.

*2)Error Classification- (κ) fair*



Congratulations! Your file has passed a basic error-check and is probably OK. But please doublecheck it if the output below seems off.

ReCal 0.1 Alpha for 3+ Coders
results for file "fleiss kappa IRR error classification time 2.csv"

| | |
|---|---|
| File size: | 135 bytes |
| N coders: | 2 |
| N cases: | 27 |
| N decisions: | 54 |

Average Pairwise Percent Agreement

| **Average pairwise percent agr.** | Pairwise pct. agr. cols 1 & 2 |
|---|---|
| 59.259% | 59.259% |

Fleiss' Kappa

| **Fleiss' Kappa** | Observed Agreement | Expected Agreement |
|---|---|---|
| 0.297 | 0.593 | 0.42 |

Average Pairwise Cohen's Kappa

| **Average pairwise CK** | Pairwise CK cols 1 & 2 |
|---|---|
| 0.336 | 0.336 |

Krippendorff's Alpha (nominal)

| **Krippendorff's Alpha** | N Decisions | $\Sigma_c o_{cc}$*** | $\Sigma_c n_c(n_c - 1)$*** |
|---|---|---|---|
| 0.31 | 54 | 32 | 1172 |

***These figures are drawn from Krippendorff (2007, case C.)

(what's this?)

Select another CSV file for reliability calculation below:

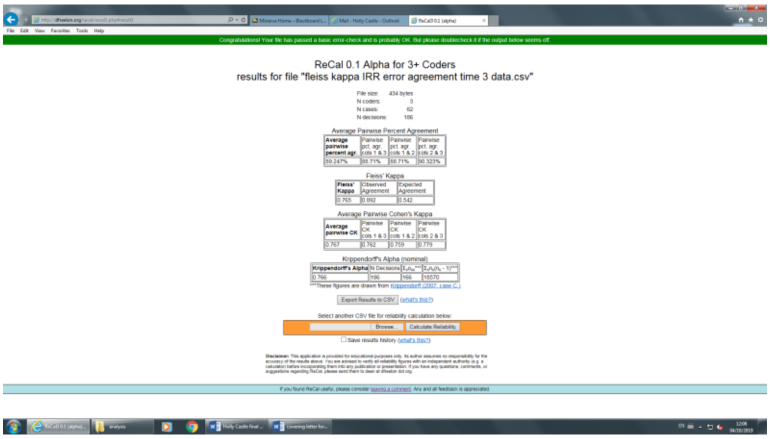
☐ Save results history (what's this?)

**Disclaimer:** This application is provided for educational purposes only. Its author assumes no responsibility for the accuracy of the results above. You are advised to verify all reliability figures with an independent authority (e.g. a calculator) before incorporating them into any publication or presentation. If you have any questions, comments, or suggestions regarding ReCal, please send them to deen at dfreelon dot org.

If you found ReCal useful, please consider leaving a comment. Any and all feedback is appreciated.

## E.5 Time Three analysis: Fleiss' Kappa (κ)

*1) Error Agreement- (κ) substantial*



Congratulations! Your file has passed a basic error-check and is probably OK. But please doublecheck it if the output below seems off.

ReCal 0.1 Alpha for 3+ Coders
results for file "fleiss kappa IRR error agreement time 3 data.csv"

| | |
|---|---|
| File size: | 434 bytes |
| N coders: | 3 |
| N cases: | 62 |
| N decisions: | 186 |

Average Pairwise Percent Agreement

| **Average pairwise percent agr.** | Pairwise pct. agr. cols 1 & 3 | Pairwise pct. agr. cols 1 & 2 | Pairwise pct. agr. cols 2 & 3 |
|---|---|---|---|
| 89.247% | 88.71% | 88.71% | 90.323% |

Fleiss' Kappa

| **Fleiss' Kappa** | Observed Agreement | Expected Agreement |
|---|---|---|
| 0.765 | 0.892 | 0.542 |

Average Pairwise Cohen's Kappa

| **Average pairwise CK** | Pairwise CK cols 1 & 3 | Pairwise CK cols 1 & 2 | Pairwise CK cols 2 & 3 |
|---|---|---|---|
| 0.767 | 0.762 | 0.759 | 0.779 |

Krippendorff's Alpha (nominal)

| **Krippendorff's Alpha** | N Decisions | $\Sigma_c o_{cc}$*** | $\Sigma_c n_c(n_c - 1)$*** |
|---|---|---|---|
| 0.766 | 186 | 166 | 18570 |

***These figures are drawn from Krippendorff (2007, case C.)

(what's this?)

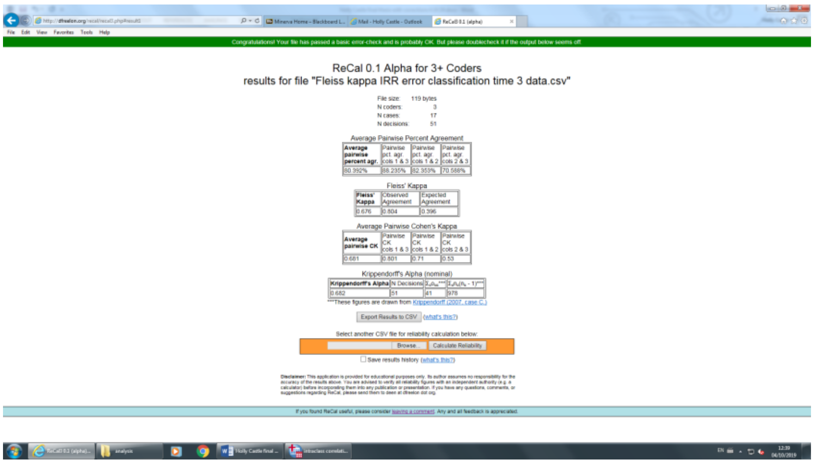
Select another CSV file for reliability calculation below:

☐ Save results history (what's this?)

**Disclaimer:** This application is provided for educational purposes only. Its author assumes no responsibility for the accuracy of the results above. You are advised to verify all reliability figures with an independent authority (e.g. a calculator) before incorporating them into any publication or presentation. If you have any questions, comments, or suggestions regarding ReCal, please send them to deen at dfreelon dot org.

If you found ReCal useful, please consider leaving a comment. Any and all feedback is appreciated.

*2)Error Classification- (κ) substantial*



Congratulations! Your file has passed a basic error-check and is probably OK. But please doublecheck it if the output below seems off.

ReCal 0.1 Alpha for 3+ Coders
results for file "Fleiss kappa IRR error classification time 3 data.csv"

| | |
|---|---|
| File size: | 119 bytes |
| N coders: | 3 |
| N cases: | 17 |
| N decisions: | 51 |

Average Pairwise Percent Agreement

| **Average pairwise percent agr.** | Pairwise pct. agr. cols 1 & 3 | Pairwise pct. agr. cols 1 & 2 | Pairwise pct. agr. cols 2 & 3 |
|---|---|---|---|
| 80.392% | 88.235% | 82.353% | 70.588% |

Fleiss' Kappa

| **Fleiss' Kappa** | Observed Agreement | Expected Agreement |
|---|---|---|
| 0.676 | 0.804 | 0.396 |

Average Pairwise Cohen's Kappa

| **Average pairwise CK** | Pairwise CK cols 1 & 3 | Pairwise CK cols 1 & 2 | Pairwise CK cols 2 & 3 |
|---|---|---|---|
| 0.681 | 0.801 | 0.71 | 0.53 |

Krippendorff's Alpha (nominal)

| **Krippendorff's Alpha** | N Decisions | $\Sigma_c o_{cc}$*** | $\Sigma_c n_c(n_c - 1)$*** |
|---|---|---|---|
| 0.682 | 51 | 41 | 978 |

***These figures are drawn from Krippendorff (2007, case C.)

(what's this?)

Select another CSV file for reliability calculation below:

☐ Save results history (what's this?)

**Disclaimer:** This application is provided for educational purposes only. Its author assumes no responsibility for the accuracy of the results above. You are advised to verify all reliability figures with an independent authority (e.g. a calculator) before incorporating them into any publication or presentation. If you have any questions, comments, or suggestions regarding ReCal, please send them to deen at dfreelon dot org.

If you found ReCal useful, please consider leaving a comment. Any and all feedback is appreciated.

# Appendix F. Raw data tables (for error classification analysis)

## F.1 Final multiple rater, interrater reliability check with reflections

*For reference purposes please note ‘conceptually inconsistent’ errors were originally named ‘ACT’ error

***Participant 1***

| Item | **Lead researcher** | **Supervisor A** | **Supervisor B** |
|---|---|---|---|
| CPAQ 1 | ✓ | ✓ | ✓ |
| CPAQ 2 | ✓ | ✓ | ✓ |
| CPAQ 3 | Logical problems<br>First part of question leads P to only talk about impact of pain on living life. Not able to answer whether she is living a normal life despite pain. | LOGICAL – doesn’t include ‘although things have changed’ | Logical; a focus on what has changed; may be item precision – leads to focus on first part |
| CPAQ 4 | Issues with question item precision<br>Misunderstands ACT concept of pain willingness, views question as implying she should have control over pain and explains that she is good at this through managing her medication. Gives a low answer but this is in inaccurate reflection of pain willingness as she already feels ‘in control’?! | ACT error – doesn’t get at acceptance/willingness as intended…as person has control over pain. | logical – focus on second part “I have to get control over my pain” and misses link to second part; doesn’t get the link to acceptance: interprets control as a good thing |
| CPAQ 5 | ✓ | ✓ | ✓ |
| CPAQ 6 | ✓ | ✓ | ✓ |
| CPAQ 7 | ✓ | ✓ | ✓ |
| CPAQ 8 | Logical error | Logical – doesn’t consider whether thoughts are ‘true’ or not within their answer. More just about how often they worry. | Logic: responding to the word ‘true’ as referring to worrying – i.e. its true that I worry for my family |
| PIPS 1 | Logical problems<br>I think the P is confused by the question, tries to adapt the wording to make more sense to her, therefore answers differently to the way the question originally intended. | ✓ | Logic: confound getting rid of pain with achieving things rather than doing nothing |
| PIPS 2 | ✓ | ✓ | ✓ |
| PIPS 3 | Issues with question item precision:<br>Doesn’t answer the question, talks about not having a ‘choice’ other people coming to her, scores low on this but actually her answer suggests she would stay away from people when in pain if she had the option. | ACT error – because assumes that person has choice to stay away…doesn’t get at choice | Focus on lack of choice rather than what they would do if they had a choice; ? item precision? Add “where you have a choice…” to beginning? |
| PIPS 4 | ✓ | ✓ | ✓ |
| PIPS 5 | ✓ | ✓ | ✓ |
| PIPS 6 | ✓ | ✓ | ✓ |

| | | | |
|---|---|---|---|
| PIPS 7 | *Lexical problems<br>Use of the word ‘say’ has caused confusion, P has interpreted as saying these things out loud to her family etc. still not clearer whether these are the kinds of statements she might say to herself? Issues with the context of the question-‘say’ means different things to different people. Score of never true does not reflect her actual inner feelings. | Lexical error – say…<br>Maybe ACT error …. Not tapping into EA/lack of committed action as intended… | Item precision? Confusion between saying out loud and thinking/doing |
| PIPS 8 | ✓ | ✓ | ✓ |
| PIPS 9 | ✓ | ✓ | ✓ |
| PIPS 10 | *Lexical problems<br>The word ‘fighting’ is not seen as fitting with the P’s answer, so she changes response into ‘ignoring pain/ or coping with pain’ | Lexical – re-casts as coping<br>ACT error … not ‘fighting’ but coping | Lexical problem: Interprets fighting pain as carrying on and ignore it |
| PIPS 11 | ✓ | ✓ | ✓ |
| PIPS 12 | ✓ | ✓ | ✓ |
| PIPS 13 | ✓ | ✓ | ✓ |
| PIPS 14 | ✓ | ✓ | ✓ |
| PIPS 15 | ✓ | ✓ | ✓ |
| PIPS 16 | ✓ | ✓ | ✓ |
| AAQ1 | ✓ | ✓ | ✓ |
| AAQ2 | ✓ | ✓ | ✓ |
| AAQ3 | *Logical problems<br>P doesn’t refer to ‘worrying about worrying’ and therefore does not answer the question in the way intended. She refers to not wanting her family to worry about her. | Logical – extent of worry, does not include idea of CONTROL<br>Maybe ACT error – extent of worry, not control of worry | Logical: focus on others’ worries |
| AAQ4 | ✓ | ✓ | ✓ |
| AAQ5 | *Logical problems<br>P scores herself low when the answer actually does indicate she does feel negative emotions cause issues- i.e. her score gives the impression of psychological flexibility when her in-depth answer reveals the opposite | Possibly an ACT Error (cover up emotions is good…reflected in score!) | Logic problem: Avoidance revealed in answer |
| AAQ6 | ✓ | ✓ | ✓ |
| AAQ7 | Lexical problems<br>The concept of ‘success’ is subjective, and P is not sure whether her answer covers this. | Lexical – not sure about def of success | Well she acknowledges worrying gets in the way of success in her relationship… maybe this is answered accurately? Success is subjective – but this is what it means for her? |

**Following Time 3 interrater reliability analysis, discrepancies between the raters were discussed. PIPS7, PIPS10, AAQ3, AAQ5 were then additionally classified within ACT error category and this was reflected in the final analysis.*

*Participant 3*

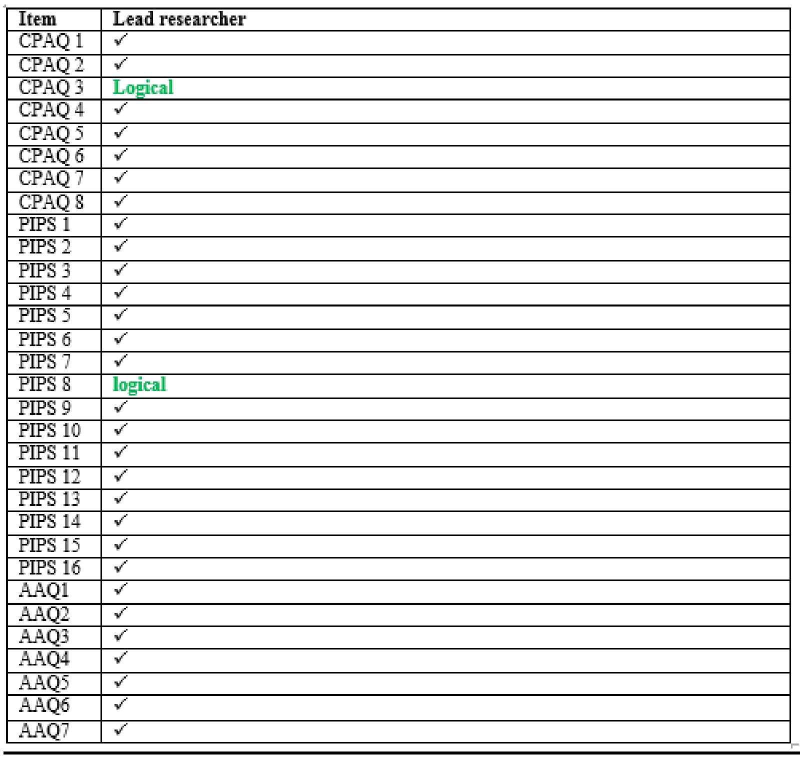
| | **Lead researcher** | **Supervisor A** | **Supervisor B** |
|---|---|---|---|
| CPAQ 1 | Lexical problems<br>P has understood the meaning of the word 'living' in a different way to that intended by the questionnaire- measure of activity engagement/living according to your values. The P refers to living in terms of his mortality- like a risk assessment questionnaire. | Lexical – by living means suicidality… | Lexical – interprets this as suicidality |
| CPAQ 2 | ✓ | ✓ | ✓ |
| CPAQ 3 | ✓ | ✓ | ✓ |
| CPAQ 4 | ✓ | ✓ | ✓ |
| CPAQ 5 | Lexical problems<br>P unclear what is meant by a 'full life'. Scores highly however is unclear whether this truly reflects what their level of activity engagement. | ✓ | Lexical problems<br>Answer indicates they have made multiple changes in life because of pain and adapted to it; unclear if it is 'full' – acknowledges that t recognises that full life depends on definition |
| CPAQ 6 | ✓ | ✓ | ✓ |
| CPAQ 7 | ✓ | ✓ | ✓ |
| CPAQ 8 | Logical problems<br>I'm not sure if he's answering the question. | Logical – don't include the concept of the 'truth' of thoughts | Logical – interpreted as worry about pain in short term |
| PIPS 1 | *Lexical problems<br>P substitutes getting 'rid' of pain for 'controlling' it and their answer reflects this. | ✓ | ✓<br>He does talk about control rather than getting rid, but I think it' still true to the spirit of the question |
| PIPS 2 | *Logical problems<br>Their answer does not reflect avoidance, scores highly for avoiding however then talks about engaging in activities but having to meticulously plan these. | Lexical – recasts avoiding<br>ACT error – doesn't tap into avoidance …. Taps into planning so avoidance doesn't happen? | Logical problem: he does do things (but plans them) |
| PIPS 3 | ✓ | ✓ | ✓ |
| PIPS 4 | ✓ | ✓ | ✓ |
| PIPS 5 | ✓ | ✓ | ✓ |
| PIPS 6 | ✓ | ✓ | ✓ |
| PIPS 7 | item precision<br>Complicated structure, leads P to go for an in-between score he agrees with half of the list but not the other half. | ✓ | item precision<br>I agree with Holly, he is choosing phrases to agree with |
| PIPS 8 | ✓ | ✓ | ✓ |
| PIPS 9 | *Lexical problems<br>The word scheduling is misleading; question asks about engaging in activities/avoiding activities, P refers to meticulously planning activities out. | Lexical – scheduling versus planning<br>ACT error – Scheduling happens for the function of avoiding pain…doesn't avoid scheduling/does avoid pain by scheduling | Lexical – scheduling used interchangeably with planning |

| | | | |
|---|---|---|---|
| PIPS 10 | *Lexical problems<br>Focus not on ‘fighting’- on something else | Lexical – recasts fighting<br>ACT error – doestnt tap into quality of struggling …. | Lexical – confounds pain with causes of pain, and refers to planning again |
| PIPS 11 | ✓ | ✓ | ✓ |
| PIPS 12 | ✓ | ✓ | Lexical<br>I think there is confusion here about ‘what is wrong’; he seems to refer to things that make pain worse, not cause of pain |
| PIPS 13 | *Lexical problems<br>Current planning referred to, not long term future | ACT error – plans serve function of avoiding pain …consequently s/he plans | Lexical<br>Future taken as short term, not long term |
| PIPS 14 | Lexical problems<br>P changes wording of question from postponing to talking about ‘planning’ | Lexical – planning and not postponing (therefore answer is a 6…!) | Lexical<br>Planning not postponing |
| PIPS 15 | ACT error | ACT error – maybe – so good at EA (via planning and pills) that doesn’t cancel. | ✓<br>Think this is OK |
| PIPS 16 | ✓ | ✓ | ✓ |
| AAQ1 | Lexical problems<br>Not personal values but instead the idea that her life has less worth | Lexical – value meaning whether their life has value | ✓<br>Hmm, is this one OK?; he is acknowledging the difficulties in living a life with value with pain, but identifying his children as giving it value? |
| AAQ2 | *Lexical problems<br>Context of the question creating lexical problems- ‘feelings’ unclear to them, is it physical sensations or emotions? | ✓ | ✓<br>Think this is OK |
| AAQ3 | ✓ | ✓ | ✓ |
| AAQ4 | *item precision | ✓ | ✓ |
| AAQ5 | ✓ | ✓ | ✓ |
| AAQ6 | *Logical problems | ✓ | ✓ |
| AAQ7 | Lexical problems<br>The concept of ‘success’ is subjective, and P is not sure whether her answer covers this. | Lexical – don’t know what success is | Lexical<br>Definition of success |

**Following Time 3 interrater reliability analysis discrepancies between the raters were discussed. The lead researcher then classified items; AAQ2, AAQ4 and AAQ6 within ‘no error’ category, PIPS1 was placed within the ‘no error’ category and PIPS 2, 9, 10 and 13 were added to the ACT error category in addition to their current category. Again, this was reflected in the final analysis.*

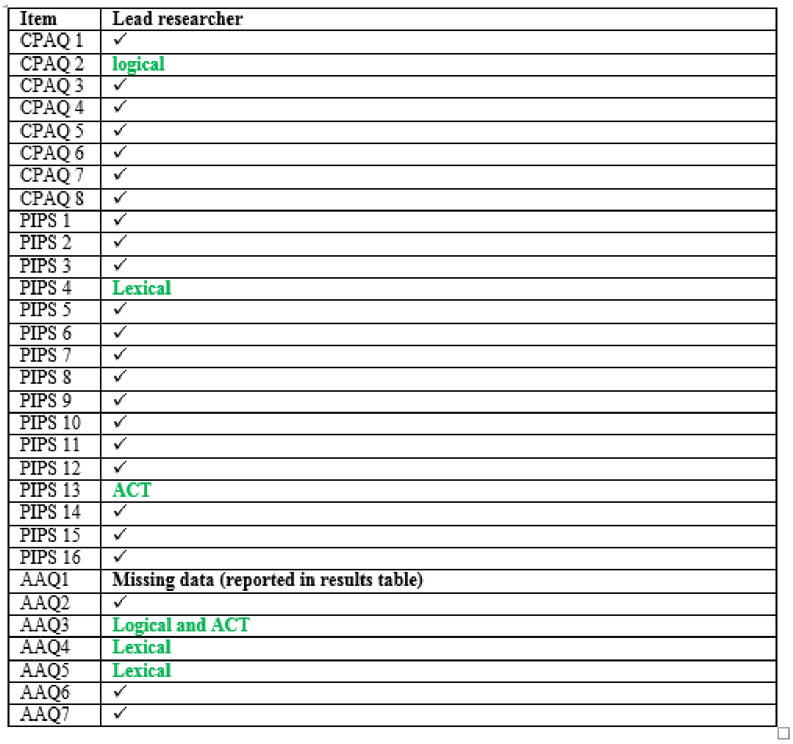
## F.2 Main analysis raw data for error classification (remaining participants)

*Participant 2*



| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | Logical |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | logical |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

*Participant 4*



| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | logical |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | Lexical |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ACT |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | Missing data (reported in results table) |
| AAQ2 | ✓ |
| AAQ3 | Logical and ACT |
| AAQ4 | Lexical |
| AAQ5 | Lexical |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

*Participant 5*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | Logical and ACT |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | logical |
| PIPS 1 | ✓ |
| PIPS 2 | Response error |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ACT |
| PIPS 11 | ✓ |
| PIPS 12 | Lexical |
| PIPS 13 | ACT |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | Lexical and logical |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | Lexical and logical |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

*Participant 6*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | Logical |
| PIPS 1 | Logical and ACT |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | Lexical |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | lexical |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ACT |
| AAQ4 | ✓ |
| AAQ5 | Logical and ACT |
| AAQ6 | ✓ |
| AAQ7 | Logical and ACT |

*Participant 7*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | Logical and ACT |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | Lexical and ACT |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ACT |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | lexical |
| AAQ2 | ✓ |
| AAQ3 | logical |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

*Participant 8*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | Logical and ACT |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | logical |
| PIPS 5 | ✓ |
| PIPS 6 | lexical |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | Lexical |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

### *Participant 9*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | logical |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | logical |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | Logical |
| PIPS 6 | ✓ |
| PIPS 7 | Logical and ACT |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ACT |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | Logical and ACT |
| AAQ4 | logical |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

### *Participant 10*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | logical |
| CPAQ 4 | Lexical, logical and ACT |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | Lexical |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | lexical |
| PIPS 10 | Lexical and ACT |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ good no error example |
| AAQ2 | Lexical |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | Logical |

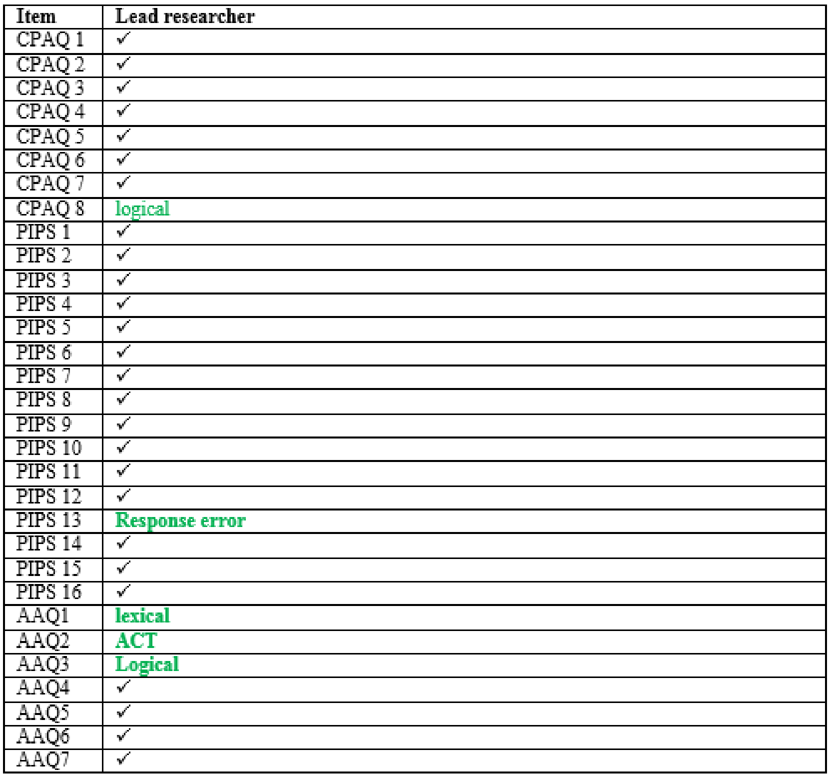
***Participant 11***

| **Item** | **Lead researcher** |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | **Logical** |
| CPAQ 5 | ✓ |
| CPAQ 6 | **Lexical and ACT** |
| CPAQ 7 | ✓ |
| CPAQ 8 | **logical** |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | **Logical and ACT** |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | **logical** |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | **logical** |
| AAQ3 | **Logical and ACT** |
| AAQ4 | ✓ |
| AAQ5 | **Lexical and logical and ACT** |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

***Participant 12***

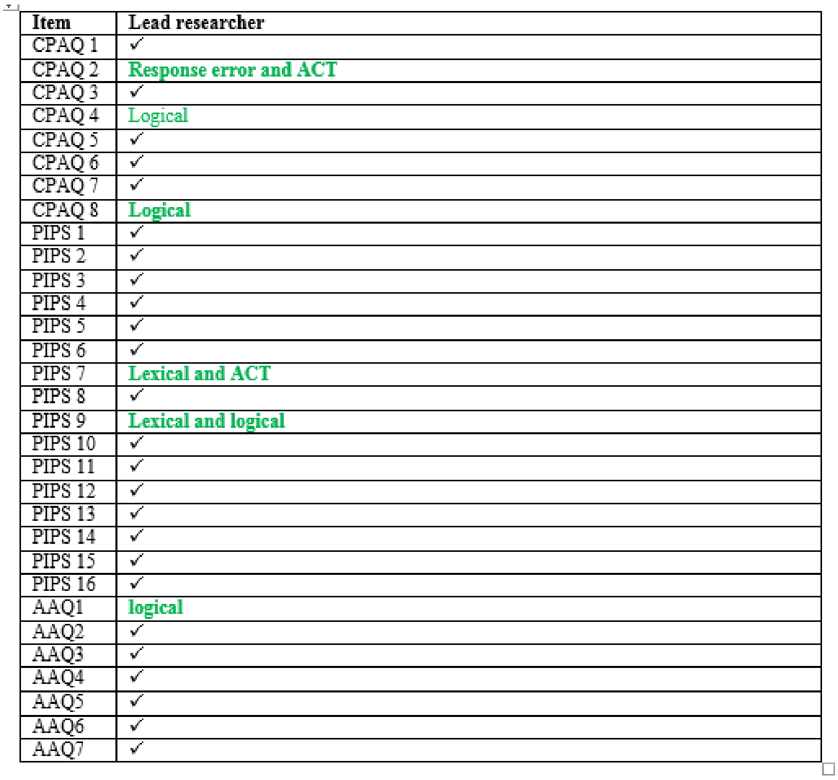
| **Item** | **Lead researcher** |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | **Lexical, logical and ACT** |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | **lexical** |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | **Logical and ACT** |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | **Logical lexical and ACT** |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

### *Participant 13*



| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | logical |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | **Response error** |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | **lexical** |
| AAQ2 | **ACT** |
| AAQ3 | **Logical** |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

### *Participant 14*



| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | **Response error and ACT** |
| CPAQ 3 | ✓ |
| CPAQ 4 | Logical |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | **Logical** |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | **Lexical and ACT** |
| PIPS 8 | ✓ |
| PIPS 9 | **Lexical and logical** |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | **logical** |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

### *Participant 15*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | logical and ACT error |
| CPAQ 3 | ACT |
| CPAQ 4 | Lexical, logical and ACT |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | logical |
| PIPS 1 | ACT |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ACT |
| PIPS 10 | Lexical and ACT |
| PIPS 11 | Response error |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | Logical / ACT error |

### *Participant 16*

| Item | Lead researcher |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | Logical and ACT |
| CPAQ 5 | Lexical |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | logical |
| PIPS 1 | ✓ |
| PIPS 2 | Response error and ACT |
| PIPS 3 | ✓ |
| PIPS 4 | Logical and ACT |
| PIPS 5 | ✓ |
| PIPS 6 | logical |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | Lexical, logical and ACT |
| PIPS 11 | ✓ |
| PIPS 12 | Lexical and logical |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | Logical |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | Lexical, logical |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | Logical and ACT |
| AAQ6 | Lexical and logical |
| AAQ7 | ✓ |

### *Participant 17*

| **Item** | **Lead researcher** |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | **logical** |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | Response error |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

---

### *Participant 18*

| **Item** | **Lead researcher** |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | **Logical and ACT** |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | **Logical and ACT** |
| PIPS 5 | ✓ |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | **Lexical** |
| PIPS 16 | ✓ |
| AAQ1 | **Logical and ACT** |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | ✓ |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

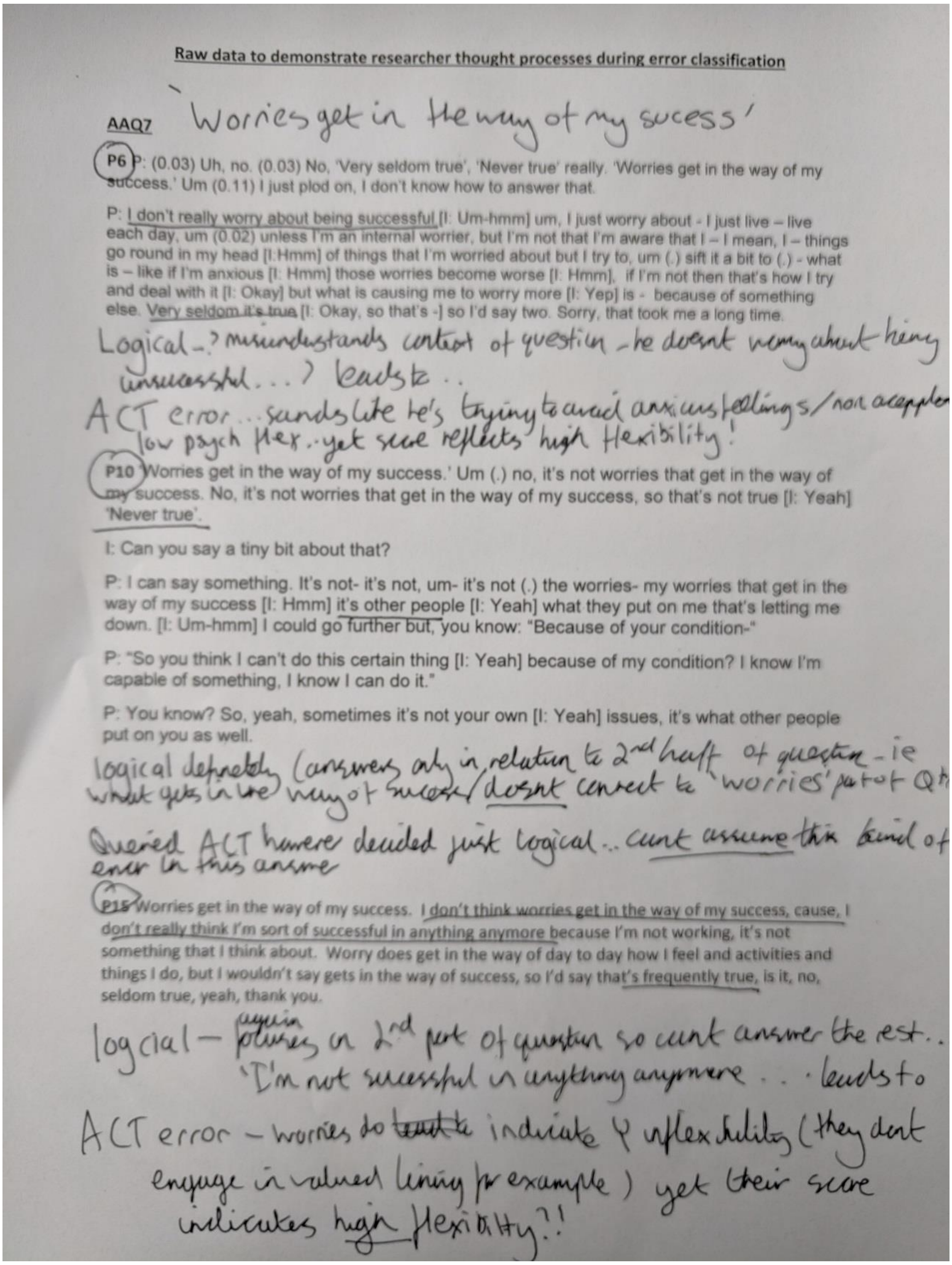
***Participant 19***

| **Item** | **Lead researcher** |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | ✓ |
| PIPS 1 | ✓ |
| PIPS 2 | **Response error and ACT** |
| PIPS 3 | ✓ |
| PIPS 4 | **Logical and ACT** |
| PIPS 5 | ✓ |
| PIPS 6 | **Lexical** |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | **Lexical** |
| PIPS 11 | ✓ |
| PIPS 12 | ✓ |
| PIPS 13 | ✓ |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | logical |
| AAQ4 | ✓ |
| AAQ5 | ACT |
| AAQ6 | Logical and ACT error |
| AAQ7 | ✓ |

***Participant 20***

| **Item** | **Lead researcher** |
|---|---|
| CPAQ 1 | ✓ |
| CPAQ 2 | ✓ |
| CPAQ 3 | ✓ |
| CPAQ 4 | ✓ |
| CPAQ 5 | ✓ |
| CPAQ 6 | ✓ |
| CPAQ 7 | ✓ |
| CPAQ 8 | **logical** |
| PIPS 1 | ✓ |
| PIPS 2 | ✓ |
| PIPS 3 | ✓ |
| PIPS 4 | ✓ |
| PIPS 5 | **Lexical, logical and ACT** |
| PIPS 6 | ✓ |
| PIPS 7 | ✓ |
| PIPS 8 | ✓ |
| PIPS 9 | ✓ |
| PIPS 10 | ✓ |
| PIPS 11 | ✓ |
| PIPS 12 | **Logical** |
| PIPS 13 | Response error |
| PIPS 14 | ✓ |
| PIPS 15 | ✓ |
| PIPS 16 | ✓ |
| AAQ1 | ✓ |
| AAQ2 | ✓ |
| AAQ3 | ✓ |
| AAQ4 | **logical** |
| AAQ5 | ✓ |
| AAQ6 | ✓ |
| AAQ7 | ✓ |

### F.3 Sample raw data to demonstrate researcher thought processes during error classification: (for AAQ-II item 7, 6 and 5)



Raw data to demonstrate researcher thought processes during error classification

AAQ7 'Worries get in the way of my sucess'

P6 P: (0.03) Uh, no. (0.03) No, 'Very seldom true', 'Never true' really. 'Worries get in the way of my success.' Um (0.11) I just plod on, I don't know how to answer that.

P: I don't really worry about being successful [I: Um-hmm] um, I just worry about - I just live – live each day, um (0.02) unless I'm an internal worrier, but I'm not that I'm aware that I – I mean, I – things go round in my head [I:Hmm] of things that I'm worried about but I try to, um (.) sift it a bit to (.) - what is – like if I'm anxious [I: Hmm] those worries become worse [I: Hmm], if I'm not then that's how I try and deal with it [I: Okay] but what is causing me to worry more [I: Yep] is - because of something else. Very seldom it's true [I: Okay, so that's -] so I'd say two. Sorry, that took me a long time.

Logical - ? misunderstands context of question - he doesn't worry about being unsuccessful... ? leads to..

ACT error... sounds like he's trying to avoid anxious feelings/non acceptan low psych flex.. yet score reflects high flexibility!

P10 'Worries get in the way of my success.' Um (.) no, it's not worries that get in the way of my success. No, it's not worries that get in the way of my success, so that's not true [I: Yeah] 'Never true'.

I: Can you say a tiny bit about that?

P: I can say something. It's not- it's not, um- it's not (.) the worries- my worries that get in the way of my success [I: Hmm] it's other people [I: Yeah] what they put on me that's letting me down. [I: Um-hmm] I could go further but, you know: "Because of your condition-"

P: "So you think I can't do this certain thing [I: Yeah] because of my condition? I know I'm capable of something, I know I can do it."

P: You know? So, yeah, sometimes it's not your own [I: Yeah] issues, it's what other people put on you as well.

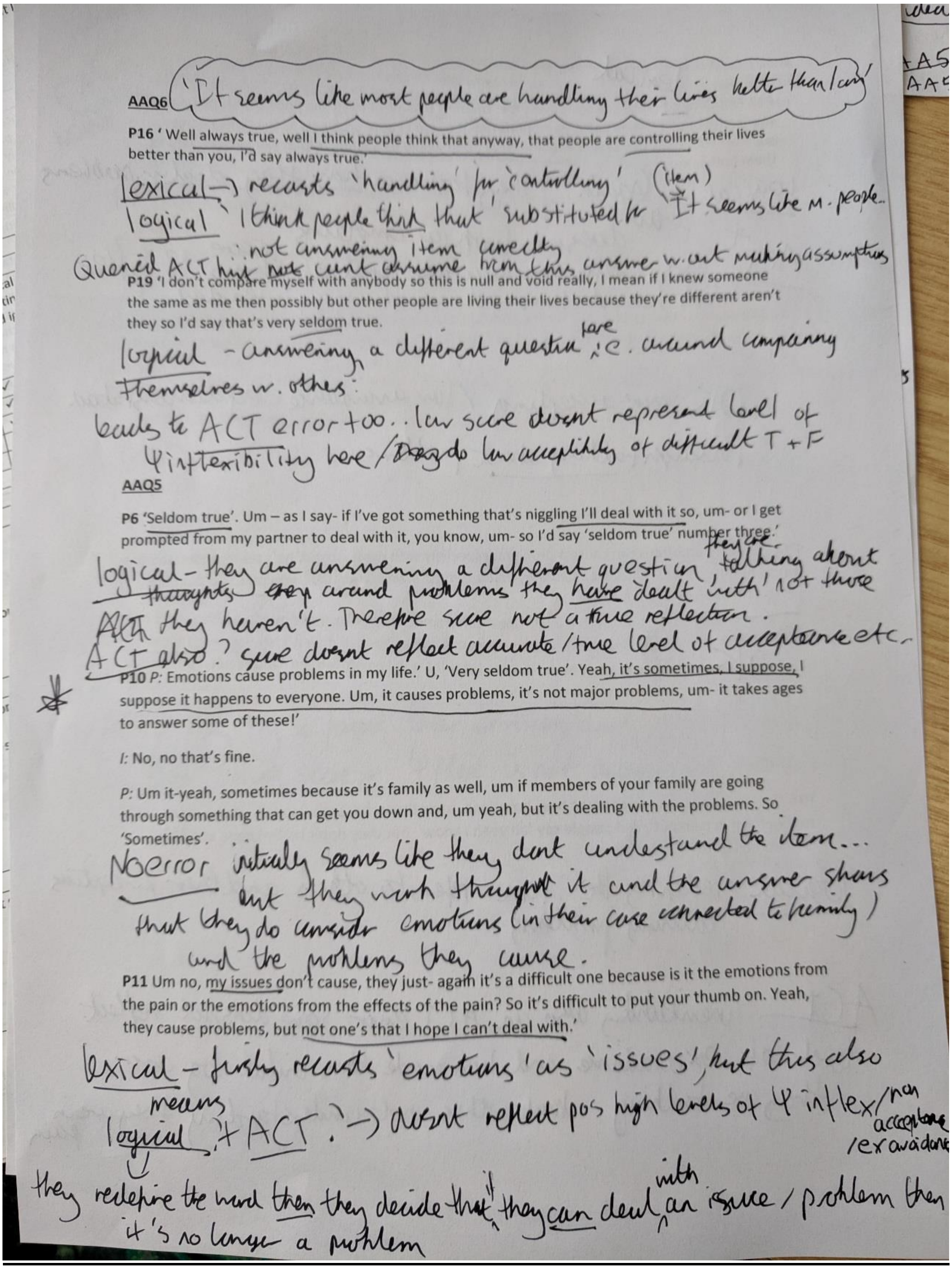
logical definitely (answers only in relation to 2nd half of question - ie what gets in the way of success/doesnt connect to 'worries' part of Q

Queried ACT however decided just logical... cant assume this kind of error in this answer

P15 Worries get in the way of my success. I don't think worries get in the way of my success, cause, I don't really think I'm sort of successful in anything anymore because I'm not working, it's not something that I think about. Worry does get in the way of day to day how I feel and activities and things I do, but I wouldn't say gets in the way of success, so I'd say that's frequently true, is it, no, seldom true, yeah, thank you.

logcial - focuses again on 2nd part of question so cant answer the rest.. 'I'm not successful in anything anymore ... leads to

ACT error - worries do ~~tend to~~ indicate ↑ inflexibility (they dont engage in valued living for example) yet their score indicates high flexibility?!



**AAQ6** 'It seems like most people are handling their lives better than I can'

**P16** ' Well always true, well I think people think that anyway, that people are controlling their lives better than you, I'd say always true.'

lexical → recasts 'handling' for 'controlling' (item)

logical 'I think people think that' substituted for 'It seems like m. people..

∴ not answering item correctly

Queried ACT ~~but not~~ cant assume from this answer w.out making assumptions

**P19** 'I don't compare myself with anybody so this is null and void really, I mean if I knew someone the same as me then possibly but other people are living their lives because they're different aren't they so I'd say that's very seldom true.

logical - answering a different question here i.e. around comparing themselves w. others.

leads to ACT error too.. low score doesnt represent level of Ψ inflexibility here / ~~Does~~ do low acceptability of difficult T + F

**AAQ5**

**P6** 'Seldom true'. Um – as I say- if I've got something that's niggling I'll deal with it so, um- or I get prompted from my partner to deal with it, you know, um- so I'd say 'seldom true' number three.'

logical - they are answering a different question, they are talking about ~~thoughts~~ ~~they~~ around problems they have 'dealt with' not those ~~Alth~~ they haven't. Therefore score not a true reflection.

ACT also? score doesnt reflect accurate / true level of acceptance etc.

**P10** *P:* Emotions cause problems in my life.' U, 'Very seldom true'. Yeah, it's sometimes, I suppose, I suppose it happens to everyone. Um, it causes problems, it's not major problems, um- it takes ages to answer some of these!'

*I:* No, no that's fine.

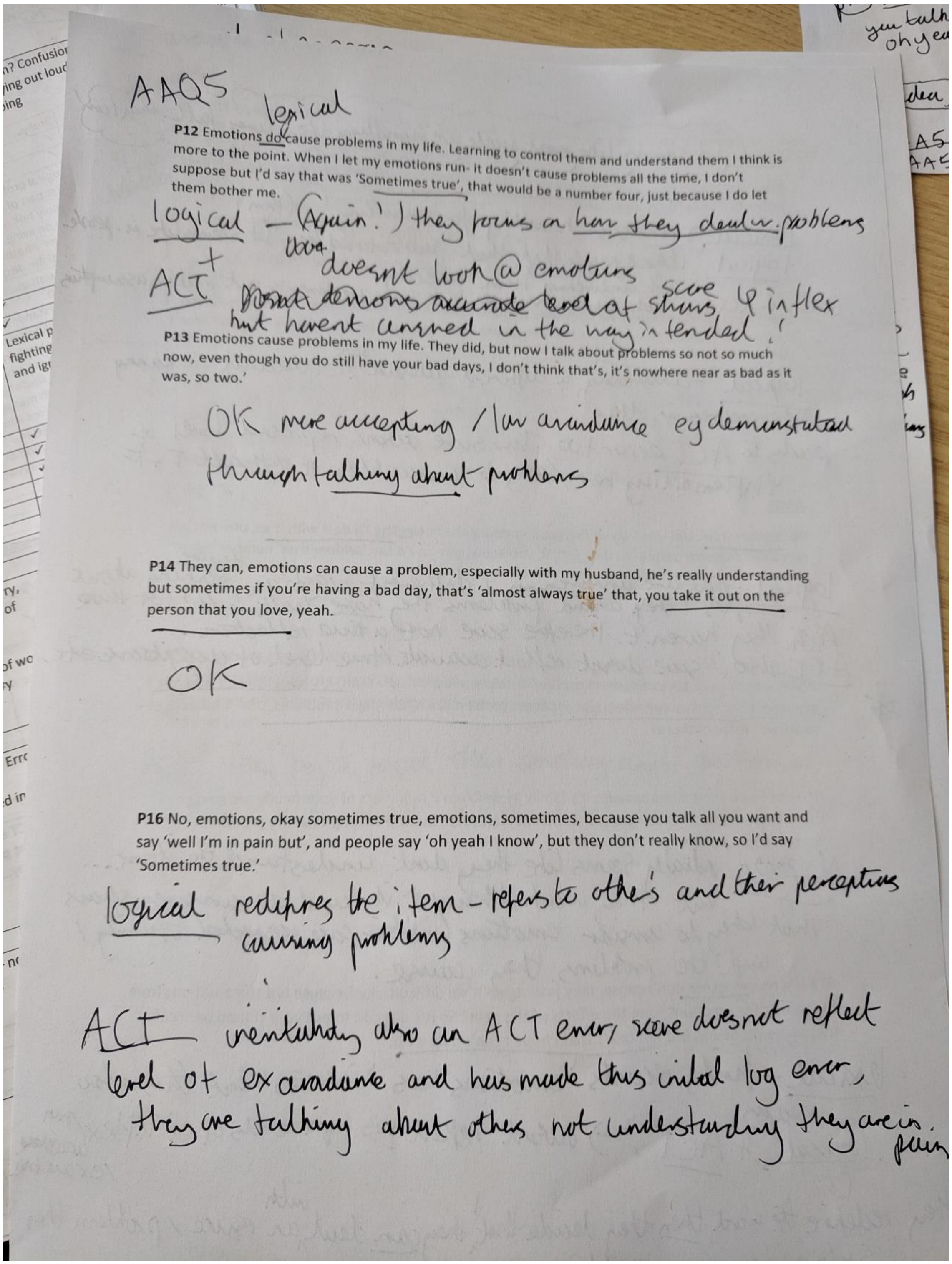
*P:* Um it-yeah, sometimes because it's family as well, um if members of your family are going through something that can get you down and, um yeah, but it's dealing with the problems. So 'Sometimes'.

No error initially seems like they dont understand the item... but they work through it and the answer shows that they do consider emotions (in their case connected to family) and the problems they cause.

**P11** Um no, my issues don't cause, they just- again it's a difficult one because is it the emotions from the pain or the emotions from the effects of the pain? So it's difficult to put your thumb on. Yeah, they cause problems, but not one's that I hope I can't deal with.'

lexical - firstly recasts 'emotions' as 'issues', but this also means logical + ACT? → doesnt reflect pos high levels of Ψ inflex / non acceptance / ex avoidance

they redefine the word then they decide that if they can deal with an issue / problem then it's no longer a problem



AAQ5

lexical

**P12** Emotions do cause problems in my life. Learning to control them and understand them I think is more to the point. When I let my emotions run- it doesn't cause problems all the time, I don't suppose but I'd say that was 'Sometimes true', that would be a number four, just because I do let them bother me.

logical – (Again!) they focus on how they deal w. problems doesnt look @ emotions

+

ACT ~~Doesnt demons accurate level~~ of score shows & inflex but havent answered in the way intended!

**P13** Emotions cause problems in my life. They did, but now I talk about problems so not so much now, even though you do still have your bad days, I don't think that's, it's nowhere near as bad as it was, so two.'

OK more accepting / low avoidance eg demonstrated through talking about problems

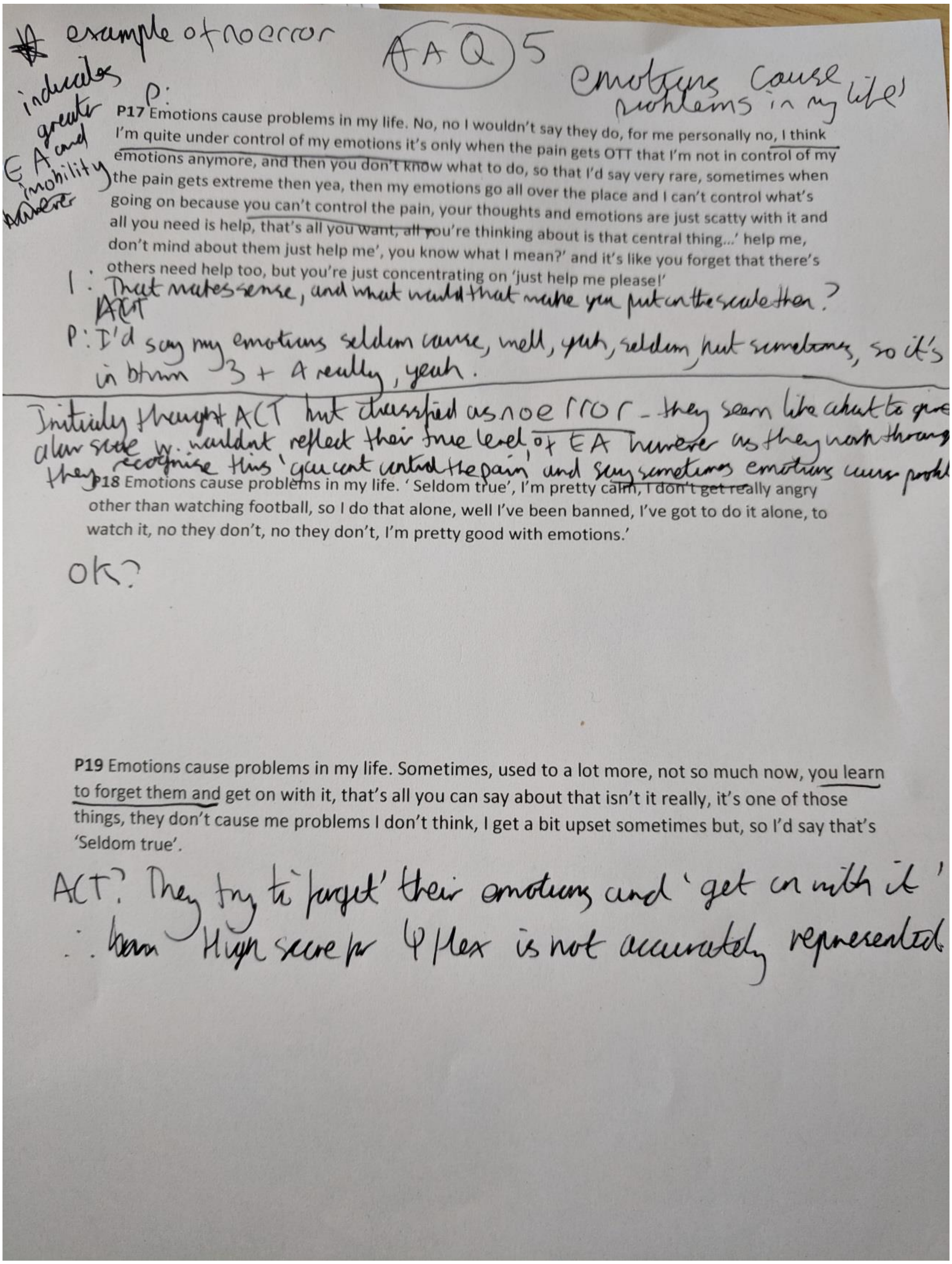
**P14** They can, emotions can cause a problem, especially with my husband, he's really understanding but sometimes if you're having a bad day, that's 'almost always true' that, you take it out on the person that you love, yeah.

OK

**P16** No, emotions, okay sometimes true, emotions, sometimes, because you talk all you want and say 'well I'm in pain but', and people say 'oh yeah I know', but they don't really know, so I'd say 'Sometimes true.'

logical redefines the item – refers to other's and their perceptions causing problems

ACT inevitably also an ACT error, score does not reflect level of exp avoidance and has made this initial log error, they are talking about others not understanding they are in pain



example of no error

AAQ 5

'emotions cause problems in my life'

indicates greater EA and imobility however

P:

**P17** Emotions cause problems in my life. No, no I wouldn't say they do, for me personally no, I think I'm quite under control of my emotions it's only when the pain gets OTT that I'm not in control of my emotions anymore, and then you don't know what to do, so that I'd say very rare, sometimes when the pain gets extreme then yea, then my emotions go all over the place and I can't control what's going on because you can't control the pain, your thoughts and emotions are just scatty with it and all you need is help, that's all you want, all you're thinking about is that central thing…' help me, don't mind about them just help me', you know what I mean?' and it's like you forget that there's others need help too, but you're just concentrating on 'just help me please!'

I: That makes sense, and what would that make you put on the scale then?

ACT

P: I'd say my emotions seldom cause, well, yeah, seldom, but sometimes, so it's in btwn 3 + 4 really, yeah.

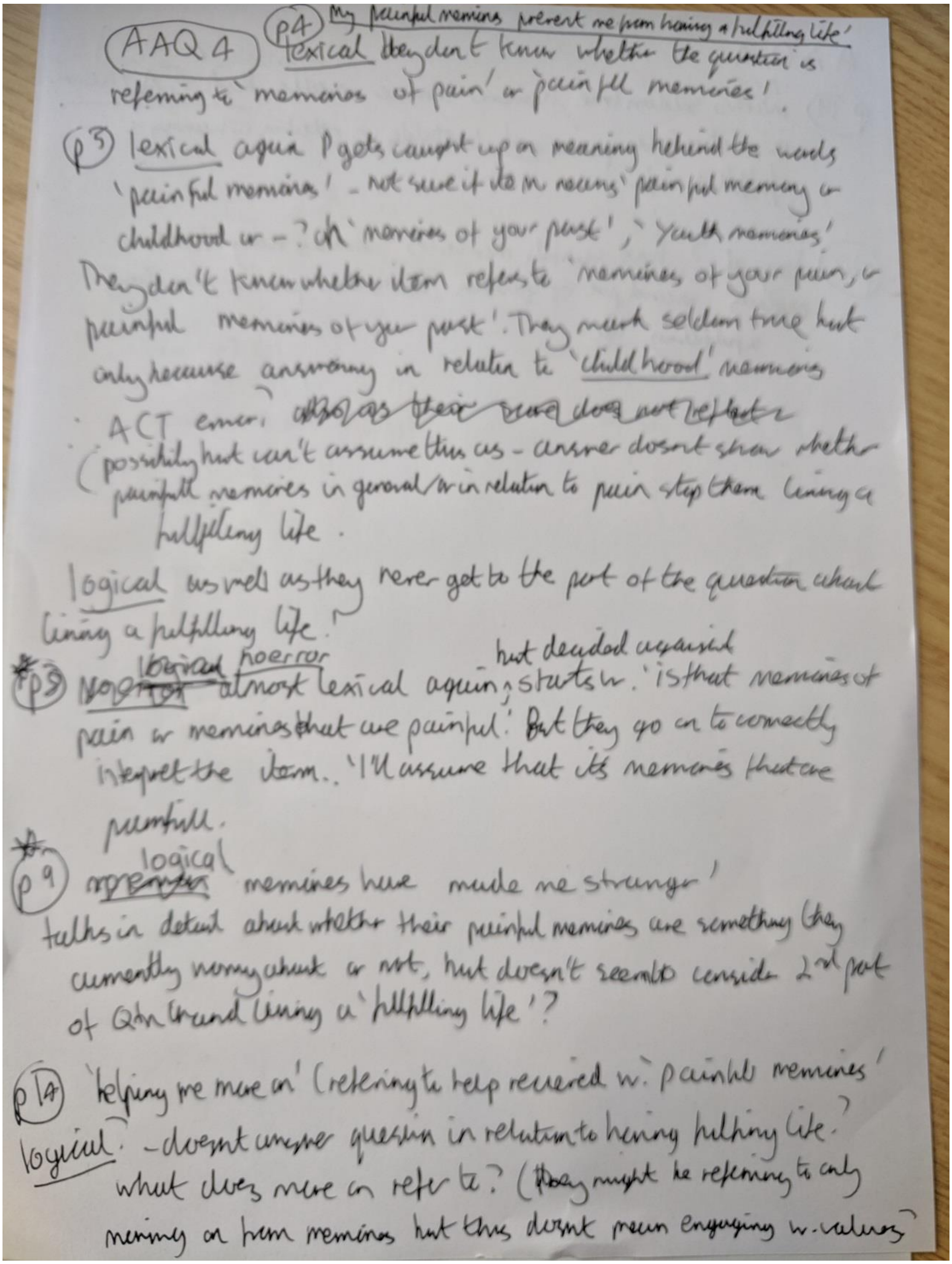
Initially thought ACT but classified as no error – they seem like what to give a low score w. wouldnt reflect their true level of EA however as they work throug they recognise this 'you cant control the pain' and say sometimes emotions cause probl

**P18** Emotions cause problems in my life. ' Seldom true', I'm pretty calm, I don't get really angry other than watching football, so I do that alone, well I've been banned, I've got to do it alone, to watch it, no they don't, no they don't, I'm pretty good with emotions.'

ok?

**P19** Emotions cause problems in my life. Sometimes, used to a lot more, not so much now, you learn to forget them and get on with it, that's all you can say about that isn't it really, it's one of those things, they don't cause me problems I don't think, I get a bit upset sometimes but, so I'd say that's 'Seldom true'.

ACT? They try to 'forget' their emotions and 'get on with it' ∴ [illegible] High score for ψ flex is not accurately represented



AAQ 4 (p4) My painful memories prevent me from having a fulfilling life'

lexical they don't know whether the question is referring to 'memories of pain' or 'painful memories'.

(p5) lexical again P gets caught up on meaning behind the words 'painful memories' - not sure if item means 'painful memory' or childhood or - ? ch 'memories of your past', 'youth memories'

They don't know whether item refers to 'memories of your pain', or 'painful memories of your past'. They mark seldom true but only because answering in relation to 'childhood' memories

· ACT error; ~~also as their score does not reflect a~~ (possibility but can't assume this as - answer doesn't show whether painful memories in general/or in relation to pain stop them living a fulfilling life.

logical as well as they never get to the part of the question about 'living a fulfilling life'.

*(p8) ~~No error~~ ~~logical~~ no error ~~almost~~ lexical again; starts w. 'is that memories of pain or memories that are painful'. But they go on to correctly but decided against interpret the item. 'I'll assume that it's memories that are painful.

*(p9) ~~no error~~ logical 'memories have made me stronger'

talks in detail about whether their painful memories are something they currently worry about or not, but doesn't seem to consider 2nd part of Qn around living a 'fulfilling life'?

(p14) 'helping me move on' (referring to help received w. 'painful memories'

logical? - doesn't answer question in relation to having fulfilling life?

what does move on refer to? (they might be referring to only moving on from memories but this doesn't mean engaging w. values)