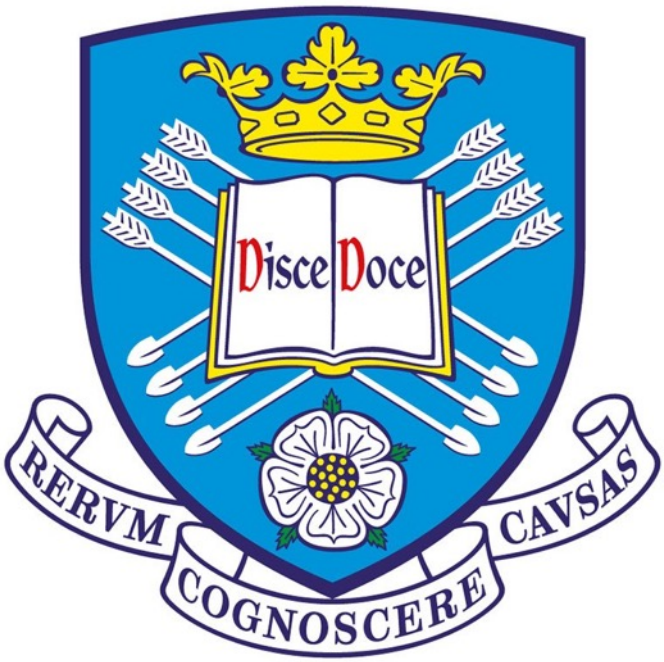


Constructing a Reliable and Valid Measure of Multidimensional Poverty

Volume One

Calum J R Webb
Department of Sociological Studies

A thesis submitted for the degree of Doctor of Philosophy

May 2019

Constructing a Reliable and Valid Measure of Multidimensional Poverty

Calum J R Webb, Department of Sociological Studies, the University of Sheffield

May 2019

Abstract: This thesis presents evidence that suggests a theoretically robust, statistically reliable and valid measure of multidimensional poverty can be constructed through the use of secondary data, factor analysis and associated statistical techniques. Competing 'quantitative operationalisations' of poverty (Lister, 2004: 6) are typically restricted by a methodologically necessary simplification to income. This ignores many of the complexities found in qualitative research and theories of multidimensional poverty, as well as central ideas about participation, which define the concept more accurately and fairly than income and consumption in isolation (Townsend, 1979; Alcock, 2006; Walker, 2014). Assumptions are often made about how broadly such measures can be applied and how accurately they reflect the underlying concepts they propose to represent. Although more recent attempts have tried to create more adequate measurements, these have been met with some considerable backlash in their transition into policy and are still constrained by a reduction to income, with limited scope for evaluating their reliability and validity.

The research presented uses data from the *United Kingdom Household Longitudinal Survey* and techniques from structural equation modelling, factor analysis, and data science, to build on the work of Tomlinson, et al. (2008) that showed it was possible to create a quantitative multidimensional construct that combines material, financial, and psychosocial dimensions into a single measure. More specifically, this study applies statistical techniques from the field of psychometrics to explore whether such an approach is robust, and can be applied equivalently across over time and between different groups. Producing such evidence may be crucial to developing our understandings of poverty in future, and breaking impasses in policy which result in selective uses of metrics and continual reimaginings of the concept which, from a social justice perspective, do very little to advance the cause and representation of people living in poverty.

Acknowledgements

I would like to acknowledge the support of friends and colleagues both within and outside of the Department of Sociological Studies, particularly Professor Kate Morris, who has been a inimitable mentor, and Professor Nathan Hughes, who, at the eleventh hour and with a great many existing commitments, provided indispensable comments on the final draft of this thesis.

I would also like to thank my family for being an unfaultering spring of good character, perseverance, and hard work, from which to draw on.

Lastly, I thank and wish to acknowledge the bravery of the uncountable number of people who have contributed to our understanding of poverty by telling their stories. I hope that this work has a place in the effort to have their voices heard.

As with all theses, unspoken between every line below is the patient encouragement, good advice, and learning experiences from more people than I can state, some of whom are, for better or worse, no longer a part of my life. This is a result of their labour too.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction
Why a real, robust, multidimensional measure of poverty is needed

> *"If measurement is arbitrary and irrational, it is impossible either to concoct the right policies for the alleviation or eradication of poverty, or monitor their effects closely"*
>
> Peter Townsend, *The International Analysis of Poverty,* 2010 [2002]: 82-83

> *"It would patently be preferable to specify a single poverty standard and hence obtain a clearcut measure of the extent of poverty. This would however represent an "all or nothing" approach, since those who disagree with the standard are likely to reject the findings out of hand. ... [A multidimensional approach] leads to less definite answers but, one hopes, commands a wider degree of support"*
>
> A. B. Atkinson, *How Should We Measure Poverty*? 1985: 16-17

> *'To move straight to definitions and measures without first considering the broader concepts can result in losing sight of wider meanings and their implications for definitions and measures. In particular, it can exclude the understandings of poverty derived from qualitative and participatory approaches. These frequently highlight aspects of poverty that lie outside definitions focused on income and material living standards and that can be difficult to measure in surveys[…] starting at the bottom with measures can encourage confusion between measures and definitions, so that arguments about competing definitions of poverty often turn out to be arguments about competing measures. […] omitting the conceptual level can encourage a myopic, technocratic approach that, in its preoccupation with measuring poverty's extent and depth, overlooks how it is experienced and understood.'*
>
> Ruth Lister*, Poverty (2004: 5-6)*

The three quotes above have shaped the development and rationale behind this thesis. Poverty is deeply complex and extends far beyond the idea of low income (Lister, 2004; Walker, 2014). However, measurements of poverty typically face restrictions that limit them to income, even in very recent designs (Social Metrics Commission, 2018). These are often the result of a mixture of political and methodological constraints. If a measure is too complex, or raises difficult questions, it is likely to be rejected for not being relevant or useful to policy (Atkinson, 1985; Alcock, 2006). If a measure combines certain indicators uncritically, and cannot be assessed to be robust and therefore relies on whether it fits with a general consensus of what the extent and depth of poverty should be, it is methodologically flawed (Spicker, 2004; Lister, 2004; Alcock, 2006; Tomlinson, et al., 2010).

This has resulted in something of an impasse in the development of measures of poverty that better reflect the concept. More and more complex ways of developing baskets of goods and establishing poverty lines have been developed, such as the Minimum Income Standard or the Social Metrics Commission measure, but this has resulted in contradictory statements about what the reality of poverty is in the UK, often diverting narratives into technocratic debates on the merits of each and away from the substantive issues they are supposed to bring to light (Piachaud, 1987; Lister, 2004). Tomlinson, et al. (2008) developed an approach that demonstrated a multidimensional measure of poverty, appropriately weighted for different dimensions, could be developed using factor analysis methods. Furthermore, the approach showed that each dimension could be retained and explored independently, to better understand the shifting patterns of multidimensional poverty. The development of this approach within the specific methodological framework meant that a potential quantitative measure of poverty could be applied and assessed using powerful statistical tools. However these assessments of reliability, validity, and measurement invariance were not conducted by Tomlinson, et al. (2008) at the time of developing their measure.

This thesis extends the methodology of Tomlinson, et al. (2008) to create a similar measure of multidimensional poverty using the *Understanding Society* survey (ISER, et al., 2018). It incorporates a number of methods from the fields of psychometrics and data science to test whether the approach to measuring multidimensional poverty in the way proposed by Tomlinson, et al. (2008) is statistically robust: whether the dimensions used and overall poverty index can be said to be reliable and valid. This is an slow and careful process involving visualisation; the testing of proposed dimensions on multiple sets and subsets of data; and numerous iterative models changing very small constraints in order to assess whether the same measure can be applied equivalently between different time points and groups. The extensive appendices testifies only to a part of the length of this process. The results from nearly 600 models were needed to make conclusions about the reliability and validity of the multidimensional approach. Data for over one-hundred variables for over 290,000 observations had to be merged, tidied, visualised, and adjusted for missingness. This equated to over 90,000 lines of code in STATA, R, and Mplus.

But the reader will find few revelations in this thesis. The structural equation modelling approach to multidimensional poverty undertaken broadly reflects what we would expect from qualitative research and theories of poverty: it works well, and it can be universally applied to all time points there was data available for and between different household types, genders, and cohorts. So why do this? Why put myself, and the reader, through this barrage of tests and tables to not show some radical new knowledge about the depth and extent of poverty? Because people living with, or who have lived in, poverty deserve a quantitative operationalisation of poverty that does justice to their real experiences (Lister, 2004; Tomlinson, et al., 2010). Part of doing that involves going through this process; someone had to do this technical work to demonstrate that a multidimensional quantitative measure is possible and desirable. This is a necessary progression towards developing multidimensional measures that are neither arbitrary nor irrational (Townsend, 2010 [2002]); these techniques show such measures to be both meaningful and rational.

The first section of this thesis outlines existing approaches to the measurement of poverty and why, although some are better than others, the income-centric approach used by all is necessarily impossible to validate statistically and in many cases is irrational when considered in reference to what we actually mean when we talk about poverty. The section then discusses theories and conceptualisations of poverty, bringing the focus back to what poverty actually means, by drawing on the research and writing of, in particular, Peter Townsend, Amartya Sen, Bob Baulch, Ruth Lister, Paul Spicker, and Robert Walker. This provides the fundamental foundation for the construction of a truly multidimensional measure that does not, as Ruth Lister (2004: 5) argues, lose sight of the wider meanings of the concept of poverty and their implications. Finally, the section also considers the dynamic nature of poverty, especially between different groups of people and at different points in time. Measures of poverty should be subject to change as the meaning of the concept changes in society over time, or if it differs between certain groups.

The following section then considers how the limitations in existing measures of poverty and how the challenges of measuring such a complex and dynamic concept may be met through the use of more advanced quantitative methods. The following research questions are identified:

1. Is there clear evidence for underlying ‘dimensions’ of poverty within secondary data survey items, as theory would suggest?
2. Is the construction of a multidimensional poverty index that uses vastly different measures possible, and how well do these factors fit within a broader multidimensional poverty construct, and can this tell us anything about the nature of poverty?
3. Are the dimensions of poverty identified in theory valid, are they distinct enough to be real differences, when using variables from existing datasets?
4. To what extent are these dimensions of poverty consistent over time and between groups? Does poverty mean the same thing when measured in this way at different years, or is it very volatile? Can the same multidimensional measure of poverty be applied to different groups, men and women, different households or cohorts, equivalently?
5. What might be the consequences of this more complex measure of poverty for social policy?

The section then outlines various advanced quantitative methods based approaches to answer each of these questions, and places them within a methodological framework that draws on novel sociological approaches to epistemological claims, focusing more on reflexivity and practical objectivity as opposed to value objectivity. The chapter discusses the secondary data and its variables and structure. It considers the way that structural equation modelling based techniques work, in particular the role that fit statistics play in making comparisons between models. The use of exploratory data analysis using visualisations of data and the reduction of the use of this technique in applied statistics and its following resurgence in more modern data science workflows is also discussed. Finally, methods for testing reliability and validity, and for dealing with missing data in complex surveys, are described.

In the last section the findings are presented in three chapters. Firstly, the thesis considers what patterns between indicators emerge when a visualisation tool – a hierarchically clustered nonparametric correlation plot – is used to try and identify potential dimensions of poverty and any evidence for an overall multidimensional poverty measure. There is fairly clear evidence to support the case for a

multidimensional cluster of indicators which broadly matches what we would expect from the theoretical conceptualisation of poverty based on qualitative research into lived experiences. Five strongly interrelated factors are observed, which include financial strain, material deprivation of short-term consumable items, material deprivation of longer-term household commodities, psychosocial strain, and life satisfaction. Additional factors are also observable, including social isolation, civic participation, and socioeconomic status, but there is limited evidence from the visualisation to suggest these are fundamental to a broader multidimensional poverty concept.

These suggested clutsers of items are then taken forward to exploratory and confirmatory factor analysis to firstly construct each dimension and test for reliability, validity, and time-based measurement invariance. Exploratory and confirmatory factor analysis is then used to construct a higher-order multidimensional poverty measure from the individual dimensions, which is also tested for measurement invariance and reliability. The research found that each individual dimension constructed was an effective, reliable, and time-invariant measure of the underlying construct, and that the higher-order multidimensional poverty measure also seemed to approximate a real underlying construct well, showing measurement invariance and reliability. Findings for validity were somewhat more mixed: the majority of dimensions showed discriminant validity from each other, suggesting the dimensions measured can be identified sufficiently distinctly from similar dimensions, but some dimensions, such as financial strain and material deprivation, were so strongly related that they were not necessarily distinguishable from one another quantitatively.

Finally, the last chapter of the findings section outlines whether the proposed multidimensional poverty factor can be applied equivalently to different household types, genders, and cohorts. This limited analysis of social divisions could have been extended much further if adequate data was available, but unfortunately there were technical reasons why divisions like disability and ethnicity could not be assessed for invariance. This is an important unknown that is often assumed to be true by a large number of conventional metrics. The analysis suggests that this way of measuring poverty can be applied equivalently to seventeen different household structures, to both men and women, and across seven different age groups.

All of these chapters tell us something substantive about the nature of multidimensional poverty. Firstly, that theories of multidimensional poverty are broadly correct, with a possible lack of clarification on what is meant by social isolation and how this is translated to measurement. Secondly, that there is evidence that a psychological and sociorelational dimension of poverty, approximated by life satisfaction, can be considered almost as important to the underlying structure of multidimensional poverty as financial strain itself. This dimension may reflect Walker's (2014) identification of shame as a universal dimension reflected in experiences of poverty. Thirdly, it suggests that there is still much we do not know about the nature of *becoming poor*, and how each dimension relates to the others in this process. The findings suggest that financial strain seems to characterise a 'material core' (Lister, 2004), whereas life satisfaction forms a 'sociorelational core'. These then ripple outwards. Financial strain manifests in a decline in participation and, eventually, a reduction in household commodities. Poor life satisfaction manifests in psychosocial strain – increased general anxiety and depression, a loss of a sense of control, and increased social dysfunction. Further research may wish to focus more on these transitory states between initial strain and loss of self-actualisation, followed by externalised relational withdrawal from full participation and the manifestation of psychological distress which may lead to clinical morbidities.

To summarise, with regard to measurement, the thesis suggests that:

- A multidimensional measure of poverty using *Understanding Society* data and factor analysis is possible and desirable. It reflects theoretical perspectives and successfully combines material and nonmaterial dimensions.
- The measure is both reliable and valid. It does not seem to reflect random patterns in the data and shows that the dimensions identified are real and measurable using the technique. The dimensions themselves are also largely representing unique concepts, they are not merely alternate ways of measuring the same thing.
- The measure can be applied at any time point in *Understanding Society* data, and can be used to compare different households, genders, and

cohorts with the confidence that such comparisons reflect real differences in multidimensional poverty.

- While such a measure is demonstrably possible, it may still not be desirable for political and practical reasons. There are still numerous technical difficulties with employing this method, and the interpretation of any resultant statistic is reasonably complex. Interpretation can be made easier using certain transformations, for example, choosing a cut off in terms of standard deviation, or inverting the equations used to derive its predictions. But underlying this is the fundamental question of whether a 'cut off' that allows us to quantify the extent of poverty is valid at a conceptual level. Based on our theoretical understandings, poverty should be measured as a continuum of experience, rather than as a binary. A continuous measure is clearly less attractive to policy, as it is harder to interpret and does not imply the same 'goal' of reducing poverty figures to zero. However, conceptualising poverty as a spectrum is dominant in research into the effects and lived experiences of poverty and, although binary classifications may serve a function in policy, there is a risk that these blunt distinctions do little to further our understanding of the true impact of poverty when they are used in sociological research. As Lister (2004) has warned, these measures often end up being incorrectly interpreted as containing complete meaning, potentially doing a great disservice to the descriptions of those who are currently or have ever experienced poverty, and therefore to these people themselves.

Chapter Two:
What does poverty mean and how should we measure it?

2.1 The conceptualisation and measurement of poverty in the United Kingdom

A clear understanding and assessment of existing conceptualisations of and approaches to the measurement of poverty is crucial to constructing arguments for multidimensional measures. Measurement of poverty must reflect what we understand poverty to be; returning to the quote used earlier, Peter Townsend (2010 [2002]) argues that if measurement is arbitrary or irrational, correct policies for poverty alleviation cannot be developed. One way to judge whether measurement meets the criteria of being meaningful and rational is by comparing a given operationalisation of poverty to the lived experiences and theoretical understandings of poverty found in the body of literature concerned with the study of poverty. This can be done by firstly considering our historical ontology of poverty, what does poverty mean, before considering our methodological operationalisations of poverty. While income-based measurements appear broadly in agreement with the distinction between 'absolute' and 'relative' poverty, they begin to look wholly inadequate when contrasted with multidimensional conceptualisations.

This chapter begins with a discussion around what is meant by the term poverty, and the various sociological approaches to defining and conceptualising deprivation and needs, and how these translate to study and measurement. This is framed by broad ideas of absolute and relative poverty and how such ideas shape public and political understandings of poverty (Spicker, 2007; Tomlinson, et al., 2010). The literature reviewed here problematizes existing measures of poverty in developed countries such as the UK, while acknowledging their benefits. A comparison of various measures demonstrates that how poverty is measured matters; these views and measures can have an almost myopic focus on consumption and minimise the multidimensional aspects of poverty: assets, dignity, autonomy, respect, stigma, rights, citizenship, and others which are frequently stressed in narratives of those living in poverty (Baulch, 1996; Beresford, et al., 1999; Lister, 2004; Spicker, 2007). Limitations in methodology and a focus on personal consumption means that non-material dimensions of poverty are often ignored (Tomlinson, et al., 2008; Walker, 2014). When these multidimensional elements are incorporated into analyses (*ibid*), or measures are adjusted sensitive to the cost of a consensual minimum standard of living (Bradshaw, et al., 2008), as opposed to an

arbitrary threshold, the picture and trajectory of poverty changes substantially – as, Townsend (2010 [2002]) argued, would any approaches to research and subsequent conclusions and policy reactions.

The following literature lays the theoretical foundation for the conceptualisation of poverty throughout the research. What follows from this grounding are questions around social divisions and dynamics of poverty. Because of this scope there is not the capacity to discuss in detail the full development of poverty and poverty studies throughout history or across geographies, however, it is noted that these are important provisos to a full understanding of the study. These theoretical debates and their empirical evidence are the spine of the thesis and inform the choice of methodology, where poverty is treated as a multidimensional concept. This is reflected in the choice of method. The major contribution to the field from this review involves taking a step back from the focus on existing methods of quantifying poverty and synthesizing the theoretical literature and its original critiques of official measures with more contemporary practices, using empirical comparisons from a range of different approaches to measurement to show that the problems inherent in the conceptualisation of poverty are still relevant to how it is operationalised in quantitative social science. A question of whether it is possible to construct theoretically robust multidimensional quantitative operationalisations of poverty for the purpose of understanding its impact in the UK emerges, and the subsequent consequences of theoretically weak operationalisations and their potential impact on subsequent political and policy decisions are discussed.

Before considering the substantial gaps that form the motivation for this research, however, it is important to understand the basis on which most disagreements with the measurement of poverty exist: the distinction between absolute and relative definitions of poverty. While there is an attempt to distinguish between them made here, it is very difficult to speak about one without reference to the other. Therefore, as context prior to the following two short sections, it should be noted that 'absolute poverty' refers to a specific conceptualisation of poverty that is primarily concerned with a level of subsistence needs – food, water, shelter - required to live, as measured by *continued physiological existence* (Alcock, 2006). In contrast, 'relative poverty' can be imagined as a conceptualisation of poverty where the poor are not afforded the adequate

conditions required to fully *take part in life* (Townsend, 1979; Alcock, 2006). What does it mean, then, to live a life free from poverty?

In popular and political discourse poverty is not often thought of as a problematic concept to define and the use of the term often implies a shared definition uncritically, however, poverty itself does not have a single fixed meaning. There is on-going rhetoric about an 'all-out assault on poverty' (Wintour & Watt, 2015) and, recently, figures reporting that a third of people in the UK have experienced poverty since 2008 (Tonkin, 2016), however, these discussions do not often explicitly state of what is meant by 'poverty', and typically reduce it to one singular, albeit important, dimension: low household income. Furthermore, some would argue that the idea of poverty in developed countries is itself a misnomer when compared to the experiences of the developing world: 'People who hold that poverty is different in rich and poor countries are saying that… what passes for poverty in a rich country would not pass for poverty in a poor one' (Spicker, 2007: 11). Does poverty refer to deprivation relative to a person's peers in society or does poverty mean total destitution where people are unable to sustain themselves? What of other, non-material, dimensions of poverty? Almost certainly most would agree a person could be considered deprived of social as well as material components that exclude them from full participation in society? Furthermore, why is it that we often conceptualise poverty as a cliff edge, wherein an increase in household income of the smallest denomination can be considered the difference between being 'poor' and 'not poor'? What people, organisations or systems decide which ideas of poverty are valid, and what is the justification for where such poverty lines are drawn?

2.2 Absolute Approaches

John Maynard Keynes wrote about the difference between two 'classes' of need - absolute needs and relative needs - in discussing the potential economic relations of future generations (Keynes, 1963 [1930]). For Keynes, absolute needs referred to needs of subsistence that had plagued civilisations throughout history and had the potential to be eliminated through economic and technological progress, needs that existed independent of the situations of others. Relative needs, on the other hand, exist as a 'desire for superiority' that 'may be insatiable' because of a tendency for them to shift ever upwards (Keynes, 1963 [1930]: 365-366). Keynes saw a real possibility that these

absolute needs might be eradicated in the 100 years following his writing and that people might become preoccupied by satisfying relative needs through an abundance of leisure made possible by technological progress. Over 85 years have passed since Keynes' essay on 'economic possibilities for our grandchildren' and neither absolute nor relative needs have been eradicated, and they continue to preoccupy the concerns of global and national social policy.

A globally influential but strongly contested idea of poverty is that of 'absolute poverty'; a condition that the Copenhagen Declaration of the World Summit for Social Development characterises as: 'severe deprivation of basic human needs, including food, safe drinking water, sanitation facilities, health, shelter, education and information' (Spicker, et al., 2007: 7). A key benefit of measuring poverty against absolute criteria is that there is a logical 'end-goal', a basket of goods, that can be broadly approximated using levels of income against the cost of services and resources that is obscured in relative measures by shifts in the overall shape of income distributions. In theory, these rates can be compared globally as they are measures of universal physiological human needs (Anand, et al., 2010; Tomlinson, et al., 2010). Amartya Sen (1983: 159) argued that there is an 'irreducible absolutist core in the idea of poverty', that the inability to meet nutritional requirements, avoid preventable disease and have shelter, clothes, transport, education and a life free of shame, were fundamental components of poverty that should not be assumed to be captured by the picture of relative poverty. However, it has been argued that the claim of this 'universality' of manifestations of absolute poverty can itself be considered relative. Peter Townsend (1985) responded that it is important to consider the relativity to society of these seemingly 'absolute' characteristics. For example, nutritional standards are heavily centred around the nature of work in different societies and shelter implies a set of cultured notions around societally acceptable levels of privacy, access to tools and leisure, and non-strenuous climate.

Absolute poverty, for some time, has been considered a problem predominently confined to developing and poor countries in the Global South, especially by the British public (Castell & Thompson, 2007). This is especially clear when considering the ways in which global organisations compare poverty rates in different countries. The United Nations Development Programme (UNDP), for instance, uses two different indices of

poverty for developing and advanced economic counties, HPI-1 & HPI-2 respectively (Spicker, 2007: 11). Arguably, the most influential measure of absolute poverty, the World Bank's PPP (Purchasing Power Parity) $1-a-day line, sometimes set at $2 or $1.25, has also not typically been considered of major importance for poverty analysis in advanced economic countries and poverty rates using this measure are usually not even published for these countries (Anand, et al., 2010).

However, there has been a revival of absolutist measures of poverty in recent years, especially in Europe and increasingly in the United Kingdom after the 2008 fiscal crises and subsequent austerity measures. Tomlinson, et al., (2010: 357) identify the reintroduction of absolute poverty in measures used by the DWP, and their prominent feature in the Laeken indicators used to measure social inclusion in Europe. Notably, measurement of severe poverty in the United Kingdom has been based on people being unable to afford four or more of nine items: rent/mortgage, utility bills, or loan repayments; fuel for warmth; unplanned expenses; regular consumption of foods containing protein; a holiday the duration of one week not staying at home or with extended family; a television set; a washing machine; a car; or a telephone (Office for National Statistics, 2013: 6-7). Of course within these we can see the relative basis which they lie upon highlighted by Townsend's (1985) criticism of Sen (1983), where items such as televisions, holidays not visiting family, and car ownership reflect to some extent relative cultural preferences and labour arrangements.

More recently, absolute poverty has been of central importance in discussions in the UK, spurred by growing cases of what is referred to as 'food poverty' (Garthwaite & Bambra, 2015). This has been reflected in an uptake in the use of food banks, charitable organisations that provide sundries, preserved and fresh food to those unable to afford these items (Cooper, et al. 2014; Dowler & Lambie-Mumford, 2015). The largest food bank charity in the UK experienced a 263 per cent increase in the number of people using food banks in 2013-14 (nearly 915,000 people) compared to the previous year, with the increase being most notable in the poorest areas (Garthwaite & Bambra, 2015). In 2010-11, the number of food bank users was 60,000 (*ibid*).

One of the outcomes of this attention on absolute poverty has been a redefinition and evaluation of destitution in the UK by Suzanne Fitzpatrick, et al. (2016), for the Joseph

Rowntree Foundation. Destitution in the UK in a given month is defined by the authors as the lack of two of more of the essentials of: shelter, measured as 'sleeping rough' for more than one night; food, measured by having fewer than two meals a day for two or more days; heating, measured as the inability to heat ones home for five or more days; lighting, measured by an inability to light ones home for five or more days; clothing and footwear, specifically lack of weather appropriate attire; and access to basic toiletries such as soap, toothpaste, shampoo and toothbrushes (Fitzpatrick, et al., 2016: 2). This definition was reached, following a review of the literature, through a consensual approach involving synthesizing the views of 50 experts across England, Wales, Scotland and Northern Ireland, with an omnibus survey of 2,000 members of the public to gauge their opinion on the appropriate definition of destitution in the UK (Fitzpatrick, et al., 2016: 1). There is a stark contradiction between the process of reaching this measure and what it claims to measure that should be obvious and neatly illustrates the irrationality of absolute definitions of poverty. If an absolute measure based on destitution is objective and scientific, as such measures propose to be, why did it need to be decided upon through a consensus of experts and the public?

Researchers determined whether the reason respondents forwent these items was because they could not afford them by asking directly whether this was the reason, checking that income was already below the 60 per cent of median relative poverty line, and finding out whether respondents had zero or negligible savings (*ibid*: 2). Respondents were also defined as destitute if their income was lower than a set absolute minimum, adjusted for family size. This income level was determined by taking an average of: the 'spend on these essentials of the poorest 10 per cent of the population; 80 per cent of the JRF Minimum Income Standard costs for equivalent items; and the amount that the general public thought was required for a… household to avoid destitution' (*ibid*: 2). The resulting 'extremely low income' criteria were set, after housing costs, at £70 per week for a single person living alone, £90 for a lone parent with one child and £140 for a couple with two dependent children. They estimate that 668,000 households, comprising of around 1.25 million people, 312,000 of whom are children, meet their criteria of destitution. This is argued to be a conservative estimate of destitution in the UK as it captures primarily those who had to utilise crisis services. Over time, the authors identify a fall in severe poverty since 1996-7, followed by a steep rise in 2009 where the percentage of households in severe poverty reached almost

2.5 per cent, the highest it had been since 1997. This rate did not dip below 2.25 per cent in the time between 2009 and 2012 (the latest available data for the longitudinal analysis, *ibid*: 26-27).

2.3 Relative Approaches

Central to measurements of relative poverty are ideas around participation and relative capability. Theoretical approaches developed by Peter Townsend (1979) and Amartya Sen (1989) respectively claim that poverty can be understood as the inability to participate fully in society, or undertake activities and achieve a good and valued state of life. Although these approaches stress aspects of poverty beyond economic metrics (discussed later), income and resources play a critical role in determining the extent to which a person can participate in society, and the extent of the capabilities accessible to achieve a good life. These approaches are both inherently relative and, because participation in society is about more than economic activity, multidimensional. As will be seen in the comparison of relative poverty measurements, relative low income is often assumed to approximate many dimensions of deprivation.

The degree to which a person is considered deprived of participation in society is itself relative to the ways that one can participate and the point at which one can be considered to be able to 'fully participate', with benchmarks often being the participation of peers and what causes certain groups to be able to participate less or more in society (Townsend, 1979; Alcock, 2006). Similarly, capabilities required for a life of value shift as what is considered valuable at individual and societal levels is reshaped, although Sen (2004) does acknowledge 'basic capabilities' such as education, health, nutrition and shelter and sought to reconcile ideas absolute and relative poverty within a capabilities framework (Sen, 1999). However, this reconciliation was central even to the approaches of pioneers of poverty studies, especially Seebohm Rowntree, who recognised that relative social needs are deeply important and comparable to physiological needs, both of which form an integral part of understanding poverty as a whole (Lister, 2004: 27-28). In the absence of more advanced methods of analysis, technological capacity and better data, and for ease of use and interpretation, the established methods of measuring relative poverty have been focused on approximating the associated deprivations of lack of capabilities and participation by using variables derived from income, such as the 'Households Below Average Income' poverty line

(Tomlinson, et al, 2010; Ferragina, et al., 2013). The acknowledgement of being constrained in this way is often made, but there has been little movement away from a reliance on approximation-by-income.

This is partly due to how greatly received and widely adopted relative poverty measures have been. In response to arguments from Townsend and others, relative income poverty measurements have been eventually introduced and accepted across the UK and much of Europe and are now, typically, the primary measure of poverty (Tomlinson, et al., 2010). These approaches acknowledge that poverty is seen as relative to the standards of the society in which it is found – classically, Townsend (1979) defines poverty as the following:

> "*People are relatively deprived if they cannot obtain, at all, or sufficiently, the conditions of life – that is, the diets, amenities, standards and services which allow them to play the roles, participate in the relatonships and follow the customary behaviour which is expected of them by virtue of their membership of society. If they lack or are denied the incomes, or more exactly the resources… to obtain access to these conditions of life they can be defined to be in poverty."*
>
> Townsend, *Poverty in the United Kingdom*, 1979: 31

Townsend places emphasis on the state of poverty as being intrinsically connected to the expectations of the quality of life in society as a whole, the costs of attaining this, and how this changes over time. As Paul Spicker (2007: 15) explains, in a relative definition of poverty, 'as society changes, so does the standard of poverty. The richer the society becomes, the higher that standard is likely to be'. This means that a second important component captured willingly or unwillingly by relative poverty is inequality. As this conceptualisation of poverty is relative to the the conditions and wealth of society as a whole, the points at which people are considered poor are adjusted according to the quality of life that is attainable in the upper echelons of society.

While subsistance based absolute measures have been linked to evidence in human sciences on physiological needs, relative poverty has also been developed as a valid construct based on our philosophical and sociological understandings of need. Doyal and Gough (1984) introduce a dynamic relationship between what they term basic individual needs and basic societal needs, prerequisites which must be attained to achieve any other goal. These definitions aim to distinguish between needs and wants,

and are therefore helpful in operationalising the prerequisites of capability and participation, while distinguishing questions about the universal attainment of basic physiological, autonomical, and societal needs from those about the unequal achievement of 'wants'. Whereas wants are easily subject to political and individual preferential sensitivity, needs 'refer to a particular category of *goals* which in common parlance are believed to apply to *all* people' (*ibid*: 11, emphasis in original), however, this does not mean that political actors are not able to attempt to delegitimise needs by constructing them as wants, as has been the case of food banks and other welfare areas in the UK.

The authors move away from individualistic based approaches to needs which culminate in self-actualisation, such as those found in Maslow's hierarchy of needs (Maslow, 1943), to stress the dynamic relationship between individual and societal needs. Survival/health, and autonomy/learning, mutually reinforce societal needs of production, reproduction, culture/communication and political authority (Doyal and Gough, 1984: 10). Some needs here are also defined as the strategies required to attain other needs or wants. Needs cannot be considered to be satisfied by only attaining the essentials of physiological and social life, as they are essential stepping stones towards higher needs of autonomy and self-actualisation. These mechanisms and strategies to achieve higher needs or wants are also needs in themself. In other words, certain needs must be met in order to enable a person to act autonomically, according to their own choice, to an extent that is deemed societally necessary. This is exemplified by the value afforded to the act of chosing, of self-determination, especially so in modern Western societies which place a high significance on individual expression and consumer identity, and an inability to satisfy this need reflects poorly on an individual's view of the self, and their position in society and many communities (Verhaeghe, 2014). Examples can be taken from areas of personal choice around food, shelter, or clothing. A person may need food, but also have an implicit need to satisfy their hunger through choice according to their want for a particular taste, style of cuisine, and desire for their diet to facilitate a healthy lifestyle, or a lifestyle reflecting specific ethical choices.

A clear definition of what constitutes an experience of poverty is therefore vitally important in any description of adversity, as is the reader's understanding of what an author means when they use one concept or another, and the ways in which they have

operationalised said concept to try and measure it. In reference to social policy, Doyal and Gough write:

> "*... what began as a relatively clear distinction in ordinary discourse – that between needs and wants – becomes extremely muddled when its basis is challenged."*
>
> Doyal & Gough, *A Theory of Human Needs*, 1984: 13

A socio-political process – an essentially relativist process - often then determines the standard at which a person can be considered to be 'impoverished', based on which needs a society or group deems essential.

Although it might be comforting politically to confine the idea of absolute poverty or destitution as a problem restricted to the developing world, there is undoubtedly evidence of severe poverty in the UK that would not have been made apparent if the proposed 'objective value' of such measures no longer held any sway due to their irrationality. A focus purely on relative poverty may risk masking growing rates of absolute poverty due to changes in inequality (Tomlinson, et al., 2010) or changes in the cost of basic human needs and services that outstrip wage income growth, such as the 43.5 per cent increase in food prices from 2005 to 2013 or the 37 per cent increase in energy prices between 2010 and 2013 (Cooper, et al., 2014: 4). However, relative measures do capture vital dimensions of poverty, especially when considered longitudinally. These measures capture inequality and social mobility – ideas central to societies that value systems based on meritocracy and equity of opportunity (Spicker, 2007). It would be difficult to argue that a group in society that has enough income to afford only shelter, food, & fuel, relative to peers who are able to afford travel, leisure and other goods and services taken for granted by most in that society, should not be considered impoverished. However, nonwithstanding whether an absolute or relative definition is used, there is a clear disconnect between the complexity and nuance of theory; the binary classifications of absolute and relative poverty that stem from theory; and, most importantly for our understanding of the extent of poverty, the eventual binary metrics these theories inspire.

2.4.1 Comparing Measures: The Households Below Average Income Line

The most common of poverty measures, in the UK and many OECD countries, is the Households Below Average Income (HBAI) poverty line. This measure sets a threshold of 60% of median income and counts any household with income lower than this as poor (Gordon, 2006; Spicker, 2007). This threshold is theoretically designed to measure the point at which a person is unable to sustain a good quality of living and can be considered deprived relative to peers in society, based on an original analysis of the cost of achieving subsistence in addition to some additional needs of societal convention (Abel-Smith & Townsend, 2010 [1965]). Indeed, the "60%" threshold is so synonymous with relative poverty that a great many textbooks, and presumably their authors, are oblivious to its actual genesis.

The best historical basis for the origin of the 60 per cent threshold that could be found comes from Seebohm Rowntree's third survey of poverty in York; and this solidifies its inadequacies for use in the modern day. In 1950, Rowntree produced a further definition of poverty, comparing family expenditure on subsistence with a modest allowance for non-subsistence needs to the average earnings of manual workers (*ibid*). Rowntree found that the cost of meeting these needs for a family of five in 1899 was approximately 79 per cent of average earnings; 69 per cent in 1936; and *60 per cent* in 1950 (*ibid*: 176). The line is typically used for international comparisons, particularly between EU countries, to assist in policy learning. Clearly, the measure makes several assumptions about required income that may not adequately reflect the proportion of the population living in poverty, or the idea of poverty itself. To think that a threshold developed nearly seventy years ago, based on one specific family structure, in one city in England, would still be used uncritically to determine poverty rates across the entire developed world, is staggering. Further, to know that the threshold had changed by almost 20 per cent in the half-century between the first survey of York and the third, but that the current threshold has been assumed to be appropriate for the last seven decades, begs the question of where the rationality is in the Households Below Average Income measure.

Why is it irrational that this measure has not changed? Firstly, the threshold of 60% is not usually adjusted to the relationship between income/resources and low standards of living, or deprivation (Townsend & Gordon, 1989; Gordon, 2006), and is now essentially arbitrary (Tomlinson, et al., 2010). This means that thresholds are not

adjusted in line with the rising or falling costs of attaining a non-deprived quality of life, which may result in under- or over-estimating the number of people in poverty. For example, factors such as aforementioned rises in food prices (DEFRA, 2014) and other goods, in excess of wages and inflation, are not reflected in changes to the HBAI poverty line. Similarly, changes in what is required for a good standard of living, relative to society, are not adjusted over time as certain goods and services may become a part of essential daily life – examples of these may be personal computing goods or childcare services. Household needs change over time, and they change in different ways for different households. For example, in establishing a minimum income standard for 2014 compared to previous years, researchers identified that:

- The cost of a minimum food shopping basket has increased faster than general food inflation.
- Transport costs have increased, especially as public transport begins to service fewer routes. Many people have had to move to the use of taxis to travel to areas that used to be served by public transport.
- There is now an expectation for people, including pensioners, to have access to certain home technology, computing devices and high-speed internet, in order to participate fully in society. Access to the internet and technology has been largely normalised but remains inaccessible for some.
- Prices have risen for the costs of common recreational activities such as eating at restaurants or ordering takeaway food. Fewer families now engage in these kinds of activities on a weekly basis.
- Low-cost private renting is now considered the minimum standard for childless couples and single persons. Social rents for smaller properties rose by 18 per cent between 2011 and 2014, and private rents outside London rose by 2.6 per cent (Davis, et al., 2014: 27). However, due to differences in energy efficiency between social and private rental properties, this was associated with a 25 per cent increase in fuel costs.

(Davis, et al., 2014: 4)

The cost of achieving a good quality of living is therefore not stable over time or necessarily in line with inflation. Wage and benefit growth has lagged significantly behind price inflation; especially for consensually agreed minimum essentials, the cost

of which has increased at a faster rate than the consumer price index (CPI) as a whole. Even when just considering CPI inflation of 19 per cent between 2008 and 2014, earnings have only risen by 9 per cent in the same time period (*ibid*: 39). Benefits growth has been markedly slower, around 1% per year (Sutherland, et al., 2008), as robust uprating of benefits has not been pursued in British social policy – where uprating policies are in place at all. Current uprating policies estimated over a twenty year period from 2006/7 with assumed annual earnings growth of 2 per cent suggests that there would be a fall in real income percentage of 17 per cent for the poorest quintile of households, and regressive income falls for all subsequent households moving up the income distribution, as income becomes redistributed from the poorest and middle earners to the richest (Sutherland, et al., 2008: 62). Despite Townsend's (1962) insistance that the concept of poverty is dynamic, relative poverty measures curiously remain fixed. Fixed threshold based measures therefore may not accurately reflect a true measure of poverty and any subsequently derived poverty rates that are comparable between years: although the method may stay the same, the reality may deviate considerably.

Likewise, targeted means-tested benefits that are often based on these measures are likely to result in under-coverage, where 'someone who deserves the benefits is denied them' (Mkandawire, 2005: 9). Townsend (2010 [2002]: 84) argues that such policies to target welfare on only the poorest have 'increased inequality and perpetuated poverty'. If the HBAI line, in its current form, is inadequate for consistently measuring the concept of poverty in a single country over time, it is difficult to argue that it is also an appropriate measure for making international comparisons of poverty rates or understanding poverty in quantitative sociological research; a more nuanced indicator of poverty is needed than the HBAI line. Typically, such more nuanced studies (for example, Tomlinson, Walker & Williams' 2008 multidimensional approach) do not support any clear 'cut-off' or threshold and might instead support more universal or tiered provisions; however, in considering the development of welfare policies Townsend (2010 [2002]) identifies that once evidence begins to mount against existing targeted and neoliberal policies they are frequently only reconceived under different guises. A well-documented example of this being the continued re-dressing of 'troubled families' style programmes since the post-war period in Britain, policies that perpetuate a narrative of feckless underclass families that disproportionately rely on public funds

(Lambert & Crossley, 2016), another being the perpetuation of families with 'generations' of worklessness who require intensive, harsh, targeted interventions (MacDonald, et al., 2013).

2.4.2 Comparing Measures: The Minimum Income Standard

In contrast, the minimum income standard (MIS) is a consensual attitudinal approach designed to adjust for these changes while involving members of society in determining a minimum acceptable level 'that is socially unacceptable for *any* individual to live below' (Bradshaw, et al., 2008: 4, *emphasis in original*). Contemporary consensual measures have been developed from original critiques of survey techniques to establish attitudinal definitions of poverty that had relied on the judgements of experts rather than societally agreed standards (Walker, 1987). Robert Walker (1987) proposed an alternative approach of democratising the process of poverty definitions based on sustained dialogue between experts and respondents from the public, exchanging knowledge and views, which allows for an informed, nuanced definition 'by consent or compromise' (1987: 221):

> *"…to be true to the consensual approach, people must be given the scope to express their views[…] Researchers must equally be prepared to listen to their respondents and to work with their 'real-world' concepts[...] Opinions grounded in ignorance, while interesting in themselves… have little utility as a basis for policy not least because they are likely to be very unstable. […] Researchers are therefore obliged to provide respondents with the information which they need in order to make reasoned choices and, as far as possible, to provide feedback on the consequences of the choices made."*
>
> Walker, Consensual Approaches to the Definition of Poverty, 1987: 222

This approach has been used by the JRF to establish a minimum income standard since 2008, with annual updates and full 'rebasing' of calculations for each household type every four years (Bradshaw, et al., 2008; Centre for Research in Social Policy, N.D.: online). This standard is set through consultation with socioeconomically diverse members of the public and experts to determine a minimum 'basket' of goods and services required to 'participate in society', clearly reflecting the consensual method and Townsend's (1979) conceptualisation of poverty (Hirsch, 2015: 7-8). This standard is adjusted yearly for inflation and the contents of budgets are reviewed biannually between rebasing (*ibid*: 7). This means that the minimum income standard is adjusted frequently in order to account for changes not only in the prices of goods and services,

changes in wages and inflation, but also in response to informed public perceptions of what constututes a morally unacceptable standard for people to fall below. The detail and nuance here, therefore, is a stark contrast to the HBAI measure.

Considering this had been the original intention of the HBAI line, the gap between the poverty thresholds of the two is meaningful when considering the ways in which poverty is measured and conceptualised (Table 1). The original report found that MIS thesholds before housing costs for each household type were all higher than the 60 per cent of median income line (Bradshaw, et al., 2008: 37). The most recent comparison between the HBAI and MIS suggests that this gap has widen further. This demonstrates that the cost of meeting a minimum standard of living has risen over time relative to income, and therefore the dominant HBAI line is likely to further underestimate the prevalence of relative poverty in the UK, if it is accepted that the methodology of the MIS is closer to accurately reflecting the concept of poverty than the methodology of the HBAI.

This has real consequences for the targeted design of welfare in the UK and our understanding of poverty and the experiences of those in poverty. If poverty alleviation policies are largely targeted to those below arbitrary HBAI lines, this means that attempts to bring people out of poverty (in a real sense, not just a statistical sense) are likely to have high levels of what Thandika Mkandawire refers to as type I errors of underpayment (2005: 9). Similarly, if the effectiveness of policies designed to alleviate poverty are judged based on changes in poverty rates based on the HBAI the conclusions reached may be incorrect, reflecting the tendencies of the measure more than the effects of the policy interventions. Depending on the measure and conceptualisation of poverty used, the proportion of households in poverty, and subsequent policy reactions, change significantly.

Table 1: Comparisons of the percentage of median income required to meet the minimum income standard, compared to unchanging HBAI measure.

	Proportion of median income required to be 'not poor' (%)	
Measure	2008*	2015**
Households Below Average Income BHC	60	60
Minimum Income Standard BHC (Working Age Household)	74	86
Minimum Income Standard BHC (Pensioner Couple)	63	67
Minimum Income Standard BHC (Couple with two or more dependent children)	76	85
Minimum Income Standard BHC (Lone Parent with one or more dependent children)	75	87

Sources: * Bradshaw, et al. (2008: 37) ** Hirsch, (2015: 17)
BHC = Before Housing Costs

2.4.3 Comparing Measures: Why measurement matters

This is evident in Table 2, which shows the proportion of the population considered to be in poverty by each measure since 2008/09. With the weaknesses of the distinctions between absolute and relative definitions in mind, the comparison between the measures shows, firstly, that there is a significant gap between the proportion of the population in "absolute" poverty, measured by destitution in 2014/15, severe poverty rates from 2008/09 to 2012/13, and severe material deprivation from 2008/09 to 2011/12, and the proportion of the population in relative poverty, measured by the HBAI before housing costs (BHC) and after housing costs (AHC) lines and the Joseph Rowntree Foundation's minimum income standard (MIS) between 2008/09 and 2013/14. Estimates of relative poverty are in excess of four times the rate of severe material deprivation and around ten times the magnitude of severe poverty and destitution; however, these rates of extreme poverty are far from negligible.

These differences demonstrate the importance of a clear ontology of poverty and problematise approaches that uncritically conflate absolute and relative measures or use the term 'poverty' in empirical study without clearly outlining the assumptions their specific measurement lies upon. Another feature of this comparison, if assuming destitution rates were similar prior to the 2015 study, is the large gap between these estimates of absolute poverty. This is likely because of the uncritical methodological approach of the ONS' severe material deprivation measure, where a person can be counted as 'not in absolute poverty' if they can afford a telephone, television, car and washing machine, goods which have become increasingly more affordable and variable in quality since the measure was implemented, but not nutritious food, warmth or shelter, the costs of which have outstripped income growth, without going into debt.

Table 2: Comparison of estimates of the proportion of population below different poverty lines.

	Proportion of Population in Poverty Per Year (%)						
Measure	'08/9	'09/10	'10/11	'11/12	'12/13	'13/14	'14/15
JRF Destitution (Absolute Poverty) *							2.4
Severe Poverty (Absolute Poverty, BHPS) *	0.9	2.45	2.4	2.3	2.3		
Severe Material Deprivation (Absolute Poverty) ** [1]	5.1	5.1	5.1	5.1			
Household Below Average Income (Relative Poverty) BHC ***	18	17	16	16	15	15	
Household Below Average Income (Relative Poverty) AHC ***	22	22	21	21	21	21	
Minimum Income Standard (Consensual Relative Poverty) ****	23.7	24.2	24.8	27.2	29.3	29.6	

Sources: * Fitzpatrick, et al. (2016: 19, 26) ** Office for National Statistics (2013: 1) *** Shale, et al. (2015: 39) **** Padley & Hirsch (2016: 12)

[1] The report on severe material deprivation by the Office for National Statistics found that the 'UK's severe material deprivation rate is broadly unchanged since 2005 when comparable figures were first produced' (ONS, 2013: 1). This rate was 5.1 in 2011, but the ONS do not give the specific proportions or sources for the preceding years.

Robust studies of absolute poverty in the UK are lacking in coverage, despite their importance as an integral component to understanding poverty as basic social and physiological needs (Lister, 2004: 27-28). Indeed, no data for the proportion of the UK living below the World Bank's globally comparable Purchasing Power Parity (PPP) line, covering 1981 to 2015, has been published by the institution (World Bank, 2016: online). This perhaps reflects Spicker's (2007: 11) observation of publicly held claims that " 'poverty' does not seem to mean the same thing in advanced economic countries as it does in developing ones". It is equally strange in light of the fact that in John Hills' (2001: 4) analysis of responses to British Social Attitudes data in the years 1986, 1989, 1994 & 2000, only between 25 to 28 per cent of respondents agreed that 'enough to buy the things [people] really needed, but not enough to buy the things most people take for granted' constituted poverty compared to between 55 & 60 per cent who agreed having only enough to buy things necessary to survival but not enough to buy other things that are needed constituted poverty, and between 90 and 95 per cent who felt not having enough to eat and live without getting into debt is considered poverty. This suggests people feel absolute poverty is a more valid conceptualisation than relative poverty in lay understanding, reflecting Walker's (1987) warning that opinions without information may not be valuable policy bases, and by extension, may not be helpful conceptualisations for research. What attempts there have been to establish the rates of absolute poverty in the UK, especially the JRF's study of destitution are few and far between.

Comparisons between relative measures demonstrate that two things matter in moving accurately from concept to measurement: firstly, the importance of the conceptualisation of poverty adopted and the degree of sophistication applied to its operationalisation; and secondly, the effects of adjustments in publicly acceptable minimum standards and the changing costs of achieving these. As previously discussed, the public and expert consultations used in the MIS' more sophisticated methodology create a line that is set at a greater threshold of median income than the HBAI line,

which is reflected in higher estimates of the proportion of the population living in relative poverty. What is more striking is how the trajectory of poverty rates differs between measurements. Due to the combination of the JRF's rebasing of minimum standards through consultation, but largely due to more frequent and sensitive yearly adjustments for price and wage inflation, a very different trajectory of poverty emerges to that of the HBAI lines. Whereas HBAI lines show mostly stable poverty rates with some decline, MIS rates show a steep increase in poverty over the same six years. These differences strongly suggest that the same underlying concept of poverty does not appear to be being approximated across all measures, and any subsequent choice between measures is likely to reflect certain political standpoints (Tomlinson, et al., 2010). One explanation for this may be that a single, sometimes volatile, proxy may not be a reliable reflection of the complex, shifting, idea of and state of poverty (Tomlinson & Walker, 2009).

This is what Joel Best calls the problem of 'unchanging measures', where 'changes in the larger society make an established measure less accurate or less meaningful' (Best, 2012: 103). This is particularly the case for measures that are not assessed statistically for their reliability and validity (Millsap, 2011). Often, this issue is compounded when non-academic interests play a role in the plasticity of the measures. Best (2012) uses the example of the US federal Bureau of Labor's measure of productivity and its link to wages, where management interests opposed any adjustments to the measure that would result in higher productivity being reported, and trade unions opposed any adjustments that would result in the opposite. A similar case might be argued for the influence of political interests on the measure of poverty; where governments can have an incentive to report stable or decreasing poverty rates and opposition and activists may have an incentive to report rising poverty rates. Many of these politically sensitive approaches to measurement are, nonetheless, used in academic scholarship. If political interests outweigh academic interests, a measure might reasonably become an unchanging measure. A way out of this impasse is to ensure that measurements accurately and robustly reflect their theoretical conceptualisations, with as few caveats as possible; one of the greatest weaknesses of existing measures is that they fail to do justice to more than one or, at most, two of the many dimensions of poverty.

2.5.1 Multidimensional Poverty

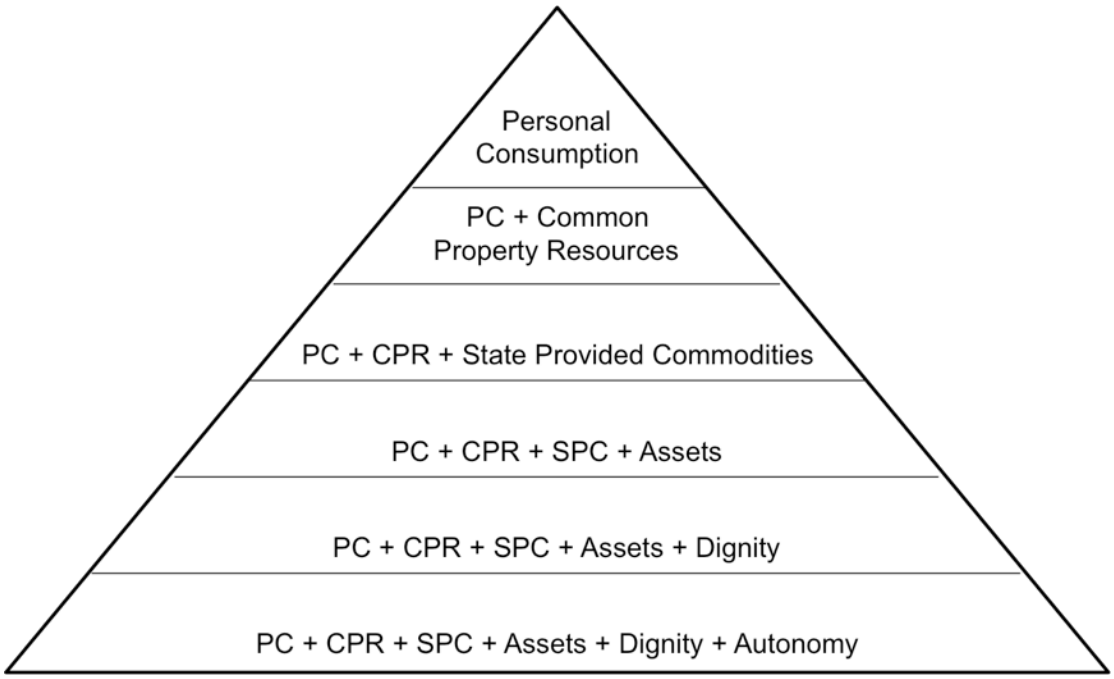
The ideas of poverty explored here are clearly multidimensional in approach, whereas measures of the extent of poverty have often focused on income-based approximations of this multidimensional idea, partly due to methodological limitations discussed in the next chapter and partly because of the the benefits of easier interpretation and temporal and geographical comparison. The complexity of poverty is reduced to a singular dimension of measurement. However, it is also clear that some dimensions of poverty are not so strongly connected with income, and some are more tenuously linked than others. In discussions of existing measures this multidimensionality has been touched upon – Townsend and Sen discuss the multiplicitous nature of participation and capability, and at the core of Seebohm Rowntree's (1901) approach was the combination of physiological and social basic needs - but these dimensions have largely been limited to the extent to which they can be measured by consumption and material deprivation and costed against prices and income. Complex immaterial dimensions, however, are well documented within qualitative literature on the experience of poverty (for example, Beresford, et al., 1999).

Furthermore, these unchanging measures go against the time-sensitive dynamics of poverty that are so important to scholars' sociological understanding of poverty. Poverty is described not only as a dynamic, rather than persistent, event or state of being in peoples' lives that can be transitioned into and out of through the lifecourse (Hills, 2015; Tomlinson & Walker, 2010), but as a dynamic concept in itself. Although poverty can have a fixed ontological definition, the ways in which in manifests and is subsequently epistemically justified change relative to culture and context. In other words, poverty does not mean the same experiences regardless of context, even though the root of those experiences may unchanging. In Townsend's (1962) own words:

"*Poverty is a dynamic, not a static, concept. … there is no list of the absolute necessities of life to maintain even physical efficiency or health which applies at any time and in any society, without reference to the structure, organization, physical environment, and available resources of society."*

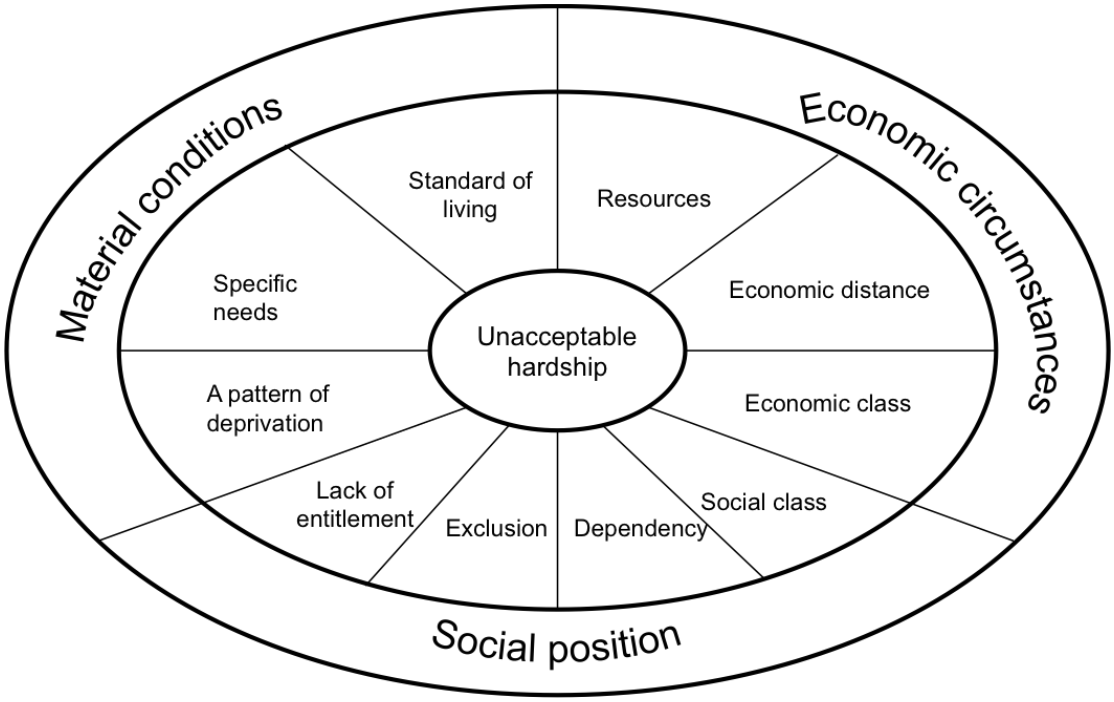
Townsend, The Meaning of Poverty (1962: 219)

Figure 1: Bob Baulch's pyramid of poverty concepts



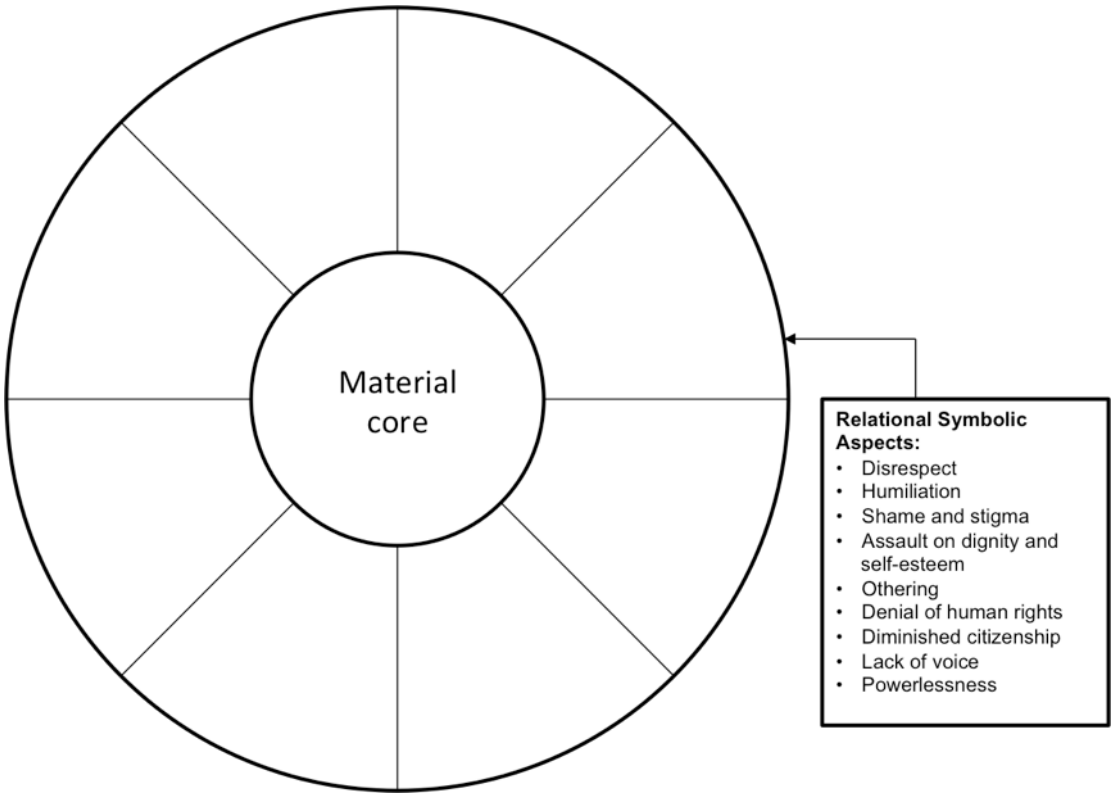
Adapted from Baulch, 1996: 2

Figure 2: Paul Spicker's clusters of concepts of poverty



Adapted from Spicker, 2007: 6

Figure 3: Ruth Lister's material and non-material wheel of poverty



Adapted from Lister, 2004: 8

2.5.2 Multidimensional Poverty: Baulch's Pyramid, Spicker's Clusters & Lister's Wheel

Figures 1 through 3 show various conceptualisations of multidimensional poverty from Bob Baulch, Paul Spicker and Ruth Lister respectively (Baulch, 1996: 2; Spicker, 2007: 6; Lister, 2004: 8). These conceptualisations aim to represent poverty as is experienced by people in participatory assessments and other in-depth qualitative research (Beresford, et al., 1999; Lister, 2004). All build on a material apex or core component, with intrinsically linked material and non-material aspects. As the concept of poverty widens to become more inclusive of the circumstances, experiences and conditions of people living in poverty beyond consumption, this singular central core or apex becomes a less suitable way to encapsulate the concept. This movement from the tangible to the complex and intangible reflects many of the difficulties inherent in the operationalisation of the concept.

Baulch's (1996) pyramid stems from discussions around the inadequacy of existing concepts of poverty as necessitated by restrictions of how to operationalise the concept. The restrictions on definitions imposed by operationalisation become problematic as broader and more accurate understandings of poverty are ignored, and restricted, typically, to the apex of Baulch's pyramid. As Ruth Lister (2004) warns:

> *'To move straight to definitions and measures without first considering the broader concepts can result in losing sight of wider meanings and their implications for definitions and measures. In particular, it can exclude the understandings of poverty derived from qualitative and participatory approaches. These frequently highlight aspects of poverty that lie outside definitions focused on income and material living standards and that can be difficult to measure in surveys[…] starting at the bottom with measures can encourage confusion between measures and definitions, so that arguments about competing definitions of poverty often turn out to be arguments about competing measures. […] omitting the conceptual level can encourage a myopic, technocratic approach that, in its preoccupation with measuring poverty's extent and depth, overlooks how it is experienced and understood.'*
>
> *Lister, Poverty (2004: 5-6)*

Measurements, therefore, often inadvertently become definitions of poverty when, in actuality, they act as only technologically and cognitively limited approximations of a multifaceted concept.

Beyond the factors commonly associated with income: personal consumption, common property resources, and state provided commodities, the pyramid extends to encompass assets, dignity and autonomy, although each of these should not be considered 'strictly additive' (Tomlinson, et al., 2008: 599). Here, assets refer to 'human and physical capital, stores, and claims' that can determine the rate at which people 'reach the threshold of collapse' (Baulch, 1996: 2-3). Physical capital here might refer to financial assets, typically wealth, in the form of savings or illiquid assets in terms of possessions, often homes, but also items such as jewellery, that can be drawn upon to support personal consumption in the case of low income; these are largely concentrated in those at the higher ends of income distributions (Hills, 2015).

This is reflected in the economic cluster and closely related social class dimensions of Spicker's (2007) concepts. Human capital refers to embodied forms of capital accrued through training, education and healthcare, which in turn can yield income and other fiscal capital outputs over time – this is perhaps most obvious in the relationship

between education and earnings (Becker, 1993). Unlike financial capital and wealth, transmission of human capital is more complicated, as 'the heir must make some effort' (Picketty, 2014: 420). Human capital, however, also has value outside of its effects on fiscal capital. Human capital essentially reflects a persons participation in human knowledge accumulation and production, culture, and in shaping development and advancement. A person's human capital is therefore important independent of its effects on income. Deprivations of human capital could, therefore, be considered a form of poverty despite an adequate financial position, if accepting Townsend's argument that poverty is the inability to participate fully in society.

Dignity, in Baulch's pyramid, refers to freedom from tasks and activities that 'are regarded as subservient' (Baulch, 1996: 3). This experience of poverty is reflected in Spicker's (2007) groupings of social position, including lack of entitlement, exclusion, dependency and social class. A more nuanced overview of deprivations of dignity as a feature of poverty is found in Lister's (2004) relational symbolic aspects of poverty, which bring attention to the experiences of disrespect, shame and stigma, low self-esteem and powerlessness. Negative associations are often made between those living in poverty and undesirable traits. Traits such as fecklessness and idleness are ascribed to those living in poverty universally in an 'Othering' process, resulting in often internalised stigma and dehumanisation that results in an assault on poor peoples' dignity regardless of individual merits (Beresford, et al., 1999; Lister, 2004).

This process of 'othering', creating a 'them' and 'us' between the poor and non-poor, is considered a construct of the dominant 'non-poor', imposed through certain institutions, especially in the case of the media (Lister, 2004). This is further reinforced by historical perceptions of a 'deserving' and 'undeserving' poor; a moral divide documented in the United Kingdom and United States especially, drawn between the 'undeserving' 'able-bodied' but irresponsible and lazy 'pauper' and the 'deserving' 'impotent or incapable', infirmed 'pauper' (Lister, 2004: 105). These classifications go so far as to suggest the existence of an 'underclass' of people characterised by elective exclusion from society and persistent dependency on structures of welfare (Lister, 2004; Alcock, 2006: 27-29). Despite little sociological evidence of such an underclass the idea frequently permeates British political and popular discourse, and therefore affects the dignity of and respect towards those living in poverty (for examples, see

Morris, 1995; Macdonald, et al., 2013, Macmillan, 2010 and Shildrick, et al., 2012, who test the claim of attitudes towards and prevalence of intergenerational worklessness among the poor; or Hills, 2015, for evidence on the high turnover of benefit claims and very low rates of dependency on prolonged welfare).

However, the idea that othering processes are imposed purely by the non-poor is problematic, especially as people as individuals have their own definitions and understandings of poverty, as alluded to earlier in discussions around what constitutes poverty in developed countries and as evidenced in Jan Flaherty's (2008) and Tracy Shildrick and Robert MacDonald's (2013) research into the rejection of the label of poverty by those living below established poverty lines. In this case, othering and stigmatisation is performed not only by the indisputably non-poor, but also the poor who self-identify as non-poor and distance themselves from a likely nonexistent 'underclass' of people.

Experiences of stigma are also evident in peoples' perceptions and feelings of claiming benefits in order to subsidise low income in periods of involuntary economic inactivity or in the case of chronic low-wages. Ben Baumberg (2016: 181) reports survey data that suggests 34 per cent of people feel that there is personal stigma (stigmatised by themselves) or stigmatisation (perceived stigma from others) attached to receiving at least one form of benefits; and this figure roughly matches reports of stigma from claimants. Over one in four respondents reported that a stigma-related reason reduces their likelihood of claiming welfare support they are entitled to (*ibid*). Townsend (1979: 841) referred to this as the shame of 'pleading poverty'.

There are material consequences of such stigma and othering that are likely to exacerbate peoples' risk of poverty. Recent figured from the Department for Work and Pensions estimate that over £13 billion of pension credit, employment support allowance, housing benefit and income-based Job Seekers Allowance benefits that households and individuals were entitled to went unclaimed in the year 2014/15 (DWP, 2016: 3). The lowest take-up was of Job Seekers Allowance (DWP, 2016: 1), a benefit with higher reported stigma, especially amongst vulnerable groups (Patrick, 2011; Friedli & Stearn, 2015; Milton, et al., 2015; Baumberg, 2016). Even this estimate ignores other kinds of benefits that go unclaimed. Although robust investigation into

the uptake of benefits is lacking, a report for an independent organisation of financial advisors, named Unbiased, estimated that £2.2 billion of child tax credits and £2.7 billion worth of working tax credits went unclaimed in the 2012-13 financial year (Wall, 2012). Increasing conditionality across previously universal benefits is likely to exacerbate these figures (Simmons, 2011), and might possibly be deliberate as a form of 'stealth' cut to the welfare bill (Gamble, 2016). Such a cyclical relationship between resources, assistance and stigma is a distinct experience of poverty in the UK and other countries where the view of many people in poverty was that they were more likely to be 'treated with contempt' than with respect during their engagement with public services, especially those related to employment (Walker, 2014: 151-153). As such examples demonstrate, loss of dignity, a feature of poverty not tenably linked with income, is a substantial part of the experience of poverty as a whole.

This leads to the metaphorical foundation of Baulch's pyramid: autonomy, the ability to 'choose self-fulfilling and rewarding life-styles' (Baulch, 1996: 3). This dimension strongly reflects Amartya Sen's (1995) ideas of capabilities as freedoms; the presence of attainable options and alternatives to fulfil a good quality of life, and access to the strategies to achieve them (Doyal & Gough, 1984). Poverty restricts this autonomy both financially, by barring access to opportunities and strategies that require money or resources, and socially, in the complex ways in which structures like social class operate to exclude certain groups from specific activities or from employing specific strategies to attain these (Alcock, et al., 2012). Limited income and assests can preclude children and adults from building sufficient human capital, which restricts opportunities to access future sources of income (Tomlinson & Walker, 2009; Hills, 2015), as can subsequent debt incurred from accessing such opportunities in lieu of sufficient income.

As a consequence of being unable to attain socially and materially essential commodities, rights, dignity, and autonomy, the resultant shame people in poverty often experience manifests as a universal psychosocial experience which extends to those who may not live in poverty but witness it around them; the importance of this shame should not be ignored due to its intangibility (Walker, 2014). Walker (2014: 65-66) incorporates this inexorable link within a theoretical model of the 'poverty-shame' nexus, which demonstrates the way in which shaming, as a process, operates on a societal-relational level and is institutionalised through stigma as a policy tool to

'discourage' poverty – treating it as a result of an individual's choices. These mechanisms of shame, and the experience of poverty itself, manifests itself in the shame and sense of failure people feel, exacerbated for those living in poverty, which then results in social exclusion, low self-worth, a lack of agency and low social capital. There is a reinforced relationship between exclusion and low self worth, resulting in a lack of volition and a restricted ability to form beneficial relationships with others, which mimic clinical indicators of psychological distress. These factors cycle back to compound a person's poverty (*ibid*).

2.6. What really lies at the heart of poverty?

Although most conceptualisations of poverty identify unacceptable material hardship as the core of poverty, there has recently been internationally comparative research that suggests there is also a fundamental socio-psycho-relational core of poverty that coexists alongside material hardship. Walker (2014) identifies this experience as shame, and he and his colleagues' international work suggests that it may not be a peripheral characteristic of poverty but a central and universal one that holds a less priveleged epistemic claim. Robert Walker (2014) uses the example of the comparison between academic or policy, and creative arts, approaches to defining poverty:

> "*While the mainstream academic and policy debate has focussed on lack of income as the defining feature of poverty… creative writers and film-makers appreciate that, for many individuals, the hardships engendered by poverty are as much social and emotional as they are material. People in poverty feel themselves to be despised by others and come to despise themselves."*
>
> Robert Walker, *Conceiving of Poverty without Shame*, 2014: 85

Most notably, Walker (2014) argues that people can feel shame in the absence of explicit shaming: it is a societally reinforced consequence of poverty. Support for such a conclusion however can also be drawn when shifting the lens from the position of those being shamed to those doing the shaming. The work of Tracy Shildrick (2018) and Robert MacDonald (Shildrick & MacDonald, 2013) highlights the way that the nature and implications of poverty enable a discourse whereby people distance themselves from their own poverty by the mechanism of shaming others, whose situation or response to a situation may be interpreted very differently. This discourse is frequently amplified by the media.

Stepping back from measurements and focusing on theoretical conceptualisations of need and poverty more broadly shows that actually the centrality of this relational aspect has been present all along. Compare, for example, Baulch's (1996) pyramid of poverty and Maslow's (1943) hierarchy of needs. The first can be imagined as a conceptualisation of needs from the standpoint of a policy-maker or academic, where income and consumption take the top positions of the hierarchy and social and emotional factors like dignity are found towards the bottom of the hierarchy. In contrast, Maslow's (1943) hierarchy focuses primarily on emotional and psychological needs. Self-actualisation and esteem take the apex of the pyramid here, and subsistence needs based on consumption are found further down the hierarchy. The two are not incompatible. In contrast, they differ in their construction only by the perspective of the person doing the ordering, as has been found in international validation of Maslow's hierarchy (Tay & Diener, 2011). Would it not therefore be possible that poverty actually has two 'cores': one which represents the material, and one which represents the social and emotional?

2.7. From Multidimensional Concepts to Measurement

How then can such a nuanced reality of poverty be operationalised within a quantitative framework? Walker (2014: 22) explains that attempts to do so have been severely constrained by lack of data, and disincentivised by their limited internationally comparative capacity in favour of more simple measures. To an extent, capabilities or participation activities with clear 'price tags' can be integrated into income-based measures by calculating the cost of attaining these as an addition to a 'basket' of goods and services, but even this is problematic. Income is only partially correlated with many dimensions of poverty, more highly with consumption, less with material deprivation and less still with dimensions such as shame or social exclusion, despite these factors being hugely stressed as central to the experience by people living in poverty (ATD Fourth World, N.D.). Therefore costed lists of items, no matter how complex, should not be considered an unproblematic and unimprovable approximation of poverty, regardless of the complex methodologies used to determine the costs of 'baskets' or 'budgets' (Townsend, 1962; Calandrino, 2003; Tomlinson, 2008; Tomlinson, et al., 2010).

Why should we take these deficiencies seriously rather than accept the limits of methodological convention? As Tomlinson, et al., (2008: 600, emphasis added) explain, '[t]he partial measures used to date necessarily *fail adequately to do justice to the experience of poor people* and, to the extent that they distort through omission, may result in implementation of inappropriate policies'. These aspects of poverty can be considered oversimplified through approximation or are simply ignored in some cases, for example, where assigning monetary value to dimensions would be impossible, very difficult or illogical. How can a researcher possibly calculate monetary equivalents for civic participation or voice (Lister, 2004: 165-168)? The weighting of each dimension of poverty, deciding which dimensions matter more or less than the others, is also often uncritical (Tomlinson, et al., 2008). Inevitably, then, dimensions associated more strongly with income are given undue prominence in the operationalisation of poverty compared to those that may be less directly linked to purchasing power but no less difficult to obtain. These may be considered more important features by those experiencing poverty but may have very little bearing on poverty measurement (Beresford, et al., 1999). A futher complication is that combining several measures in order to try and create multidimensional indices through traditional means exacerbates error in the final index based on the error in the measurement of each individual component (Tomlinson, et al., 2008).

Using new statistical techniques, Tomlinson, et al., (2008) modelled poverty as a multidimensional concept over the 1990s and early 2000s. With structural equation modelling (SEM), the authors created a higher-order latent poverty index based on financial pressure, as modelled by financial strain and material deprivation; social isolation; civic participation; psychosocial strain, modelled by low confidence, social dysfunction and anxiety/depression; and, with newly integrated housing and neighbourhood variables from 1997 onwards, environment (*ibid:* 604-608). Using this index they estimate the proportion of people in poverty based on relative and fixed thresholds. These are shown in comparison with HBAI estimates in table 3. Similar to the earlier observation of differences between simpler measures of poverty and the more complex MIS and destitution measures, different patterns emerge between the thresholds. While the relative HBAI line shows very stable poverty rates over the years, with some decline, the relative poverty index scores show a strong decline year on year. With the poverty index rates set at 25 per cent in 1991, equivalent to absolute poverty

rates by income measures, the index also shows a multidimensional estimation of falling absolute poverty. However, Tomlinson, et al. (2008) do not report any reliability, validity, or invariance testing of their constructs, which means that accurate comparisons over time and between groups may not be appropriate.

This, and an associated analysis by Emanuele Ferragina, Mark Tomlinson & Robert Walker (2013), also raise important questions about the way in which poverty is imagined as a binary concept. This idea has been constant throughout this review, and some of the problematic elements of this have been touched on in reference to targeted welfare benefits. When considering the experiences of people living in poverty, and the stark differences between the lines drawn by each measure, it seems obvious that any 'cliff-edge' grouping is problematic as poverty is experienced on a sliding scale (Gordon, 2006). This is clear from Tomlinson, et al.'s, (2008: 613) comparison of multidimensional poverty index scores across income deciles. Rather than a steep increase in poverty index scores at a certain point of low-income, scores decrease steadily as income increases.

Similarly, in their analysis investigating a 'breakpoint' in participation when income is below a certain point, a theory expressed by Townsend (1979), Ferragina, et al. (2013: 8-15) find, rather than a breakpoint, a floor of low participation scores at lower income groups followed by steady increases in participation as income increases, with a sharp increase in social participation once a certain level of income is reached. In light of this, splitting people into poor and non-poor therefore seems counter-intuitive to sociological theory, and potentially marginalises and omits the experiences of dimensions of poverty and deprivation of those with middling-incomes or moderate experiences.

Table 3: Comparison of relative poverty rates in the United Kingdom when measured as a multidimensional concept compared to Households Below Average Income criteria.

	Proportion of Population in Poverty in Year (%)												
Measure	'91/92	'92/93	'93/94	'94/95	'95/96	'96/97	'97/98	'98/99	'99/00	'00/01	'01/02	'02/03	'03/04
HBAI BHC (less than 60% of Median Income before housing costs)*				19	18	19	20	19	19	19	18	18	17
HBAI AHC (less than 60% of Median Income after housing costs)*				24	24	25	24	24	24	23	23	22	21
Less than 80% of Median Multidimensional Poverty Index Score **	13.5		12.6		11.6		11.5		10.9		10.2		9.7
Less than 85% of Median Multidimensional Poverty Index Score **	19.2		18.3		17.4		17.2		16.7		15.4		14.8
Less than 90% of Median Multidimensional Poverty Index Score **	27.5		26.8		26.3		26.1		25.0		23.8		23.0
Change from fixed 1991 25% line in Multidimensional Poverty Index **	25		23		20		17		16		14		12
Change from fixed 1997 25% line in Multidimensional Poverty Index including Environmental conditions **							25		22		19		17

Sources: * Shale, et al. (2015: 39) ** Tomlinson, et al. (2008: 610-612)

The approach also limits how much understanding of poverty can be gleaned from such analyses: if poverty is clearly multidimensional, and clearly experienced to differing degrees, it seems wasteful in a statistical and real sense to impose such a threshold (as beneficial as it may be in the case of targeted policy design). As will be discussed in a later section, experiences of recurrent poverty over time further problematize the binary approach. This has lead to some experts to focus away from poverty, and the attempts to draw such a line between poor and non-poor, to focus on concepts with greater emphasis on continuous scales of experience and exposure such as social quality & wellbeing (van der Maesen & Walker, 2012). In order to better understand poverty, this research aims to investigate poverty as a multidimensional concept where people face diverse continuums of deprivation.

2.8 Social divisions and poverty

Thus far, these discussions of poverty have assumed that poverty is experienced equally across people of all characteristics. A reliable and valid measure of poverty should be applicable to all groups equally, or should have clear adjustable parameters to better reflect the experience of other groups. This section briefly explores the social divisions that emerge in existing studies of poverty. As the UK is a western capitalist advanced economic society that stresses meritocracy and equity of opportunity as the defining predeterminants of success, there should, in theory, be no reason why arbitrary characteristics unrelated to merit such as gender, age, ethnic origin, or disability should determine a person's chances of experiencing poverty. Obviously this has not been the case in the UK or similar countries and there are great and substantial social divisions in the risk of poverty (Alcock, 2006; Lister, 2004). Women, children, the elderly, certain household structures, people with non-White heritage, the disabled and their carers have historically been at greater risk of poverty than others (Alcock, 2006).

2.8.1 Women's experiences of Poverty

Women experience a slightly higher incidence of poverty as measured by the HBAI metric than men (MacInnes, et al., 2013), however, the risks of poverty for women are likely to be underestimated due to assumptions around intra-household distribution of income and divisions of labour (Pantazis & Ruspini, 2006). Claims that there is an obfuscation of women's risks and experiences of poverty have centred around the critique that income based measures of poverty often assume that women and men receive equal access and derive equal benefit from household resources (Millar & Glendinning, 1989). For example, the idea that women who are the wives or partners of non-poor men are also not poor is based on the assumption

that household income is split evenly between partners, an assumption that is made of all family structures, but may not be the case. Existing research has argued that there is not equitable access to household resources between couples; money is not always equally split between 'the purse and the wallet' (Pahl, 1989; Goode, et al., 1998; Bradshaw, et al., 2003; Burton, et al, 2007; Bennett, 2008).

Sophie Ponthieux (2013: 18) estimates that the assumption of equally shared income between partners does not hold when considering equally shared income would be completely pooled, for approximately 30 per cent of households in Europe, with this rate being higher, around 35 per cent, in the UK. Bennett & Daly (2014: 36) claim that this may lead to "'hidden' poverty'", that is obscured by the assumption of equal access to household income (Kempson, et al., 1994; Goode, et al., 1998; Bradshaw, et al., 2003; Alcock, 2006).

Although not always the case, women and women's finances in particular are often expected to fulfil a greater number of domestic roles within the home than men's, all whilst their earning status and role is seen as secondary, resulting in greater risks of poverty in unequal households (Alcock, 2006). Money management and the gendered division of care work are central to these gendered experiences of poverty (Alcock, 2006). Women in low income households are more likely to be responsible for managing budgets for household spending, which encompasses goods such as food & groceries, fuel, clothes and children's school-related goods (Kempson, et al., 1994; Goode, et al., 1998; Maplethorpe, et al., 2010).

There are numerous examples documented in research where shortfalls in household budgets are often made up from women's personal incomes, or through women depriving themselves of food, fuel (in heating or transport), to make ends meet for the benefit of other family members as an act of what Land and Rose (1985: 86) call 'compulsory altruism' (Kempson, et al., 1994; Graham, 1986; Graham, 1992; Pantazis & Ruspini, 2006; Burchardt & Karagiannaki, 2018). Moreover, the effect of this unpaid work is cyclical as it restricts women's capacities to undertake paid work (Shildrick, et al., 2012). This unequal behaviour of self-sacrifice is far from consigned to history as, in a period of prolonged austerity and resultant food poverty, despite the wealth of the UK, women in low-income households continue to go without meals and warmth (Dowler & Lambie-Mumford, 2015) and are more likely to invest any excess income into the human capital of their partner or children (Alcock, et al., 2012: 50). These

deprivations are fundamentally linked to societal expectations of women being responsible for caregiving, and the stigmatising costs of breaching such expectations (Finch & Groves, 1983).

These stigmatising and non-material dimensions form core components of experiences of poverty for women. Already, gendered differences in household income distribution and budgeting responsibilities allude to dependency and diminished autonomy and respect that becomes reinforced by gendered societal structures. Central to this is the idea that these responsibilities also result in deprivations in free time, or 'time poverty', due to the vast amount of unpaid work undertaken by women; this is especially the case for employed women who often face a triple burden of paid work, unpaid care and unpaid emotional work on their time (Oakley, 1976; Lister, 2004; Chatzitheochari & Arber, 2012; Burchardt & Karagiannaki, 2018). This raises a question of whether women's experiences of poverty are more centred around dependency and lack of, or constrained, autonomy which can further heighten the powerlessness, exclusion, and diminished dignity experienced by those in poverty (Lister, 2004; Alcock, 2006).

However, in terms of gendered differences in what constitutes poverty, Pantazis, Gordon & Townsend (2006: 98-100) found there was a strong consensus between men and women in what they deem necessities of life ($R^2 = .98$). Where there was less consensus between men and women was in areas of personal consumption, where men were slightly more likely to view having a hobby, money to spend on oneself and fortnightly visits to the local pub as necessities. One of the challenges of incorporating these differences into quantitative analyses of poverty is that many surveys do not collect data on individual income or within-household distribution and even those that have cannot retroactively adjust previous years' data to allow longitudinal comparisons. The use of a multidimensional measure of poverty should elucidate to some extent the ways in which additional burdens faced by women increase their risk of deprivations of non-material dimensions of poverty. However, there is a risk that without a statistically robust way to analyse whether differences between men and women's underlying constructs of poverty differ enough to warrant different conceptual understandings they may be incorrectly treated as incompatible and measured differently when they may, in fact, on quantitative scale be comparable. Does the experience of poverty for men and women differ so substantively that they ought to be measured on different scales?

2.8.2 Household Structure, Children's and Older Peoples' Poverty

The age a person is has a profound impact on their risk of poverty, however, the question of whether poverty means substantively different things to different generations is one which has been left more unexplored. Children and older people have historically been at greater risk of poverty than working-age adults (*Children's poverty:* Platt, 2005. *Poverty and Social Exclusion in Old Age:* Townsend, 1979; Walker, 1980; Walker, 1993). Rates of relative poverty for children have been consistently higher than the rates for adults since the beginning of the 1980s, and the rates of pensioners have been higher than working age adults for far longer (Figure 4), only matching the rates of working-age adults very recently in around 2012 (Joyce, 2014). Although the criticisms of this measure still stand, the HBAI line can still give a general indicator of these differences in poverty risks. Although poverty for these groups has been decreasing, by this measure, since the 1990s, the latest figures from the Department for Work and Pensions suggest there has been a sudden growth, of around 2%, in the proportion of children in relative poverty (Department for Work & Pensions, 2016). Considering the aforementioned inadequacies of these measures to capture the complexity of poverty, this increase may reflect far more worsening conditions for children (Tomlinson & Walker, 2009).

Figure 4: Relative poverty (60% HBAI) rates in Great Britain

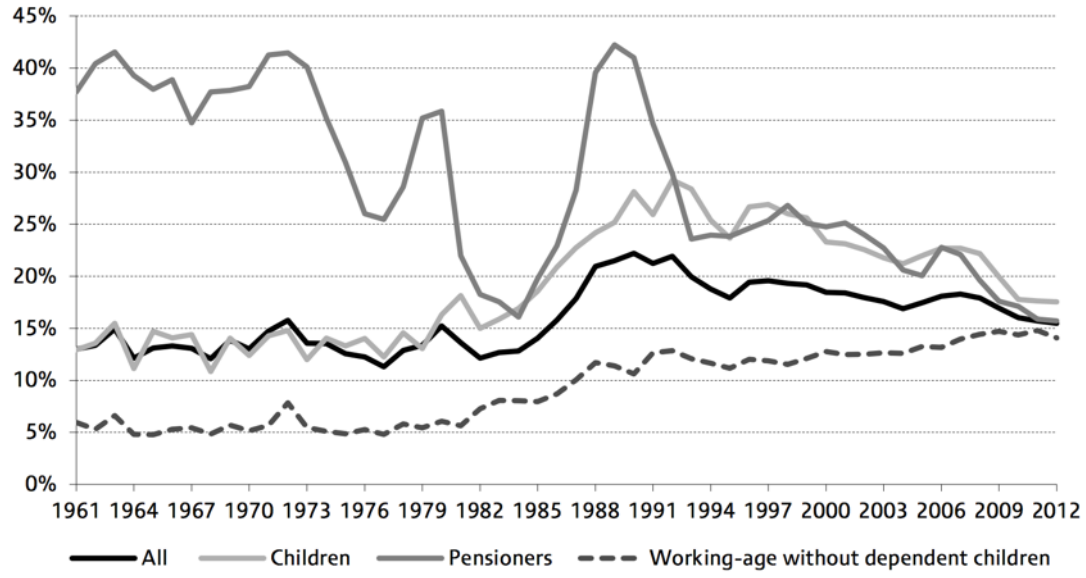
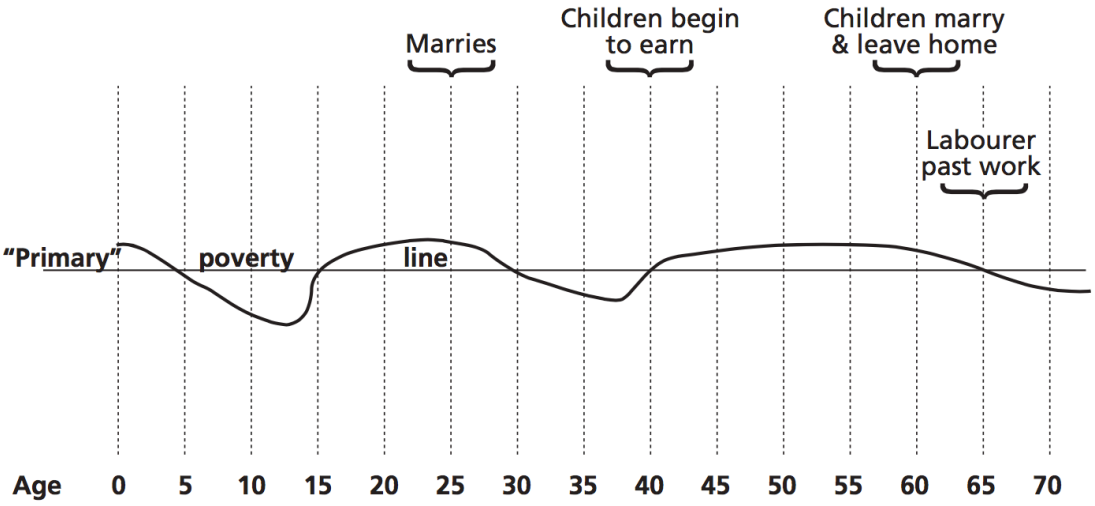


Figure reproduced from Joyce, 2014: 5.

These inequalities are fundamentally linked to the exclusion of people based on their age from the labour market, as well as shortfalls in welfare and work remuneration policies to bring above the poverty line those who are not able to, not obligated to, or not expected to be working. A central goal of welfare policy is to protect people at ages where this is the case from poverty

(Alcock, et al., 2012). For wider public policy, the expectation is that working adults are able to earn sufficient income to prevent members of their family from falling into poverty.

Figure 5: Rowntree's Observations of Poverty over the Life Cycle



Source: Rowntree 1901: 137

This observation of age-based inequality in poverty is not a recent phenomenon, and constituted a major part of Seebohm Rowntree's studies of poverty in York at the end of the 19th century which first drew attention to age-based inequalities in poverty (Alcock, 2006: 160; Glennerster, et al., 2004). Rowntree mapped poverty to the life course with reference to the ways in which the structure of society shaped the risk of poverty, largely in relation to labour capacity. These observations, along with other observations from leading sociologists, policy scholars and architects, went on to shape the redistributive design of UK welfare systems that aim to smooth income throughout the life cycle (Hills, 2015). Indeed, as the above graph above suggests (Joyce, 2014: 5), trends in poverty over the lifecycle have tended to move towards convergence, especially through the 1990s, by the measure of HBAI.

As this and the preceding diagram show, risks of poverty are not equal across the life course. Tomlinson & Walker (2009) explain the mechanism by which the distribution of income between adults and children within the family results in greater incidence of child poverty:

> *"Many adults – indeed the majority of adults – live in households without children and so do not have to spend any of their income meeting the needs of children. Since the main forms of income (wages, investment income and certain benefits) are not determined by household size or need, adults without children are typically better off than adults with children and hence, parental sacrifice nonwithstanding, better off than children themselves."*
>
> *Tomlinson & Walker, 2009: 63*

In this case, the implication is that the life course and household living arrangements are intricately aligned. Both in Rowntree's (1901) observations and the above quote, progression through the life course is tied up not only with the experience of ageing and changes in society but this itself is reflected in the composition of living arrangements. These living arrangements also appear to matter in relation to how poverty is experienced and how great a risk of poverty a household faces. To some extent then, changes in household composition, ageing or cohort membership at a given time, and changes in societal structures are joined up in a way that suggests poverty, as a dynamic concept, might differ in its substantive meaning depending on any of these factors.

When measured as a multidimensional concept the fall in child poverty between the years 1991 and 2003 was more marked than the fall when measured using the income-based HBAI measure (Tomlinson & Walker, 2009). While the proportion of children who met the at-risk-of-poverty criteria set by the HBAI line fell seven percentage points from 28.6 per cent to 21.6 per cent between 1991 and 2003, the proportion of children who met the multidimensional poverty criteria set by Tomlinson & Walker (2009: 64) fell almost ten percentage points from 26 per cent to 16.4 per cent. This was partly due to the multidimensional approach capturing not only improving income situations for families, but also a reduction in financial strain and material deprivation (*ibid*: 64-65). However, children's experiences of poverty may be substantively different to adults. They are likely to face greater stigma, shame, lack of autonomy and dignity, but in many regards also face a distinct lack of voice and occupy a societally reinforced state of civic powerlessness, frequently being the first members of society whose public services are cut in times of austerity (Lister, 2004; Walker, 2014; Webb & Bywaters, 2018).

Historically however, an age group with a far greater risk of poverty than children or working age adults has been pensioners (Joyce, 2014). Throughout the 1960s and 70s poverty was a certainty for the majority of people in retirement age, with Townsend identifying that 64 per

cent of older people were living on income which was far below their benefit entitlement (Townsend, 1979 quoted in Alcock, 2006: 166). These rates of poverty are closely associated with older peoples' relationship with primary sources of income (Alcock, 2006), particularly income from the commodification of labour which society does not value as highly as that which can be offered by people younger than the retirement age – an age often set with broad assumptions about the personal capacity and often ignoring medical advancements in healthy longevity and resultantly being more socially, than medically, imposed (Walker, 1980; Walker, 1993). In 2010, for the first time since poverty rates began being measured by the HBAI standard, pensioners were less likely to be living in poverty than the average rates across other household types (Adams, et al., 2012: 191-204). However, this ignores many of the dimensions found in conceptualisations of poverty that go beyond income, such as diminished autonomy as a result of the social creation of dependency (Walker, 1980), that older people continue to face disproportionately. Because of poverty metrics disinterest in these dimensions there has been something of a divergence leading to a greater focus on the quality of life of older people, the field of which incorporates far more reflections on health and wellbeing, social exclusion, and isolation – dimensions that should be featured in theoretically robust poverty measures as they are not limited to older age (Walker & Mollenkopf, 2007).

What is valued and constitutes a necessity can also differ across age groups. Reporting findings from the Poverty and Social Exclusion Survey, Pantazis, Gordon, & Townsend, (2006: 100-101) show how there were some significant differences between people aged over 65 and people aged between 16 and 24 in what goods, services or activities are deemed a necessity. For example, younger respondents in general were more likely to view fewer items and activities as essential; this was particularly the case for clothing items but also in the case of home entertainment, telephones and joints of meat (or non-meat equivalent). In contrast, older respondents were less likely to consider activities such as the ability to pick up children from school or the ability to attend school sports days as essential. However, outside of this there was still a good degree of consensus between the two groups ($R^2 = .81$). Whether this is related to cultural differences between cohorts, such as an increased focus on child safety in recent generations or an expectation of weekly home-cooked family meals in older generations, or reasons to do with mobility or health, such as recognition of the need for warm clothing and communication equipment to remain in contact with friends or family, or indeed some combination of the two, it is important to recognise that there are generational differences between the kinds of dimensions poverty is comprised of. We should not therefore assume that

a generalised multidimensional measure of poverty reflects the same experience for people of different ages, or living in different households associated with the life course.

2.8.3 Ethnicity and Poverty

Many Black and minority ethnic groups are disproportionately at risk of poverty (Platt, 2007), and more frequently than the predominant White British population experience deprivations in multiple areas of life (Alcock, 2006; Platt, 2007). This is not a result of biological differences which manifest from genes through 'skin colour and cultural differences' (Alcock, 2006: 148), as some explanations from the sociobiological tradition had argued throughout the latter half of the 20th century (for example, see Herrnstein & Murray, 1994) which result in different choices and inherited traits. Discussions around the high prevalence of poverty among BME groups in the latter half of the 20th century have been dominated by arguments from geneticists and 'socio-biologists' who have had a tendency to 'announce the discovery of genes for every human characteristic… with depressing regularity', ignoring complex factors in favour of a proposed deterministic biological trait (Spicker, 2007: 112).

Such studies set the discourse around poverty that people with minority ethnic identities find themselves discussed (LSE, 2016). Rather, ethnic divisions in poverty are part of a complex system of entrenched racism and intergenerational replications of inequality, poverty and discrimination where social discrimination reinforces economic disadvantage and prevents intergenerational accumulation of wealth and human capital, even when one considers the complexity around inherited poverty (Hills, 2015; Platt, 2007). There is good reason to suspect that this entrenched disadvantage may precipitate a fundamentally different experience of poverty for some BME groups. If certain groups face substantively different, or greater, barriers, it would stand to reason that they would require different or greater resources to overcome them. For instance, if BME groups face social discrimination and exclusion because of racist perceptions and structures in society regardless of their economic or human capital, it is unlikely that alleviating financial strain would change their experience in that dimension in the same way that it might for White British groups. It should not be presumed that a measure of poverty designed on the consensus of the White British majority accurately reflects the same underlying construct that minority groups experience.

With regards to peoples views on what constitutes deprivation, research exists that suggests there is a large amount of agreement between the ethnic majority and ethnic minorities in the

UK. Pantazis, et al., (2006: 102) report a high level of consensus around the necessities of life between BME groups and the White British majority, however, they have a very limited sample size of 72 respondents in the PSE survey. Still, BME and White British responses were closely correlated ($R^2 = .84$). However, there were some cultural differences between the groups: 80 per cent of BME respondents reported that they felt being able to attend a place of worship was a necessity, compared to 40 per cent of non-minority ethnic respondents. Furthermore, close to 45 per cent of BME respondents felt that being able to take a holiday abroad was a necessity, whereas only just under 20 per cent of non-minority respondents felt the same. Similarly, 25 per cent of ethnic minority respondents felt that satellite television was a necessity, but less than 5 per cent of non-minority respondents agreed. In contrast, non-minority respondents were more likely to agree that holidays away from home or being able to visit the pub once a fortnight were necessities of life. This reflects that there may be some significant cultural differences between ethnic groups in terms of what experiences might constitute certain dimensions of poverty.

To further complicate this, we should not assume that all ethnic groups share the same experiences. Some BME groups, for instance, are more vulnerable to poverty than others, and others are less vulnerable than the ethnic majority (Spicker, 2007: 56). The tendency in quantitative research to uncritically combine all ethnic groups into a single 'non-white' category in order to construct a sample size large enough to make statistical tests meaningful obscures these differences (Aspinall, 2000; Pantazis, et al., 2006: 102). Although progress has been made to increase sample sizes to a point that comparisons can be made between more meaningful categorisations of groups, especially through the use of 'sample boosters', the group sizes required for very complex types of statistical analysis such as structural equation modelling can easily exceed five-hundred (Wolf, et al., 2013). This is a limitation that is revisited later in the thesis.

2.8.4 Disability and Poverty

Disability and poverty are often closely related, with people living with disabilities usually at a much higher risk of being in poverty than people living without disabilities (Beresford, 1996; Alcock, 2006; Palmer, 2011). Using the Households Below Average Income approach, Tinson et al. (2016: 9) estimate that 28 per cent of people in poverty have a disability, and 19 per cent of these are of working age. The authors argue that the loss of earning potential created by societally imposed disability, and incurred costs associated with impairments, are not

adequately made up through disability benefits such as Disability Living Allowance (DLA) and Personal Independence Payments (PIP). Disability and poverty end up mutually reinforcing one another (Palmer, 2011; Alcock, 2006). People with disabilities are put into positions where they face greater needs that drain income. In many cases, trying to meet this excess of needs with limited resources can result in the exacerbation of disability, especially in regards to mental health comorbidities (Niles, et al., 2015).

However, differences in the ways that disability and poverty can be conceptualised raise important questions about how deep the relationship between poverty and disability runs. While the link between low income and disability is well established, there is a strong conceptual link to suggest that disabled people are overrepresented in other dimensions of poverty (Lister, 2004). The social model of disability presents, instead, a distinction between impairment and disability (Oliver, 1983; 2013; Beresford, 1996). Where impairment relates to the medicalised features of disability, the specific physical or mental individual differences presented, disability itself relates to the way in which the structure of society deliberately excludes those living with impairments. With regards to poverty these exclusions are usually imagined to be restricted to exclusions in the labour market and from earning potential (Alcock, 2006). However, using a broader definition of poverty it can be seen that people with impairments can face societally imposed disabilities in a great many dimensions, especially those that are non-material.

People living with disabilities often face a higher incidence of shame and stigma, which can manifest in decreases in mental wellbeing and health (Matthews & Harrington, 2000). They may also disproportionately face time poverty, with regards to navigating the barriers imposed by society that result in their individual differences being disabling. Beyond this, it is likely that such a societal structure, one designed without disability in mind, is itself incredibly damaging to a persons' wellbeing, life satisfaction, participation, and sense of inclusion. The extent of disabling barriers that affect any given person might be interpreted as a dimension of poverty itself, one that may be more specifically associated with a person's physiological or psychological differences.

What's more, the additional risk and depth of poverty experienced by people living with disabilities extends to people with caring responsibilities, the vast majority of these being women (Lister, 2004; Tinson, et al., 2016). The fiscal and 'opportunity costs' associated with

those with caring responsibilities, which include their restrictions on working hours and the uptake of additional work-related responsibilities, can greatly increase the risk of income poverty (Alcock, 2006). Carers also experience similar social and psychological costs, often having to balance caring and work responsibilities and sacrificing social participation, and facing anxiety and stress from the demanding situation they find themselves in (Howard, 2001).

However, despite disability and caring responsibilities having such a strong relationship with poverty there is typically very little nuance in the measurement of differences between people living with disabilities and different types of caring responsibilities. Statistical reporting of the relationship between disability and poverty typically uses a fairly general definition of disability (and poverty) uncritically. Disability is often, for instance, approximated by using responses to a survey question asking if the respondent has any longstanding illnesses or whether they are in the receipt of disability or caring related welfare benefits.

There are quite significant problems with either approach. In the former, the prerequisite of a longstanding illness assumes that all kinds of individual physiological and psychological differences have the same disabling barriers in society, which is not the case. A person with a clinically diagnosed antisocial personality disorder is likely to face very different disabling barrers to someone diagnosed with chronic pain syndrome. These differences may also affect different dimensions of poverty more than others. Using such a criteria may therefore exaggerate the extent of disability-related poverty and underestimate the depth of poverty experienced by disabled people. In contrast, if an analyst classed only those receiving a disability related benefit as disabled, they would likely exclude a great many people who face disabling barriers but do not meet the criteria to qualify for DLA or PIP. This would likely underestimate the extent of the relationship between disability and poverty and over-exaggerate the depth of poverty faced by those with disabilities. A measure that attempts to gauge the level of societally imposed disabling barriers related to physiological or psychological differences, without the assumption of a clinical diagnosis, would be a patently better way to investigate this important link.

2.9. Moving measurement forwards

In summary, this section has brought together competing and complementary conceptualisations of poverty and compared these to the measures commonly used to estimate and approximate poverty. This included theories of participation and capabilities (Townsend,

1979; Sen, 1995) and the ways in which people living in the UK, including the poor themselves, define and conceptualise poverty. Not only is poverty multidimensional, these dimensions are complex and shift over time; they involve processes of labelling and othering, stigma, and resistance. Definitions used are also often strongly linked to the political viewpoints of those involved (Tomlinson, et al., 2010). Furthermore, this conceptualisation raises the issue of whether income and consumption can accurately approximate the extent of non-material deprivation experienced by those in poverty: experiences of stigma, shame, exclusion, low human capital, dignity and powerlessness.

A comparison of different measurements of poverty has identified discrepencies between different measures and weaknesses in the ways in which they have been designed because of methodological restrictions, as well as the embeddedness of income measures in UK social policy. In some cases it could be argued these measurements have taken on a life of their own that may no longer accurately reflect the underlying concepts and experiences (Lister, 2004). Not all of these competing measures can be correct when they show very different things. With new methodological innovations this section has also explored the ways in which recent research has attempted to develop more accurate quantitative measures of poverty that are sensitive to changes in prices, wages, benefits and social consensus, and can be inclusive of the non-material elements of poverty (Bradshaw, et al., 2008; Tomlinson, et al., 2008; Ferragina, et al., 2012). In addition to demonstrating the potential for these new measures, the findings also raise fundamental questions about the way that we think of poverty as a binary classification of the 'poor' and the 'not poor', and the rammifications of this for targeted welfare systems based on means-testing. After problematising existing approaches, a compelling case for transferring multidimensional theories of poverty into measurement has been presented as a way of researching poverty which does more 'justice to the experiences' of those living in poverty (Tomlinson, et al., 2008: 600).

Although the frequent reimagining of measures fosters healthy scepticism of poverty metrics, and produces valuable and interesting research in itself, these re-investigations and reimaginings can risk devaluing the concept of poverty altogether. Indeed, this has happened historically and is common in everyday understandings of poverty, where measurements of poverty can bear little resemblence to lived experience (Shildrick & MacDonald, 2013). In *the Poor and the Poorest*, Brian Abel-Smith and Peter Townsend make the following argument:

“Two assumptions have governed much economic thinking in Britain since the war. The first is that we have ‘abolished’ poverty. The second is that we are a much more equal society: that the differences between the living standards of the rich and the poor are much smaller than they used to be.”

Abel-Smith & Townsend, 1965, in Banks & Purdy (2005: 66)

Thirteen years later, Margaret Thatcher was quoted as saying that ‘nowadays there really is no primary poverty left in this country… In Western countries we are left with problems that aren’t poverty’ (Margaret Thatcher, cited in Jones, 2012: 64). A similar denial of poverty is playing out again, in 2018, as the Conservative government under Theresa May refute allegations of poverty in the UK in response to a report from the United Nations’ special rapporteur on poverty, Philip Alston, with a spokesperson for the Department of Work and Pensions arguing that ‘household incomes have "never been higher", income inequality has fallen and there are one million fewer people living in absolute poverty compared with 2010’ (BBC, 2018: online).

Creating such devaluation in the concept by repeated reimaginings of equally problematic measures can be a great injustice to the people the study of poverty aims to give a voice to. Because such strong moral arguments are usually made for certain approaches to measurement they can be easily countered by equally passionate moral arguments. This is a point of contention that may require rather more caution from academics who seek to redefine poverty without an acute awareness of how such repeated attempts can affect the metanarrative in a way that creates more, not fewer, deaf ears to the plight of those who are or who have lived in poverty. This risks turning the discussion of the definition of poverty into what Piachaud called ‘a semantic and statistical squabble that is parasitic, voyueristic and utterly unconstructive’ (Piachaud, 1987: 162). There are real concerns about the consequences of such unconstructive endeavors, especially when one considers the disproportionate risks that certain marginalised groups face.

Therefore, movement towards additional valid claims to knowledge about the existence and nature of poverty, especially those that utilise quantitative methods, are crucial for securing a definition of poverty that strengthens the voice of the poor. It is a labourious process. It does not provoke emotive responses for social change in the way that life stories do. It does not foster moral outrage and social protest in the way that ethnographies, documentaries, or participatory action research can. However, metrics provide the solid canvas that these rich

stories can be painted, without which they are easily dismissed. A broader canvas provides a better foundation to capture diversities of experience. A stronger canvas means that these experiences are better supported and harder to break. But more attention is often given to finding more vibrant ways to describe lived experiences of poverty than is given to strengthening the frame, supports, and weave of our canvas, although both are equally important. The value of metrics is frequently overlooked but this is their central importance: the better our canvas the better the entire picture can be seen and contextualised, including the details. This chapter has explored the ways in which poverty measurements fail to do justice to theories of poverty; the following chapter explains why methodological sophistication is essential to understanding why this is so, and introduces an alternative methodology.

Chapter Three: Methodology
How to construct a multidimensional poverty measure?

3. Introduction, Research Aims and Questions

The primary methodological aim of this research is to apply the methods used in Tomlinson, et al., (2008) to more recent data collected by the Institute of Social and Economic Research in the form of the *Understanding Society* survey, and then to test the resultant factors for reliability and validity. This involves constructing a multidimensional measure of poverty that takes account of non-monetary facets of deprivation that are found in conceptualisations of poverty developed by Townsend (1979), Baulch (1996), Lister (2004), Spicker (2007), Walker (2014), and others. The research builds on this previous work by introducing procedures for dealing with missing data, reliability, validity, and invariance testing to better understand the stability of the measure as both a dynamic concept and state of being. The primary research aim was therefore to construct a multidimensional poverty index using structural equation modelling based on theories of poverty in order to inspect its goodness of fit, how well the theoretical arrangement of items fits the reality of a representative sample of the population, how sensitive each domain is to changes in time and between different groups, and whether these differences are so pronounced that measures which use a different underlying construction of poverty should be applied to different time periods or different groups of people. This chapter provides an overview of existing methods and their logic in approaching such questions, however, some methodological developments are proposed as a result of the empirical research in the thesis. These are focused on within the context of the substantive findings, whereas this chapter focuses more on providing the reader an introduction to the methods used. The research questions are as follows:

1. Is there clear evidence for underlying 'dimensions' of poverty within secondary data survey items, as theory would suggest?
2. Is the construction of a multidimensional poverty index that uses vastly different measures possible, and how well do these factors fit within a broader multidimensional poverty construct, and can this tell us anything about the nature of poverty?
3. Are the dimensions of poverty identified in theory valid, are they distinct enough to be real differences, when using variables from existing datasets?
4. To what extent are these dimensions of poverty consistent over time and between groups? Can the same multidimensional measure of poverty be applied to different groups, men and women, different households or cohorts, equivalently?

5. What might be the consequences of this more complex measure of poverty for social policy?

One additional research question makes a specific methodological contribution to the existing literature:

6. Do different imputations of missing data change the inference of a structural equation model using *Understanding Society* data?

Research questions one through five can be investigated using the structural equation model fit and composition, model fit comparisons, factor invariance testing, and factor validity testing. Question six requires the comparison of model fit, effect sizes, and standard errors across models using different imputations of data. The first part of this chapter, methodology as applied epistemology, discusses the methodological approach that is the foundation of the choice of methods used and the remit of their discussion within the thesis. The second part outlines the secondary data used for the research, including its benefits and weaknesses, and the variables included in the data analysis. Part three outlines the basic logic and use of structural equation modelling and factor analysis, and why it is so applicable to the study of multidimensional poverty. Parts four, five, and six outline the various approaches taken to handle features inherent in the structure of data and the challenges of employing advanced quantitative methods with complex survey data. Part four outlines a workflow approach that draws on the principles of exploratory data analysis, five discusses the approaches to invariance and validity testing, and part six discusses competing ways of handling missing data. Finally, the chapter offers a reflexive summary of the unique contribution that the application of these methods can make to better understanding and investigating multidimensional poverty and answering the above research questions.

3.1 Methodology as applied epistemology

This section intends to outline the methodological approach of the research project. It subscribes to Jenkins' (2002: 7) definition of methodology as 'systematic applied epistemology' as an attempt to build evidence to make claims about things that correspond with the realities of the human world. In what way do the methods chosen in this research build empirical support with which to make a claim to knowledge of the subject matter? To this end, the methodology employed here aims to produce knowledge that is plausible, making sense when applied to the real world, and defensible, robust in its approach to the production of knowledge. The

principles used can be outlined as follows: the first principle is that the methodological rigour traditionally associated with quantitative research that typically involves the checking of assumptions and errors and employing the most suitable statistical methods, should be retained. The second principle involves making explicit the workflow decisions and processes that aim to reduce problematic trajectories researchers can take through a 'garden of forking paths' (Gelman & Loken, 2013), concerns that have historically been associated with statistical research but are now seeing something of a renaissance in the field of data science and have long had equivalent counterparts in qualitative social research (Mason, 2002). Thirdly, this chapter recognises the inherently qualitative decisions that are made in all stages of the quantitative research process and argues that engagement with reflexivity can offer benefits to quantitative methodology without substantial departure from methodological objectivity or realist interpretations of the human world.

The topic under investigation and the means by which it is investigated offer a real opportunity to engage with the three ways that contemporary sociology should be challenged identified by Jenkins (2002: 2):

> "*First, general theory and empirical research have become too weakly connected to each other;*
> *second, sociologists routinely fail, at least in part because of the ways in which they write, to address the wide audience that the discipline's subject matter surely demands; and*
> *third, the sociological aspiration to objectivity – problematic as it is – which is so vital if we are to defend successfully our claims to understand the human world, has been particularly threatened in recent years."*
>
> Jenkins, 2002, Foundations of Sociology: 2

The research aims to address these challenges, but cannot claim to meet them fully. Firstly, the research is centrally concerned with reconnecting theories of poverty with quantitative evidence by using methods that move towards models that better resemble the complexity of poverty as it is conceptualised and experienced. Secondly, the methods used throughout can be highly technical, however, the general reasons for their application can be explained in relatively accessible ways, appropriate for a diverse audience; the principles and aims of the techniques used should not be obfuscated by the technical procedures involved. Certainly, these technical explanations should not be given in isolation with the expectation being that clarity will emerge from the details, or that the reader should go away and become an expert in what is a very specialised and complex area of applied statistics. Thirdly, this methodology

recognises that a constructive definition of objectivity is one that acknowledges the fallibility of the researcher, to hold and be influenced by their own views and experiences, and factors this into consideration, by reporting honestly rather than with disinterest why certain decisions have been made and ways in which critical distance has been maintained (*ibid*). This is contradictory to traditional understandings of objectivity and the positionality of quantitative research within a strictly positivist, scientific framework.

3.2 Secondary data analysis, questions and compromise

Secondary data analysis involves the use of data, usually quantitative, that has previously been collected for some other purpose in order to answer research questions that it had otherwise not been explicitly intended to answer (Bryman, 2008). The idea of this kind of analysis as 'secondary' is historically rooted in instances when research teams would conduct their own surveys using bespoke questionnaire designs and the resulting data could then be reused to answer further questions in other projects without further fieldwork (MacInnes, 2017). In other words, data collected could be analysed for answering the primary research questions that the data was intended to answer, but later may be used to try and answer further research questions. Although this still happens, most data used in secondary data analysis now comes from large omnibus surveys conducted by specialist governmental, or government-funded, research institutes (MacInnes, 2017). There are few 'primary' research questions associated now with the largest surveys, replaced by overarching themes and modules, which challenges the 'secondary' status of secondary data analysis. Nonetheless, there are clear drawbacks and benefits associated with using secondary data in research.

Secondary data analysis requires a substantial commitment of time in the scoping of potential datasets, understanding their design and data collection protocols, sampling methods, examining which variables are included and how these have been operationalised. Understanding missing data, attrition over time, and data structures further complicates this. To put this undertaking in perspective, the individual response survey file for *Understanding Society (UKHLS)* contains 946 variables. These can be linked to other data files that include questions at the household level, or related just to modules completed only by a youth cohort. The user guide for waves 1-5 of *Understanding Society* is 105 pages long, documentation on questionnaires for each wave are over 300 pages long, and their accompanying technical documents contain over 60 pages of additional information and provisions. The aspiring end user must therefore be willing to dedicate a substantial amount of time familiarising themselves

with the secondary data, without any tacit knowledge from the process of undertaking the data collection or survey design to draw upon (Cheng & Phillips, 2014). In many instances, the operationalisation of variables or the frequency of their measurement is not ideally suited to the research in mind – some variables may be missing entirely – and the secondary data analyst needs to either work with or around this to take advantage of the many benefits that secondary data offers.

Chief amongst these advantages are the quality and quantity of data that is available. Research institutes that design and administer surveys employ extensive teams of survey methodologists with expertise in designing, piloting, and administering survey questionnaires (Bryman, 2012). They employ techniques that assess the robustness of indicator measures, purchase psychometric and econometric scales, decrease non-response, and have the resources and knowledge to establish and maintain very large samples of the population for longitudinal analysis (MacInnes, 2017). The scope and depth such survey data offers is far beyond what the resources available to small projects, such as this, would be able to procure. Although there are difficulties that accompany the use of secondary data – it may be frustrating to 'decode' and learn about them – being able to work with data that is so carefully curated by experts is valuable.

Understanding Society (UKHLS) is a longitudinal survey of approximately 40,000 households in the United Kingdom. It amalgamates some of the topics from several other surveys that were discontinued in 2009, the most important for this research being the *British Household Panel Survey (BHPS)*, with an expanded sample size as well as booster samples of minority groups in the population (Knies, 2017). MacInnes (2017: 112-114) recommends answering the following twelve questions about secondary data using the accompanying documentation to provide a clear overview of the data and assess its eligibility for a specific secondary data analysis research project.

3.2.1 Understanding Society: twelve questions to ask about data (MacInnes, 2017)

1 **When was the survey fieldwork carried out?**

Understanding Society's data collection operates on a four year staggered repeating development, data collection, data processing and release cycle.

Development of survey design is completed by the Institute for Social & Economic Research (ISER, Essex), fieldwork is completed by NatCen Social Research and the Central Survey Unit of NISRA (waves 1-5) and later by Kantar Public and Millward Brown Ulster (waves 6 and onwards). Survey development lasts for one year and involves pilot studies, including interviews that use translated versions of the questionnaire, the briefing and training of interviewers and fieldwork managers. Fieldwork commences at the end of development and continues for two years and five months, with the final five months dedicated to trying to complete surveys for participants who are eligible but could not complete the questionnaire at first contact. After the initial two years of data collection, data processing begins with the aim to release data in quarter four of the last year for that wave. Each new wave begins a year after the previous wave starts, for example, wave two's questionnaire development began at the end of wave one's development, at the same time as wave one moved into the data collection phase. The research unit also conduct follow up quality assurance interviews with approximately 10 per cent of the sample. (Knies, 2017: 15-21)

2 **Is the survey one of a repeated series?**

Understanding Society is repeated annually with the same target population and sample, however, not all questions are repeated each year to reduce respondent fatigue (Lavrakas, 2008). Examples are question sets about material deprivation, which were asked in wave one, two, four, and six, but not in waves three and five. Key questions are asked each year, such as those about household income or composition, other questions are derived from responses in previous waves. In order to analyse the data within a 'tidy' structure and use consistent time function estimates, new variables were constructed here that included a column of 'missing values' for years where questions were not asked. These missing data were then imputed using multiple impitation by chained equations (discussed in a later section).

3 **Is there a panel element to the survey?**

Understanding Society is a panel survey where all original participants from the first wave are treated as permanent original sample members (OSMs). Any

children who are born to women who are original sample members are also given original sample member status. People who join the household after the first wave are invited to complete the survey as temporary sample members (TSMs), but are transferred to permanent sample member (PSM) status if they parent a child with an original sample member. This means that they and their child are both followed up by the survey regardless of whether there is a later dissolution of their partnership with the original sample member (Knies, 2017: 21).

4 & 5 **What was the target population? What was the sampling method and weighting strategy?**

The survey intends to construct a sample that represents the general UK population based on residential addresses. The sample for England, Wales, and Scotland is selected following a proportionally stratified design using a clustered sample of addresses from the Postcode Address File (a database of all Royal Mail postal addresses). Therefore, each household within these clusters of addresses had an equal probability of being selected for the survey. In Northern Ireland, the survey employs an unclustered systematic random sample of addresses from the Land and Property Register Services Agency list of domestic addresses (Knies, 2017: 10-12). Clusters were stratified by proportions of the population in non-manual occupations, these sub-strata were then sorted further into groups based on population density before finally being sorted by ethnic minority density (*ibid*). The final sample should therefore represent more broadly the demographics of the entire population of the United Kingdom, although there are certain individuals that are not within the sampling frame, notably people who do not have a registered address. This has consequences for studies about poverty that rely on this data, so any inferences should not be made that relate to the 'entire' population experiencing poverty, but rather the population that have a registered address. Those without residences are very likely to be experiencing deprivation.

Understanding Society contains many survey weights that adjust responses and statistics to better represent the target population and make more accurate inferences about the UK population. This is important for adjusting for identified nonrandom nonresponse, for example, due to attrition or low response rates from

ethnic minority groups. However, the incorporation of such survey weights to more complex analyses of data, such as the structural equation models used here, is still not possible in most statistical packages. There is contention around the best way to deal with weighting data in analytical studies (Lee & Forthofer, 2006).

The introduction of sampling data weight variables to analytical models can introduce confounding covariance amongst independent variables and in complex models can be inefficient by introducing greater variance (*ibid:* 20-23). However, the omission of survey weights assumes that the estimates in models would be the same regardless of attrition and estimates for groups with higher rates of nonresponse are the same as, or more statistically similar, to groups with lower rates of nonresponse (because of higher standard errors) (Knies, 2017). If nonresponse introduces bias into any analytical models the absence of weights is more problematic. Model estimates can be accurately generalised to the population if models include enough independent variables to adjust effects for sociodemographic biases, but models can still be mis-specified if some effects are omitted (such as interaction terms, e.g., between age and education) (Lee & Forthofer, 2006: 75-76). When weighted point estimates cannot be used in analytical models it is therefore important to control for variables associated with nonresponse – such as time, gender, and ethnicity. In this case we are explicitly interested in differences between these groups and there is a strong theoretical justification for controlling for them in any models, which mirrors the approach taken by Tomlinson, et al. (2008; 2009). The fact that *Understanding Society* includes many of the demographic variables used for weighting data means that estimates derived using the methods and factors developed here can be adjusted appropriately for nonresponse bias in the sample.

6 **What was the response rate?**

There are few figures on 'average' response rates for national surveys to put *Understanding Society's* reponse rate in perspective. Some surveys, such as postal self-completion questionnaires, can have low response rates of 10-15 per cent of the sample (Bryman, 2008). Surveys conducted by research institutes using face-to-face questionnaires typically achieve response rates of 50-60 per cent (*ibid*:

199). Nonresponse is problematic when the reasons for it are not random, for example, if participants from certain socioeconomic groups have less time to dedicate to the survey. It is unlikely that nonresponse here is random and mechanisms of missingness are discussed further in a later section. *Understanding Society* achieved a response rate for households of 57.1 per cent in Great Britain and 60.9 per cent in Northern Ireland in wave one (Knies, 2017: 22). Within households the survey had response rates from all eligible members of the household of 82 per cent in Great Britain and 77.3 per cent in Northern Ireland (*ibid*). This high response rate is partly due to the strategies that the survey methodologists employ to reduce nonresponse, which includes a high minimum number of face-to-face visits to establish contact, telephone interviewing options, follow-ups, interviewer training, 'mop-up' interviews, and six month updates provided to participants along with regular opportunities to update their contact details to avoid attrition (*ibid*).

7 **Who answered the questions?**

Each eligible adult is surveyed, and a sample of young people are provided a youth questionnaire (Knies, 2017). Data can be collected by proxy, but this is treated as missing in the datasets if not confirmed as correct by the participant. Each household is designated a 'household reference person', although this is not necessarily the person that answers the household level questionnaire. This can introduce a division in the gendered dimension of survey response. In this instance, the household reference person is the eldest owner or renter of the accommodation in which the household lives (Knies, 2017). In many cases, this person is male. However, women may be more likely to complete survey questionnaires. Davies (2011) notes that, in her research, women tended to answer the door to home visits more frequently and after being asked whether they would like to take part in the research project men were more likely to defer the responsibility to their female partners.

8 & 9 **Does the survey contain modules administered only to subsets of respondents? What data did the survey collect?**

Understanding Society contains a set of questions that are asked on first contact that are considered time-invariant, for example, country of origin, year of birth, or ethic group category. For many participants these are asked in wave one, but new entrants at other time points are asked these when appropriate. The survey then has many modules that are repeated yearly, some of these include questions along the themes of disability, caring, employment, childcare, politics, income and benefits that are closely linked to the previous focus of the BHPS. Because of the very broad focus of the UKHLS, there are several rotating modules that have been asked in different waves that are based on different themes (a full outline of which can be found in the user guide, Knies, 2017: 48-53). In this research, the least frequently recurring modules are those that include questions about friendship (asked in waves two and five), membership of organisations (asked in waves three and six), standards of local services (asked in waves three and six), and questions about self-esteem (asked in waves four and six) – the exact variables included in the analyses are listed and described in table 4 along with any transformations.

10 **What use have others made of the data?**

This research builds directly on the work of Tomlinson, et al. (2008; 2009) but using the successor to the BHPS, the UKHLS, including a number of legacy variables. The *Understanding Society* team keep a publication catalogue (https://www.understandingsociety.ac.uk/research) that documents the use of the data and, because of its wide scope and novel introduction of biomarkers and other biological data, its application has been in fields of anthropology, public health, sociology, epigenetics, economics, and beyond. Researchers and the survey methodologists have also used the findings to develop methodological advancements in social research, for example, to explore non-response in biomarker data collection (Cernat, et al., 2016). The team produce an annual report series (Insights) that documents research that has been made possible using the *Understanding Society* data (Understanding Society, 2017).

11 **How can I access the data? Can I review it online without downloading?**

Understanding Society data can be accessed online via the UK Data Service (previously the Economic and Social Data Service). Access requires users to

create an account, agree to and abide by terms of service that forbid the use of the data for commercial purposes, and register a legitimate use of the data for the purpose of record keeping (UK Data Service, N.D.). Versions of the data with identification risk factors, such as location or travel-to-work data, require special licensing, training, and reporting agreements. Due to the size of the data it cannot be explored online using the UKDS' NESSTAR service, but Understanding Society offer documentation containing information about variables and design separately on their web page.

12 **How do I cite the data I use?**

The bibliographic citation preferred by Understanding Society for their data is as follows (also found in the bibliography of this thesis):

University of Essex. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. (2017). *Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 9th Edition.* UK Data Service. SN: 6614

3.3 Variables

Selection criteria for variables was firstly based on matching variables that linked to Tomlinson, et al.'s (2008) paper, *Measuring Poverty in Britain as a Multi-dimensional Concept, 1991 to 2003*, prior to selecting other variables of interest based on whether they had some theoretical justification implied from the literature review. In some instances, questions measuring the same variables had been unchanged since their use in the BHPS. In other cases, questions have been discontinued, but other variables have been identified that may approximate the same underlying concept when used in a structural equation model. The table below documents the variables that are used in this research, as well as a description and any transformations they have undergone. Variables are grouped by their proposed dimension, following Tomlinson et al.'s (2008; 2009) groupings, that are informed by the theoretical frameworks of the work of Peter Townsend, Bob Baulch, Ruth Lister, & Paul Spicker discussed in the previous chapter. Later groupings to develop latent variables are based on these with the addition of exploratory data analysis using a visualised correlation matrix sorted using hierarchical clustering to inform their construction. The levels of measurement in the cross-classified multiple membership multilevel data structure are also identified in the following table. Although the technical

capacity to extend the model to a multilevel framework was not available in this research, the identification of different levels is highlighted as valuable for understanding some potential weaknesses in the findings.

The poverty dimension refers to the theoretical grouping that variable can be imagined under, and the descriptions specify how the question has been coded by *Understanding Society*. The variable name refers to the original name of the variable in the *Understanding Society* data, which can be searched by anyone downloading the dataset. Finally, the level refers to the clustered hierarchical groupings within the data, which are explained in more detail in section 3.5. Briefly, *ti* refers to a variable that is measured at specific time points for each individual (time points are nested within individuals); *tj* refers to a variable that is measured at specific time points for specific households (time points are nested within households, but are the same for all individuals in the household); and *tij* refers to data that is answered at a specific time point which can be aggregated up to the level of the individual (the average across all time points), and aggregated again to the level of the household (the average across all individuals in each household). *i* refers to variables that are assumed to not change over time, and are at the individual level, an example being ethnic origin. Because these variables can often be expressed as percentages within a household, *i* level variables can exist at both levels *i* and *j*, hence *ij*.

Table 4. A comparison of Multidimensional Poverty related variables used in Tomlinson, et al., (2008) and in *Understanding Society*.

	Tomlinson, et al., (2008)	**Wave 1-6 Understanding Society**		
Poverty Dimension	Description (Coding)	Variable/s	Description (Coding)	Level
Financial Strain	Whether a housing bill has been missed in the last 12 months (0/1)	*xphsdb*	Whether household is behind on rent/mortgage payments (0/1). All waves.	*tj*
		xphsdct	Whether household is behind on council tax payments (0/1). All waves.	*tj*
		xphsdba	Whether the household is not up to date with all bills (0/1). All waves.	*tj*

		matdepf	Whether the household isn't able to put some money aside for savings regularly, but would like to be able to if they could afford it (0/1). Waves 1, 2, 4, & 6	*tj*
	Whether respondents considered their financial status to be good or bad (5-point scale) Whether respondents felt their financial status would get better, worse, or stay the same over the future (3-point scale)	*finnow*	Whether respondents feel like their financial status is good or bad (5-point scale, higher = worse). All waves.	*tij*
	Natural logarithm of equivalised household income (continuous, normally distributed)	*hhnetinc3* *ieqmoecd_dv*	Household net income (after taxes and cash transfers) before housing costs, adjusted for family size using OECD equivalence scale (Anyaegbu, 2010). National logarithm plus constant value for incomes below 0. All waves.	*tj*
Material deprivation	Household does not possess: a CD Player (0/1)	*cduse11*	Household does not possess: a CD Player (0/1) All waves.	*tj*
	a VCR (0/1)	*cduse2*	a VCR/DVD Player (0/1) All waves.	*tj*
	a Washing machine (0/1)	*cduse6*	a Washing machine (0/1) All waves.	*tj*
	a Tumble dryer (0/1)	*cduse7*	a Tumble dryer (0/1) All waves.	*tj*
	a Microwave oven (0/1)	*cduse9*	Microwave oven (0/1) All waves.	*tj*

a Dishwasher (0/1)	*cduse8*	a Dishwasher (0/1) All waves.	*tj*
a Personal computer (0/1)	*cduse10*	a Personal computer (0/1) All waves.	*tj*
	cduse5	a freezer (0/1) All waves.	*tj*
Central heating (0/1)	*heatch*	Central heating (0/1) All waves.	*tij*
Use of a car (0/1)	*ncars*	Number of cars owned, transformed into whether they own at least one car (0/1) All waves.	*tj*
	matdepa	Can't afford a holiday for a week at least once a year, but would like this (0/1). Waves 1, 2, 4, & 6	*tj*
	matdepg	Can't afford to replace furniture if it breaks, but would like to be able to (0/1). Waves 1, 2, 4, & 6.	*tj*
	matdepb	Can't afford to feed friends or family that visit at least once per month (0/1). Only asked in wave 1 and 2.	*tj*
	hheat	Able to keep home adequately warm (0/1). Waves 1, 2, 4, & 6.	*tj*
	matdepc	Can't afford two pairs of all-weather shoes for each member of the family, but would like to (0/1). Waves 1 and 2 only.	*tj*
	matdepd	Not able to keep their house in a decent state of repair, but would like to (0/1). Waves 1, 2, 4, & 6.	*tj*

		matdepe	Is not able to afford contents insurance, but would like to be able to (0/1). Waves 1, 2, 4, & 6.	*tj*
		matdepf	Cannot afford to put aside savings regularly, but would like to be able to (0/1). Waves 1, 2, 4, & 6.	*tj*
		matdeph	Cannot afford to replace major electrical goods if they break, but would like to be able to (0/1). Waves 1, 2, 4, & 6.	*tj*
	Natural logarithm of equivalised household income (continuous, normally distributed)	*hhnetinc3* *ieqmoecd_dv*	Household net income (after taxes and cash transfers) before housing costs, adjusted for family size using OECD equivalence scale (Anyaegbu, 2010). Natural logarithm plus constant value for incomes below 0. All waves.	*tj*
Social isolation	Does the respondent feel like they have someone who: will listen to them (0/1) Will help them in a crisis? (0/1) They can relax with? (0/1) Who appreciates them? (0/1) Who comforts them? (0/1)			
		scwemwbf	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: close	*tij*

	to other people (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	
scfrendany	Whether participant would say that they do not have any friends (0/1). Asked in waves 2 and 5.	*tij*
scfundstnd	How much the participant feels that their friends really understand the way they feel about things (4-point scale, higher = understand less), Asked in waves 2 and 5.	*tij*
scfrely	How much the participant feels that they can rely on their friends (4-point scale, higher = can rely on less), Asked in waves 2 and 5.	*tij*
scfopenup	How much the participant feels like they can open up to their friends (4-point scale, higher = can open up to less), Asked in waves 2 and 5.	*tij*
scfcritic	How much the participant feels like their friends criticise them (4-point scale, higher = are critiqued by friends often), Asked in waves 2 and 5.	*tij*
visfrnds	Whether the participant goes out for social events or not (0/1). Asked in waves 3 and 6.	*tij*

		visfrndsy2	Whether financial reasons restrict respondent from visiting friends (0/1) Asked in waves 3 and 6.	*tij*
		visfamy2	Whether financial reasons restrict respondent from visiting family (0/1). Asked in wave 3 only.	
Civic participation	Membership of the following organisations: political party, trade union, environmental group, parents association, tenants or residents group, religious group, voluntary service, community group, social group, sports club, women's institute, women's group, or any other organisation. (Count of N membership, continuous variable)	*orgm1* *orgm2* *orgm3* *orgm4* *orgm5* *orgm6* *orgm7* *orgm11* *orgm12* *orgm13* *orgm14* *orgm15* *orgm16*	Membership of the following organisations: political party, trade union, environmental group, parents association, tenants or residents group, religious group, community group, social group, sports club, women's institute, women's group, or any other organisation. (Count of N membership, continuous variable). Asked in waves 3 and 6.	*tij*
	Whether an active member of the following organisations: political party, trade union, environmental group, parents association, tenants or residents group, religious group, voluntary service, community group, social group, sports club, women's institute, women's group, or any other organisation. (Count of N active membership, continuous variable)			
		orgm7	Whether the participant takes part in voluntary service	*tij*

			(0/1). Asked in waves 3 and 6.	
		vote8	Whether the participant voted in the last general election (0/1). Asked in wave 2 only.	*tij*
Psychological strain – anxiety/depression	Whether the participant has: GHQ12-2 lost much sleep (0-3 scale)	*scghqb*	Whether the participant has: GHQ12-2 lost much sleep (0-3 scale). Asked in all waves.	*tij*
	GHQ12-5 been under stress (0-3 scale)	*scghqe*	GHQ12-5 been under stress (0-3 scale). Asked in all waves.	*tij*
	GHQ12-6 been able to overcome difficulties (0-3 scale)	*scghqf*	GHQ12-6 been able to overcome difficulties (0-3 scale) Asked in all waves.	*tij*
	GHQ12-9 been feeling unhappy or depressed (0-3 scale)	*scghqi*	GHQ12-9 been feeling unhappy or depressed (0-3 scale) Asked in all waves.	*tij*
		scwemwbb	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: useful (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	*tij*
		scwemwbc	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: relaxed (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	*tij*
		scptrt5n3	Big5 inventory: participant sees themselves as someone who is: relaxed (5-point scale	*tij*

			on strength of agreement, inverted so higher scores = less likely to agree). Asked in wave 3	
Psychological strain – social dysfunction	GHQ12-1 able to concentrate (0-3 scale)	*scghqa*	GHQ12-1 able to concentrate (0-3 scale) Asked in all waves.	*tij*
	GHQ12-3 playing a useful part (0-3 scale)	*scghqc*	GHQ12-3 playing a useful part (0-3 scale) Asked in all waves.	*tij*
	GHQ12-4 capable of making decisions (0-3 scale)	*scghqd*	GHQ12-4 capable of making decisions (0-3 scale) Asked in all waves.	*tij*
	GHQ12-7 enjoy normal activities (0-3 scale)	*scghqg*	GHQ12-7 enjoy normal activities (0-3 scale) Asked in all waves.	*tij*
	GHQ12-8 can face up to problems (0-3 scale)	*scghqh*	GHQ12-8 can face up to problems (0-3 scale) Asked in all waves.	*tij*
Psychological strain – low confidence	GHQ12-10 is losing confidence in themselves (0-3 scale) GHQ12-11 thinking of self as worthless (0-3 scale)		*Absent in Understanding Society data file from UK Data Service.*	
Life satisfaction		*sclfsat1*	Self-reported life satisfaction with: Health. 7-pont scale, inverted so higher scores = lower satisfaction. Asked in all waves.	*tij*
		sclfsat2	Self-reported life satisfaction with: income of household. 7-pont scale, inverted so higher scores = lower satisfaction. Asked in all waves.	*tij*

	sclfsat7	Self-reported life satisfaction with: amount of leisure time. 7-pont scale, inverted so higher scores = lower satisfaction. Asked in all waves.	*tij*
	sclfsato	Self-reported life satisfaction with: life overall. 7-pont scale, inverted so higher scores = lower satisfaction. Asked in all waves.	*tij*
Psychological strain - control	*scloutcont*	Locus-of-control question, 6 point item asking the extent to which the respondent agrees with the statement "what happens in life is out of my control". Higher scores = more disagreement. Asked in wave 2 only.	*tij*
	scwemwbd	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: that they have been dealing with problems well (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	*tij*
	scwemwbe	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: that they have been thinking clearly (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	*tij*

		scwemwbg	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: that they have been able to make up their mind about things (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	*tij*
		scwemwba	Warwick-Edinburgh Mental Wellbeing Scale: In the last two weeks, the respondent has been feeling: optimistic about the future (5-point scale, inverted, higher = worse wellbeing). Asked in waves 1 and 4.	*tij*
Environment – home and local area	Whether the house in which the respondent lives has problems with (0/1): bad light bad heating leaks rotting wood damp Noise from neighbours Crime levels Lack of space Respondent's subjective appraisal of the area Whether the respondent would like to move away			

Environment – quality of local services	*locserb*	Respondent's assessment of the quality of local services, 4-point scale, higher scores = worse services. Asked in waves 3 and 6. Medical services	*tij*
	locserc	Transport	*tij*
	locserd	Shopping	*tij*
	locsere	Leisure	*tij*
	crmugg	Respondents self-reported perception of local crime rates, 4-point scale, higher scores = more crime. Asked in wave 3.	*tij*
Multigroup invariance variables/identified controls	*sex*	Sex of respondent	*ij*
	birthy	Age/cohort of respondent (see discussion below about age/period/cohort analysis).	*ij*
	plbornc	Country of birth	*ij*
	jbes2000	Employment status (recoded into dummy variables)	*ti*
	racel_dv	Ethnic origin	*ij*
	health	Whether respondent is registered as having a disability/longstanding health condition	*ti*
	qfhigh	Highest educational qualification	*ti*
	jbrgsc_dv	Social class (General registrar)	*ti*

The use of structural equation modelling and confirmatory factor analysis requires, *a priori*, a strong hypothesis about the indicators used in latent variable factors, which has been developed both from the literature reviewed and using previous similar research, and is explored using exploratory data analysis (Tomlinson, et al., 2008; Ullman, 2014). Some of the above variables also come from questionnaire inventories, such as the GHQ-12, that are explicitly designed to measure latent constructs (whether using SEM or not), fit parametric assumptions, have high internal consistency and robustness over time, and have been piloted and developed extensively. The variables chosen here include items from some of these questionnaires, including the general health questionnaire (Goldberg & Williams, 1988); material deprivation indicators from the Poverty and Social Exclusion millenium survey (Pantazis, et al., 2006); and the Warwick-Edinburgh Mental Wellbeing Scale (Stewart-Brown, et al., 2011).

One other important point of note is that, although age, represented by birth cohort, is included as a control variable, any inference made from the effects of age should be treated with caution. Age, cohort, and current year, are all perfectly collinear variables and this introduces a phenomenon known as the age-period-cohort (APC) identification problem into models (Bell & Jones, 2013; O'Brien, 2011). There is no suitably robust method of estimating the relative effect sizes of age, cohort membership, and year of time, because there are infinite 'best' solutions – therefore it is important to acknowledge in any analysis that age, period, and cohort effects are interwoven in complex ways. The nature of this relationship can be elucidated over long spans of time by using specific visualisation tools (see Minton, et al., 2017), but a full APC analysis is beyond the scope of the research aims and the data available here.

3.4 Doing (some) justice to theory: Factor Analysis and Structural Equation Modelling

In considering the literature around the definition and measurement of poverty we can observe how, in the journey from concept to measurement, a large amount of nuance is lost (Lister, 2004). The goal of measurement is to approximate, using a single indicator, something that identifies the presence of a specific state of being or the extent to which something is being experienced. This begins with the operationalisation of the concept of poverty (*ibid*): there is no 'ruler' for poverty, so something measurable must be constructed. The aim is to reduce a large number of items to a smaller number of variables using some appropriate commonality. The previous chapter identified some of these attempts. They typically involve first identifying a 'basket of goods', items where ownership or participation can be measured, that meets the criteria of the definition of the concept, whether that is normative participation or a consensual

moral minimum. This creates a large number of items that must then be further reduced to a single indicator. In this case, the basket of goods is approximated simultaneously using a known underlying common factor – income. However, this leaves the question of how to include dimensions of a concept that do not share the same underlying commonality: non-monetary dimensions of poverty such as autonomy or dignity. It is not possible to put a price on these. Therefore, they are not captured explicitly by income-based measures of poverty. This section outlines the fundamental methods and statistical concepts that can be used in establishing a valid and reliable multidimensional measure, and some of these ideas are reiterated within the context of the findings in the next section.

Figure 6 shows a simplification of this process from concept to measurement. Boxes with rounded edges represent concepts or unmeasured 'unobserved' items, boxes with squared edges represent measured 'observed' items. All of the dimensions of poverty can be operationalised, but not all of them can be approximated by a single indicator. Furthermore, there is no systematic way to quantitatively assess how well the final approximation captures the underlying concept.

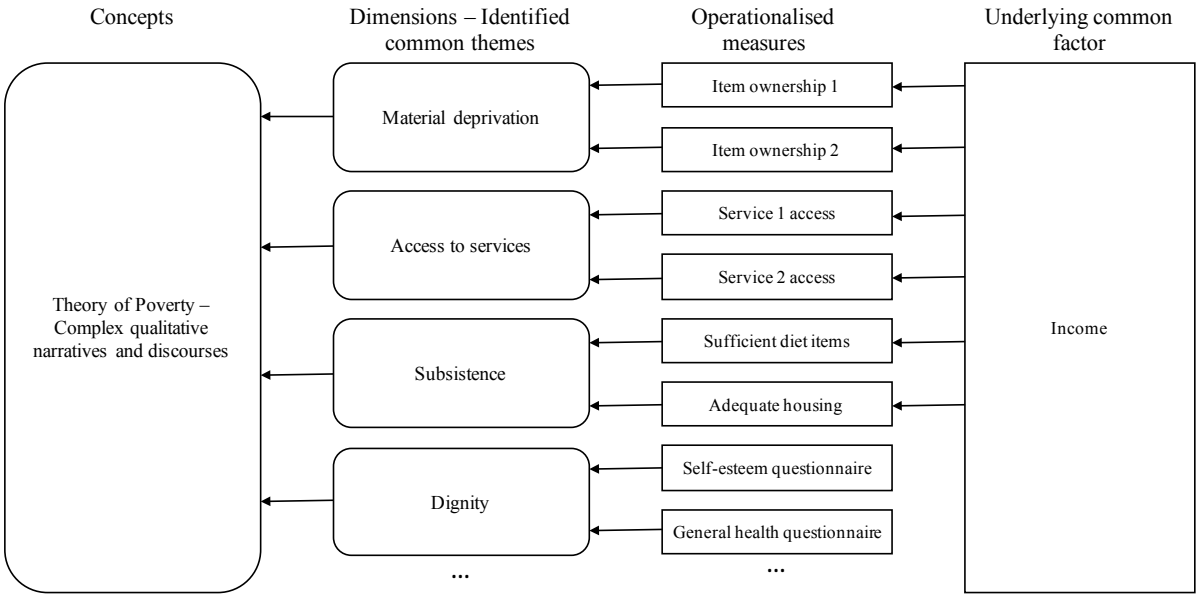


Figure 6: *The process of moving from concept to a single measurement variable using operationalisation and a common measured underlying factor without the use of latent factors.*

Contrast the above figure, where operationalised measures are simultaneously modelled using one known factor, income, to the use of factors and structural equation models offers an alternative to this, shown in figure 7. The concept of poverty still needs to be broken down into dimensions (process a) and then further operationalised into measurable variables (process b)

in order to make the collection of data possible. My theoretical position is that, if all of these 'manifest' variables are good measures of the underlying dimensions, they should share some kind of close correlation with one another. Instead of choosing an existing measured variable that approximates these simultaneously, the purpose of factor analysis is to create a new 'latent' variable that simultaneously predicts its associated observed variables (process c). The latent variable captures their commonality. The commonality of multiple latent factors can also be modelled as a 'second order' latent variable (process d). If the theory of the concept and its sub-dimensions have a good basis in reality, the dimensions have been operationalised well, data collected using operationalised measures minimises potential bias and unintended effects, and the specification of latent variables and the methods used in their construction are correct, the resulting constructs should accurately reflect the underlying concept (Little, 2013). The approximation should reflect reality. The real strength of this approach is that how well the latent variables that are specified reflect their composite indicators can be interrogated using model fit statistics. This approach allows us to include dimensions of poverty that cannot be easily approximated with income.

The benefits of using secondary data are important here. The survey has been developed by expert survey methodologists and administered by interviewers whose training had been heavily invested in using techniques that minimise bias and employ best practice. Items from the questionnaire often come from well-regarded and validated question sets. Therefore, we can have some confidence that the variables in the data have limited measurement error.

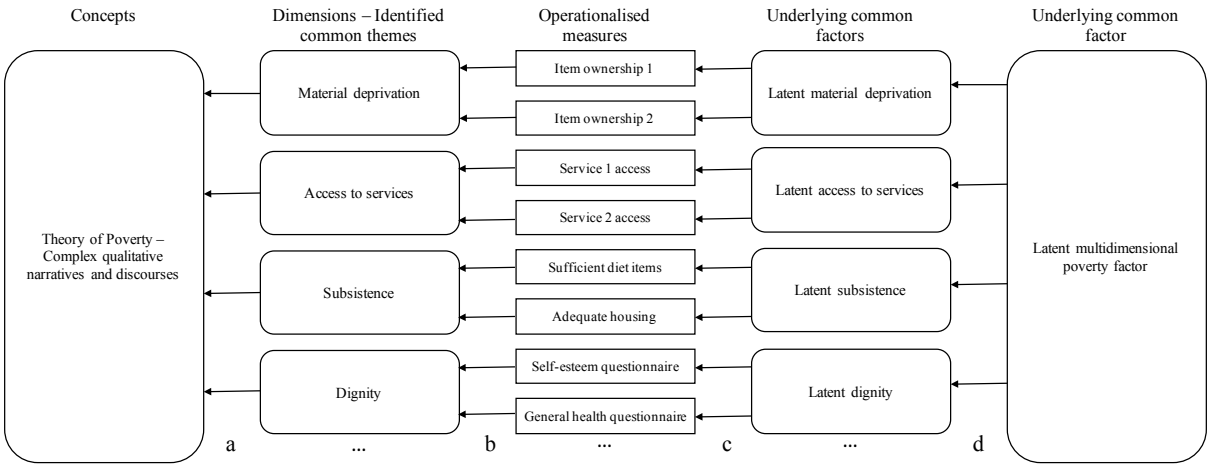


Figure 7: *The alternative process of moving from concept to a single measurement variable using operationalisation and latent factors. Boxes with rounded edges are 'unobserved', boxes with squared edges are 'observed'.*

Why is all of this sophistication necessary in the first place? For one, there is the practical justification for any statistical modelling in trying to piece together some coherent picture of trends from tens of related variables. More importantly, the inclusion of many closely related variables in one statistical model without some form of information reduction introduces the problem of multicollinearity. Multicollinearity occurs when two or more variables are highly correlated and therefore contain redundant information (Tabachnick & Fidell, 2013). Models that contain many variables that share a strong correlation have estimates of effect sizes and standard errors that are unstable and may therefore not accurately reflect the real effect of an independent variable on a dependent variable (*ibid*). This is where the 'independence' of independent variables matters. However, if we are interested in a concept that manifests in many observable ways, we accept and actually want these to be closely associated to one another. This leaves us with the two choices outlined above: reduce our information to one common existing factor or create a new common factor. The latter, used here, is factor analysis. In assessing the value of factors, Kline (2010: 165) notes that not using factors "force[s] the researcher to choose among alternative measures (if available), … basically any single indicator is susceptible to measurement error." Based on the wildly different reported poverty rates by different ways of reducing poverty to income, a factor analysis approach may yield more reliable results.

Tabachnick and Fidell (2013) define the aims of principal components analysis (PCA) and factor analysis (FA, explained below) as follows:

> "*The specific goals of PCA or FA are to summarize patterns of correlations among observed variables to a smaller number of factors, to provide an operational definition (a regression equation) for an underlying process by using observed variables, or to test a theory about the nature of an underlying process."*
>
> Tabachnick & Fidell (2013) *Principal Components and Factor Analysis*, 660-661

This operational definition refers to the fact that if all of the observed variables are highly correlated, they can be all be predicted by the resultant latent variable with varying degrees of accuracy and with associated error term that can be summarised using a regression equation. For example, the relationship between a latent variable (LV) and the three normally-distributed, continuous observed variables ($MV_{1...3}$) it simultaneously predicts can be expressed in three regression equations as follows:

$$MV_1 = B_0 + B_1LV + e_1$$
$$MV_2 = B_0 + B_2LV + e_2$$
$$MV_3 = B_0 + B_3LV + e_3$$

As the latent variables will undoubtedly have some error in their predictions an error term, e_n, is also included.

The structure of such latent variables are typically presented in diagrams such as those in figure 8. Usual conventions (Byrne, 2009; Kline, 2010; Beaujean, 2014) are that manifest or observed (directly measured) variables are denoted by squares or rectangles, latent or unobserved (not directly measured) variables are denoted by circles or ellipses. The relationship between two variables is denoted using an arrow, with the direction of the arrow indicating the direction of the relationship (for example, $x \rightarrow y$, the value associated with the arrow is equal to that of the associated change in y for a one unit increase in x). Arrows from latent variables predicting observed variables are usually referred to as factor loadings. A double-headed arrow signifies the correlation or covariance between two variables. These are also sometimes used on a single variable to signify residual variance. Finally, each variable that is predicted in any way has a residual error term (E_i) associated with it that includes unexplained variance and measurement error. Factors made up of other factors are called second-order factors, factors made up of second-order factors are called third-order factors, and so on.

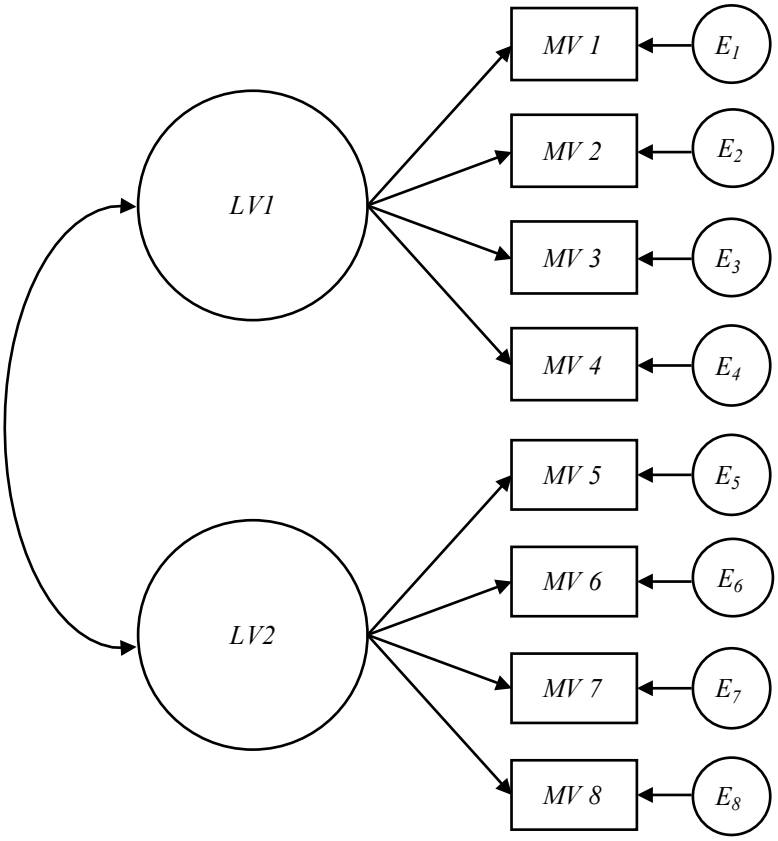


Figure 8: *An example of the structure of two latent variables modelled on four manifest variables each with associated error terms.*

When factor analysis includes specified regression paths between factors, observed variables, or both, such as in regressing a latent factor of poverty on an observed variable of income, it becomes a structural equation model (Ullman, 2014). Structural equation modelling refers then to a family of techniques that include path models and the investigation of differences in underlying factors. Any variable can be considered independent, dependent, or both (a mediator or moderator) depending on the pathways that are specified. Hence the importance of a strong theoretical foundation. This makes the method especially valuable for developing understandings of underlying theories, and testing those theories empirically with quantitative evidence (Tabachnick & Fidell, 2013). It is therefore the core method for investigating all the research questions of this thesis. There are, however, important differences in the way in which factor analysis and structural equation modelling is implemented. While factor analysis techniques can be used to tell us something about the nature of underlying structures, and to test scales or reduce the amount of information while reducing measurement error, SEM is concerned with the application of these factors to hypotheses. Often factor analysis and its associated techniques of reliability and validity testing are hastily passed by in favour of the 'applied' aspects of SEM.

Factor analysis is typically split into two approaches: confirmatory (CFA) and exploratory (EFA). In confirmatory factor analysis a pre-hypothesised structure of variables that represents an underlying factor is specified, factor analysis is then conducted and the researcher observes how well the posited hypothesized arrangement of items fits an underlying factor using model fit indices (Brown, 2015). This places a restriction on the researcher that, if followed closely, prevents them from adjusting the arrangement of items *ad hoc* and deviating from the original theoretical arrangement that is being tested. Deviation from the original structure introduces researcher 'degrees of freedom' that can result in significant results being found that may be the result of random error in the data that are justified *post hoc* and not actually a reflection of how good the underlying factor is, just how well the factor models noise, a problem that is discussed more generally in the following section. Exploratory factor analysis relaxes these restrictions in order for the researcher to more freely explore which variables correspond to one another to find a good arrangement of factors that seems to model the commonality underlying the variables (Tabachnick & Fidell, 2013; Brown, 2015). However, even if the decision making process is not automated as in the methodological definition of EFA, as soon as the researcher changes a hypothesised model to better reflect the data based on the output of some preceding analysis the method's nature changes from confirmatory to exploratory and the generalisation of goodness of fit outside the sample data cannot be assumed.

In practice, most factor analysis ends up exploratory in nature when proper safeguards are not put in place. If a researcher is using their entire dataset in their analysis and has to change one small parameter because of a mistake it can be very easy to accidentally violate the requirements of confirmatory factor analysis. The unexpected happens more frequently than one might expect. A model may converge with an illogical estimate for the correlation between two variables, for example, giving a standardised coefficient larger than 1 or a residual variance of less than 0. In this case, the researcher would want to impose a constraint on this path so that the estimate could not exceed 1 or be lower than 0. Because this was not hypothesised prior to the analysis the research now becomes exploratory, and cannot confidently be extrapolated beyond the sample data. Similarly, it is common for there to be significant correlated residual measurement error between one or more items that may not have been predicted and negatively affects the fit indices of a factor when not estimated and should be built into the model. It is unclear at that point whether this error is a true underlying feature of the measurement or noise specific to the given sample, but the consequence is that the researcher can no longer claim the

benefits inferred from confirmatory analysis. A pragmatic safeguard to this problem is to build and check models preliminarily using a subset of data designed specifically for exploratory data analysis (discussed in the following section) before using an additional subset, or the complete data set, for confirmatory factor analysis to check that any tweaks to models are not made to artificially inflate the goodness of fit.

One weakness of income-based measures of poverty already mentioned is the inability for researchers to gauge how well certain levels represent the proposed theoretical foundation. Factor analyses and structural equation models are evaluated using various fit statistics. The full history and differences between fit indices is too complex to detail here (see, Bollen & Long, 1993). Generally speaking, most fit indices that are found in contemporary use of the method fall into one of two categories: incremental fit indices and comparative fit indices, but there is also one other measure that is referred to as an absolute fit index (Kenny, 2014). Their purpose is to compare the complete explanatory power of the reduced model compared to a complex null model with many parameters (*ibid*; Bollen & Long, 1993).

In the case of incremental fit indices the alternative is the worst possible model, which is typically a model where all variables are considered to be independent and unrelated to one another. This is sometimes called a baseline model. The 'best' possible model in this case refers to a model that perfectly reproduces all the variances, covariances, and means of the variables included (Tabachnick & Fidell, 2013; Kenny, 2014). This is achieved by considering a 'saturated' model, a model where all possible covariances between variables have been included in the model and the maximum number of parameters has been estimated. This is the 'best' model only in the sense that it is able to perfectly replicate the variances and covariances between the variables, they do not usually add to our understanding of the relationship between variables. The logic behind these fit indices is that a well-specified model should accurately reproduce the variances and covariances in the underlying data while being as parsimonious as possible, similar to an adjusted R^2 statistic. The closer a model comes to replicating the underlying structure of the data the closer the value of an incremental fit index approaches a value of 1. Incremental fit index statistics have their own commonly agreed upon values that represent 'poor', 'mediocre', and 'good' fit, and different indices have different penalties for model complexity. Because of this more than one incremental fit index is usually reported (typically the Comparative Fit Index, CFI and the Tucker-Lewis Index, TLI).

While incremental fit indices are expressed on a standardised scale between zero and one, comparative fit indices are only meaningful when compared to other models (Kenny, 2014). A smaller comparative fit index value indicates a better fitting model, so the criterion score for two models can be compared to identify which model fits better. Comparative fit indices apply a penalty for the number of parameters specified, which makes them useful in exploratory factor analysis and exploratory structural equation modelling to assess which additional parameters (such as specified relationships between variables) do not add sufficiently to the explanatory power of the model to justify the additional complexity. This can be useful when deciding how much complexity should be added when estimating random effects. For example, if an additional parameter is estimated in the model and the chosen comparative fit index for the resultant model increases rather than decreases there is support for the argument that the additional parameter should not be included. The two most commonly used comparative fit indices in reporting SEM model fit are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Beaujean, 2014).

One other statistic for assessing model fit is the 'Root Mean Square Error of Approximation' (RMSEA), which is a measure of absolute fit (Kenny, 2014). These indices are operationally similar to incremental fit indices, with the key difference being that the best fitting model has a score approaching zero, rather than one. The RMSEA is based on sampling error, and has been found to not be reliable for small *N* and small *df* models, so should not be reported for these (Kenny, et al., 2015). Overall, a CFI of greater than 0.95, a TLI of greater than 0.9, and an RMSEA of less than 0.08 are usually indicators of good fit.

One other fit index that is commonly used is the chi squared goodness of fit test. A lower chi square reflects fewer differences between the variance-covariance matrix reproduced by the model and the full data's variance-covariance matrix, and therefore a better model. However, the chi square test for SEM and FA is almost always useless as a test of goodness of fit because it is so sensitive to sample size (Kenny, 2014). A sample size of 400 or more will almost always lead to a statistically significant chi square, which indicates statistically significant differences between the matrices. As a comparative fit index, the chi square can be calculated for one model and a *p*-value can be calculated for the change in the chi square with the degrees of freedom being equal to the change in the number of parameters as models become iteratively more complex. If the reduction in the chi square value between a simpler model and a more complex model is statistically significant, there is support for retaining the more complex

model. Nevetheless, the chi-square test as a measure of model fit is not very instructive for research using survey data because of the aforementioned sample size sensitivity.

The use of model fit statistics is central to answering all of the research questions outlined above. These research questions involve constructing competing models, for example, alternative arrangements of items for factors or different models for different groups or time points compared to a pooled model. To answer these questions, model fit statistics for alternative models are compared at the exploratory stage and then validated at the confirmatory stage. If distinct models for distinct groups produce significantly better model fit statistics this tells us something about how poverty is differentially experienced. For example, a pooled multidimensional poverty index might meet the criteria for a 'mediocre' fit, but separate models for men and women may each produce 'good' fits, if the factor loadings for each dimension are different. This iterative comparative approach to exploratory factor analysis is key to investigating, for example, whether the same multidimensional structure for measuring poverty holds consistent over time, or whether there is evidence that the definition and underlying structure changes significantly – whether certain dimensions become differently associated with the underlying factor. This is achieved by imposing constraints on models, for example, fixing factor loadings equal over time, and comparing the model fit index for these models compared to those where the constraints are released and allowed to be estimated freely (Little, 2013). This is typically done in invariance testing. This flexibility enables the exploration of the research questions above.

Factor invariance testing

Factor invariance testing (Widaman, et al., 2010; Little, 2013) is typically associated with longitudinal structural equation modelling but is often treated as a purely diagnostic tool relating to the reliability of measurement variables that are – in theory – supposed to be invariant over time and between groups. However, this research argues that results from this kind of testing can be instructive to our understanding of how underlying constructs change over time or differ between groups, rather than being purely diagnostic. Townsend (1979) argues that poverty can only be understood as a relative and dynamic concept. By this definition we would expect that the construct of poverty, here measured by factor loadings on multidimensional factors, is not necessarily fixed over time or between different groups, given large enough social changes, different enough groups, or long enough time spans. Factor invariance testing can help investigate whether this appears to be the case. However, the

number of years explored here is limited. This thesis aims to identify whether the dimensions of poverty are fixed over a six-year period after the 2008/9 recession or whether each ones' relative weight within an index has changed. There are limitations imposed by the availablity of certain variables in secondary data. This research relies on being able to structure factors using the same measures each year but, because of the discontinuation of the British Household Panel Survey in 2008 and the inconsistency of the wide number of variables listed above carried across from the BHPS to *Understanding Society*, it is not possible to construct factors with the desired level of complexity and comparability over a longer span of time. However, this thesis represents evidence that such a task can be undertaken, and such evidence might hopefully encourage administrators of large surveys to include multidimensional poverty related variables more frequently in future waves, so that the process undertaken here can be repeated in future with data that is higher in quantity and quality.

While these tests are used to assess reliability, a different type of test is required to measure validity. While reliability refers to the extent to which a measure or approach can be used to produce comparable results regardless of the group or context in which it is applied (Jenkins, 2002), validity refers to whether the measurement or construct has some legitimate grounding in real life (*ibid*). This is why reliability is typically associated with quantitative research, as there are finite responses to questions and standardised delivery of prompts, and validity is typically associated with qualitative research, as the relationship between researcher and participant, and the depth of the methods employed, enables far richer data grounded in the lived experiences of the participant (Bryman, 2012; Mason, 2002). However, quantitative research methods do have some test criteria for validity through the use of factor analysis which is based on the comparison between different constructs.

The most widely used method, and the one used in this research, is the Fornell-Larcker test of discriminant validity (Fornell & Larcker, 1981). This test relies on the idea that something can be labelled as a validly measured construct if it is simultaneously related to other constructs it is theorised to be related to while also being unique enough to be distinguishable from these (Cambell & Fiske, 1959; Fornell & Larcker, 1981). These ideas are referred to as *convergent* and *discriminant* validity. The Fornell-Larcker criterion is that the squared correlation between two related constructs should not exceed the 'Average Variance Extracted' value of a construct, which is calculated using the following formula:

$$AVE = \frac{\sum_{i=1}^{k} \lambda_i^2}{\sum_{i=1}^{k} \lambda_i^2 + \sum_{i=1}^{k} Var(e_i)}$$

Here, $\sum_{i=1}^{k} \lambda_i^2$ refers to the sum of all of the squared factor loadings for a given latent variable, and $\sum_{i=1}^{k} Var(e_i)$ refers to the sum of the variance of all of the residuals for each factor loading in a latent variable.

This formula gives what can broadly be interpreted as the proportion of all of the variance of the manifest variables (indicator items for a factor) that can be explained by the underlying construct that the factor analysis has produced. If a comparison of two factors does not meet the Fornell-Larcker criterion either the theory that they are distinct concepts should be revised, the conclusion should be drawn that the choice and design of indicators measuring the underlying construct was not sufficiently precise, or both. For example, the use of the test might establish whether financial strain, as a construct, is meaningfully distinct from material deprivation. As the empirical grounding for valid constructs is largely based on very valid qualitative research and many of the variables used in this research have not been explicitly designed to model underlying factors it is considered best practice here to assume that poor discriminant validity between two similar constructs is likely due to inadequate measurement.

One weakness specific to SEM is that the value of a given latent variable is not always easily interpretable. Latent variables are, unless otherwise specified, normally distributed with a mean of zero and a scale bounded by the reference marker variable. The reference marker variable is usually just the first variable specified in the model code so, for example, if it is a variable with a range of responses between 1 and 5 the consequent latent variable will be on a range between 1 and 5. Therefore, the actual values of the scale are only meaningful in the context of the manifest variables that form the factor. There is some interpretation possible in that higher scores will be associated with worse outcomes because of the way that manifest variables have been recoded. Although the additional complexity that can be captured with SEM and factor analysis is great, there are clear limitations in the availability of data and the restrictions imposed by the parameters that shape factors' distributions. For example, it may be that the underlying distribution of multidimensional poverty is skewed, in the same way that Ferragina, et al., (2013) identify a participation 'floor' in the association between participation and income. Adjusting the algorithms to change restrictions on factor formation to allow for skewness, however, requires specialist technical skills beyond this project (Liu & Lin, 2014).

However, there have been some advances in the implementation of recent developments in structural equation models and factor analysis in user software which may result in more accurate estimates compared to those made by Tomlinson, et al. (2008). This research uses the software *Mplus8* (Muthén & Muthén, 2017) to conduct factor analysis models through the *R* statistical programming language (R Core Team, 2016) in the *RStudio* (RStudio Team, 2018) integrated development environment (IDE) using the *MplusAutomation* package (Hallquist & Wiley, 2016). For example, *Mplus8* now allows users to specify the type of manifest variables to allow for ordinal and categorical dependent variables that use logistic regression in factor formation, as opposed to relying on the assumptions of linear regression to predict ordinal variables and categorical binary variables (using linear probability models). *MplusAutomation* enables models to be read into the *RStudio* environment as *R* objects, which aids better visualisation, exploration, and transcription of results, as well as allowing for parsing .init scripts for iterating through models with changing constraints without having to write multiple scripts by hand. The result of this flexibility is reflected in the large number of models in the appendices and the range of visualisation used for testing group invariance in the following chapter. Testing of this depth would not have been possible in such a timespan without drawing on the programming functionality in *R*.

3.5 Exploring data: workflow and visualisation using random subsets of data

The ethos behind the workflow used here is heavily influenced by the idea of applied factor/principal components analysis and statistics more generally as data reduction or, more broadly, information reduction, which takes the core principles of quantitative methods and integrates the associated limitations and logic of data processing and communication. All models, theoretical or statistical, qualitative or quantitative, involve reducing a large amount of information to a more manageable amount of information to aid cognition and communication. Regression models reduce variance to the expected value of a variable conditional on several independent variables (Schroeder, et al., 1986). Latent variables reduce multiple indicator measures to one continuous predictor (Ullman, 2014). Recent contributions to the field of data science have developed along these lines, where the focus on visualising and exploring patterns in data before introducing statistical tests and models has been more formalised in the pedagogical approach (Grolemund & Wickham, 2016). However, this way of doing things is not new. The strong recommendation of visualising data before building models is a feature of all undergraduate quantitative methods textbooks (e.g. Bryman, 2008;

Field, 2013; MacInnes, 2017) but is often treated as purely prescriptive and used for testing assumptions rather than for establishing strong theoretical justifications for models. Rather, the need to visualise here is grounded in John W. Tukey's (1977) ideas of *exploratory data analysis*, and is more closely related to a set of principles rather than prescriptions. In his words, " '[e]xploratory data analysis' is an attitude, a state of flexibility, a willingness to look for those things that we believe are not there, as well as those we believe to be there." (*ibid*: 806). "Exploratory data analysis… does not need probability, significance, or confidence." (*ibid*: 794).

Why are these principles important? Although there are good reasons to explore data for its inductive benefits - it allows researchers to "isolate patterns and features of the data and reveals these forcefully" (Hoaglin, et al., 1983) - there are modern practices in quantitative methods that add to the value of exploratory data analysis. Despite the growing sophistication, accessibility, and computational power of statistics computer programs, comparatively more attention has been directed towards identifying 'statistically significant' relationships in data and less effort has been directed towards improving and routinising the exploration of data using visualisation (Gelman & Loken, 2014). This is, to some extent in the field of sociology, a result of a fundamental misunderstanding of *p*-values as representing some kind of absolute truth as opposed to their actual function of testing the likelihood of a test statistic of the given magnitude being found under different (null) conditions; conditions that may be meaningless or unrealistic in the context of the research (Greenland, et al., 2016). Such an approach is incentivised by the tendency of scientific journals to publish only quantitative research where 'statistical significance' in the form of *p*-values smaller than the threshold of 0.05 can be demonstrated (Nosek, et al., 2012).

Practices that emphasize finding significance are sometimes referred to as '*p*-hacking' or 'fishing expeditions', however, critics of these definitions point out that such practices are not necessarily deliberate or malicious (Gelman & Loken, 2013). Rather, Gelman & Loken (2013; 2014) argue that quantitative researchers are faced with a 'garden of forking paths'. In a world of all the possible hypotheses that can be constructed from a set of data a certain proportion of these in any sample will be false-positives – significant relationships in the data that do not actually exist in the world and are the result of random noise. The researcher is faced with a large number of decisions that can be made throughout the design of a study: in data collection and data tidying, their choices of hypotheses to test, the number of cases that are excluded,

which variables are included and what measures are used, how missing data is treated, what extraneous variables are included in models, what interactions are included in models, what causal pathways are specified in path models, what kind of models are fit and what assumptions they make about the data, are just a select few choices that are made. The more models that are fit to the data, then chopped, changed, or adjusted before being refit *ad-hoc*, the higher the chance of finding a false positive. Gelman & Loken (2013; 2014) refer to this process of taking many undocumented twists and turns as akin to traversing a garden of forking paths. The journey involved is reported as being premeditated and deliberate, but in actuality there are many choices that create divergent paths that a reader is not made aware of.

The recommendations for dealing with these problems are as follows:

- Ensure the preregistration of any hypotheses and integrate this into research design and the operationalisation of concepts. Make explicit the data collection and analysis protocols in advance of the study (Humphries, et al., 2013). Pre-registration in this instance is not possible as the research uses secondary data analysis.
- Be explicit and transparent about the choices made in the research process, show steps taken with reproducible code, and reflect upon how these might lead to different results. (Gelman & Loken, 2014)
- Sufficiently ground statistical models in existing theory, especially qualitative literature where complex causal mechanisms and relationships between factors can be better explored. State theory prior to modelling and avoid inventing theory *ad hoc* to explain statistically significant findings. (Brady & Collier, 2010; Tabachnick & Fidell, 2014)
- When exploring relationships in the data avoid relying on tests of statistical significance. Instead, visualise relationships and investigate any clear patterns that emerge (Grolemund & Wickham, 2016). When combining exploratory data analysis and confirmatory data analysis, create a subset of data for exploratory analysis and retain another subset of data to validate any findings (Ullman, 2014). This both helps the researcher cross-validate any models and helps avoid overfitting, where complex models are just capturing relationships in random noise.

The similarities between some of the strategies undertaken should seem familiar to those engaged with qualitative methodological discussions. This refers to aims of making known the ‘implicit epistemological aims of your argument’ (Mason, 2002: 176). By engaging with and

critiquing the ways in which your own views and approach may shape their research a stronger, more productive, contribution to knowledge can be made. Data analysis is an active, not passive, process. There are opportunities for quantitative research, with its claim to objectivity recently questioned, to learn from qualitative research where effective argumentation has been fundamental. As Jenkins (2002: 9) argues, 'for a long time we seem to have been mixing up two rather different understandings of objectivity, one to do with our politics and values, the other with our knowledge'. It is possible to communicate clearly the steps that have been taken to mediate the relationship between our values and politics and our epistemological approach. These reflections are weaved through the discussion of findings.

Figure 9 shows how such an approach is, to an extent, encapsulated in common workflows within data science, with the real strength of this approach being the flexible nature of plotting packages such as *ggplot* (Wickham, 2009), and data manipulation packages such as *dplyr* (Wickham, et al., 2017). To bring this in line with the recommendations above, this thesis uses an exploratory subset of data for visualising data, model fitting and building. Modelling, in particular factor composition, is informed using both theoretical contributions from Townsend (1979; 1985), Baulch (1996), Lister (2004), and Spicker (2007), and comparisons between groups and across time are informed by the body of literature outlined earlier in the thesis.

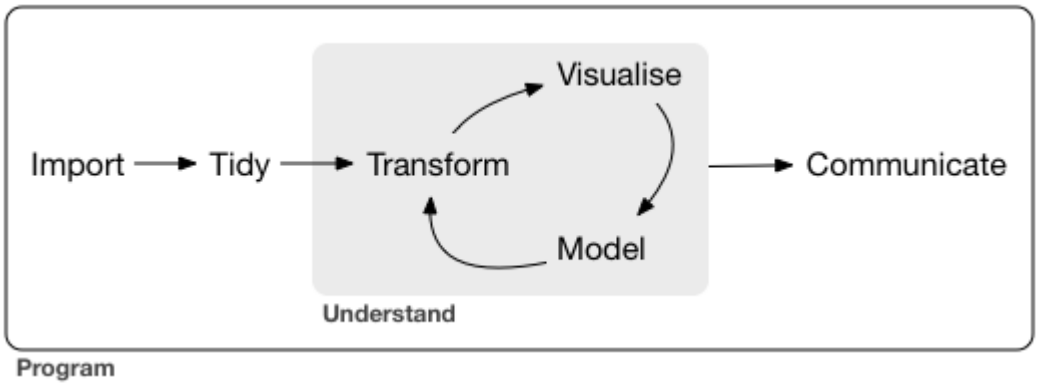


Figure 9: *Tidy approach to data science. (Adapted from Grolemund & Wickham, 2016: online)*

Using the tools outlined above it is possible to more flexibly visualise the complex relationships between variables that are often features of SEM, to identify what variables appear to be more closely related and which are less. Without the flexibility of packages such as *ggplot*, these relationships would usually have to be explored using many visualisations of pairs of variables. This is cumbersome and obscures any larger patterns. Presented in appendix 2 is a visualisation of the strength of nonparametric association that each variable shares with every other variable in the dataset. The colour represents the direction of the association (green is a positive

relationship, blue is a negative relationship), and the transparency of the colour reflects the strength of this association (closer to white is closer to no association). Here, association is determined using a nonparametric correlation coefficient (Spearman's Rho). Alone, this is not an impressive feat. However, using the native functionality in *R* and the flexibility of *ggplot*, the variables are ordered in such a way that maximises the clustering of variables based on their association with others. They are coerced to be ordered more closely to those they are similar to and simultaneously further from those they are dissimilar to using hierarchical clustering to sort the axes (Everitt, 1974; R Core Team, 2016). This aids the explorative exercise of identifying interesting patterns in the data and potential clusters of variables that can form factors in structural equation models (as seen by the annotations in appendix 2.2).

3.6 Data structure: consequences of mixed membership multilevel models

What follows is a discussion of the hierarchical clustering and its importance in modelling approaches. Unfortunately, it was not possible for a multilevel model to be fitted to such a complex set of higher-order factors, and this stands as a limitation of the research discussed in the conclusion of the thesis. Data structures are, however, an increasingly essential part of quantitative analysis and their impact is outlined below. In an effort to retain transparency and make explicit the weakness of any analysis being more valued than giving the illusion of a research project with no problems in practice, it was decided to retain this section rather than omit it. Because of the very high factor loadings for many latent constructs it is unlikely that the corrected standard errors for multilevel models would change the substantive interpretation of the findings significantly, especially as this research does not rely on tests of statistical significance. However, the following details why multilevel structures are important and it remains a weakness of this research that the technological capacity was not available to pursue such a model.

Models to address the hierarchical structures found in data have been essential in advancing quantitative methods. In their most simple form, hierarchical models take into account the clustering effects that higher level units of analysis can have on lower level units by partitioning the variance of dependent variables between the levels (Heck, 2015; Robson & Pevalin, 2015). This means that, when appropriate, group-level variance can be separated from individual-level variance. Taking a commonly used example, a researcher may wish to look at the effect of certain interventions on schoolchildren's reading scores while controlling for the confounding group-level effects of the class they belong to. When the multilevel structure of data is not

controlled for sometimes vastly different inferences can be made, often from a phenomenon called 'Simpson's paradox' (Gelman, et al., 2007). The differences in conclusions can sometimes be severe. In testing the claims of economists that GDP growth was negatively associated with country debt level, Bell, et al., (2015) found that this was only the case when time point levels of each variable for each country were pooled together and their country membership was ignored. When instead using a multilevel model where the clustering of time points within countries was taken into account when estimating the relationship between GDP growth and debt ratio for each country, there was no significant general association across all countries. The consequences of the original paper, the flawed Reinhart and Rogoff model, had become the economic justification of austerity policies across most of Europe and North America.

There are also technical consequences of not using a multilevel model for nested data (Robson & Pevalin, 2015; Heck, 2015). If data is in 'long' format, that is where each row of data represents the lowest level unit of observation, (e.g. row one refers to person one at time point one, row two refers to person one at time point two, and so on) the number of observations (n) is equal to the number of time points multiplied by the number of participants. This n is used to calculate a range of inferential statistics and confidence intervals to test whether effects and associations between variables are statistically significant. The higher the n, the smaller the standard error associated with a given statistic. However, depending on the structure of the data, the n for variables that are at higher levels will be artificially inflated, giving a false sense of confidence (*ibid*). For example, imagine a research project is designed to investigate differences in private and comprehensive school pupils' achievement. In this hypothetical project, data is collected on five hundred students from two schools, one comprehensive and one private. Five hundred students would be considered a large sample. However, there very clearly are not five hundred observations of our variable of interest: whether a school is private or not. There are only two. A regression model that does not take account of the nested levels would estimate the effect and confidence levels of school type as if there were a sample of 500 schools, that is to say the standard error will be greatly underestimated.

The lowest level unit of analysis in multilevel models up to three levels are usually denoted with the subscript i, with higher level units being ascribed the subscripts j, and k. For longitudinal multilevel models, t is usually used instead of i for the lowest level. Variance or residual variance at each level is usually denoted by the latin letters e, u, and v, or the greek

letters ε (epsilon), υ (upsilon), and ν (nu) (Robson & Pevalin, 2015). In many cases, information on group membership, such a neighbourhood, in secondary data is not readily accessible in order to protect survey respondents from identification. In any case, there is a clear hierarchical structure to the *Understanding Society* data used in this research. Time points (t) are nested within individual respondents (i), who are nested within households (j). Time points that belong to the same individual are more likely to be similar than time points belonging to different individuals. Individuals that live in the same household are more likely to have similar responses to one another. Variance at each level is what is usually trying to be explained, but is interesting in itself. Some phenomena may vary more at the time-individual level than at the individual level, for example, a significant proportion of the variance in a person's health might be dependent on the time point that it is measured at (e.g. if that time point is associated with older age), and functions of change over time can be fit to try and explain time-specific variance, but some may be attributable to their individual genetics (Grimm, et al., 2017). Being able to identify the extent to which dimensions of poverty vary differently at each level would answer fundamental questions about the nature of poverty; the extent to which its dimensions are transient, persistent or scarring, as well as the extent of intra-household poverty.

Specifically, the data structure within *Understanding Society* is cross-classified with multiple membership – time point responses to questions are nested within individual participants and within households, and time-individual observations can be nested within multiple households. On first consideration it would make intuitive sense that the structure of the data is strictly hierarchical, with time points nested within individuals that are nested within households, however, this is not the case, because households are not static entities over time (Murphy, 1996; Steele, et al., 2013). Figure 10 shows that although time points (t) are all nested within the same hierarchical unit of person (ti), they can be nested within multiple households (tj) and there is no such thing as a time invariant household. The nature of the 'longitudinal household' is often ignored in favour of a strictly hierarchical solution, such as using the household membership from the first wave of observations for defining membership for all subsequent waves, but the greater the time span that the research covers the poorer estimation of the models will be when partitioning variance accurately to the household level. Steele, et al., (2013) present a way of modelling longitudinal household measurement weighted by adult members of households, but their methodology has yet to become standard practice and could not be implemented here, in part because of the complexity of implementing the method and in part because of the complexity of the structural equation models in this thesis.

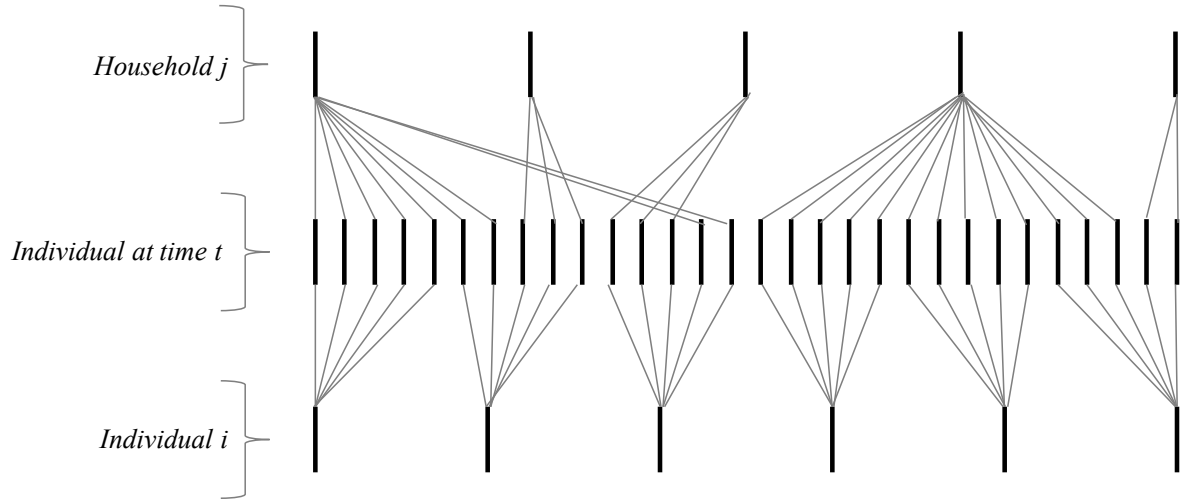


Figure 10: *Diagram showing a multiple membership cross-classified data structure. Each observation of an individual at a given time point belongs to only one individual, but the household membership can change over time.*

Understanding Society recognises in its design that household composition changes over time, so in each wave households are given new household identifier (Knies, 2017). The survey does however include an identifier for the reference person within each household, usually the member of the household responsible for the accommodation, which can be linked across waves. When a child leaves their parents' household, or a partnered couple split up, they are given their own new household identifier (with some exceptions, such as participants that joined the sample as a partner of an original participant but later left because their partnership dissolved). This leaves two possible options for modelling the hierarchical dependencies within the data structure:

- Use the wave household identifier to classify membership to a household which assumes households change in composition at least once between each wave

or

- Use the head of household identifier to identify some kind of household membership over time, with mixed membership to new households in the case of children leaving home or family composition changing (divorce, separation, and so on).

Neither of these solutions is ideal. The first would assume that there are no dimensions of poverty that manifest as persistent features of families and households, however, there is clear evidence of intergenerational transmission of poverty and disadvantage, especially in terms of wealth inheritance (Bradshaw, 2016; Hills, 2015). The second appropriately allocates

membership to survey participants who move between households but were not previously the household reference person, but does not adequately allocate new membership for changes in household composition for the household reference persons. This also means that the longest running 'households' will be referring to men. Although still problematic, the second option may be useful for defining some kind of 'longitudinal household' for the study of poverty, and was used to derive longitudinal household membership here. This should highlight any interesting partitions of variance at the household level, but may not be as accurate a representation of longitudinal household clusters as a more complex method such as the one Steele, at al., (2013) suggest provides.

3.7 Missingness: complete case analysis, multivariate imputation by chained equations, and full information maximum likelihood estimation

Data from complex survey designs often contain missing values, that is, cases where some variables do not have a 'valid' response either because the question was not answered or asked (Allison, 2001). This is especially the case when survey designs are longitudinal, or contain questions that could be considered sensitive, or are worded in such a way that makes participants disinclined to provide an answer. Ways in which to deal with missing data have grown in complexity and usefulness in the past two decades, but there are still common misconceptions that means the standard method of using complete case analysis is not considered as unproblematic as it should be (Schafer & Graham, 2002; Little, et al., 2014). The consequence of this is that much research 'deals' with missing data passively, using default procedures, without proper consideration for the way in which this can introduce bias into the data. The treatment and documentation of missing data should instead be principled and transparent (Schafer & Graham, 2002). This short section will firstly address the standard three ways in which data can be missing and the ways in which 'not missing at random' patterns can be introduced into the data by typical responses to missing data, and why this is particularly important for sociological data. Then, this section presents an accessible overview of the most common applied methods for addressing missing data, multivariate imputation by chained equations and full information maximum likelihood, and evidence for their effectiveness. Finally, problems of missingness unique to the Understanding Society data are discussed and the ways in which these missing data have been accounted for are reported.

The reason that missing data has to be considered so carefully in this research is because it has the potential to cause very real disruption to the precision of the final models for two reasons.

Reason one is that structural equation modelling and the use of latent variables for capturing multiple dimensions of poverty necessitates the use of a large number of variables that are then transformed into a smaller number of factors (described later in this chapter). The other reason is that questions related to poverty, as previously discussed, can be highly sensitive and personal, stigmatising, and may induce shame in participants which can result in non-response (Walker, 2014). Both of these factors means that the default way of dealing with missing data that is used by many statistical procedures, listwise deletion, can remove observations in a way that introduces bias.

Statistical procedures are not designed to work with data that has missing values, but many software environments will change the data frame that is being used in some way so that these procedures can be run (Allison, 2001). Listwise deletion refers to the most common way in which missing data is dealt with, and involves removing from the analysis all observations (rows) where there are missing values in one or more of the variables being used in the analysis (see figure 11 for explanation) (Bryman, 2008; Field, 2013). Because of this the number of observations removed using this method often increases exponentially with the number of variables included in the analysis.

ID	Var1	Var2	Var3
1	NA	20	NA
2	1	NA	NA
3	0	23	1
4	1	NA	0
5	0	46	1
6	1	72	0

→

ID	Var1	Var2	Var3
~~1~~	~~NA~~	~~20~~	~~NA~~
~~2~~	~~1~~	~~NA~~	~~NA~~
3	0	23	1
~~4~~	~~1~~	~~NA~~	~~0~~
5	0	46	1
6	1	72	0

Figure 11 *Rows of data with information for other variables are removed based on missing values within a data frame with listwise deletion.*

There are three recognised mechanisms by which data can be missing, originally put forward by Rubin (1976). Data can be either 'missing completely at random' (MCAR), 'missing at random' (MAR), or 'missing not at random' (MNAR). MCAR refers to a missing data mechanism where the pattern of missing data is completely random (Rubin, 1976; Little, et al., 2014). Paradoxically, this is both the least likely form of missingness to stem from an uncontrolled environment but the only form of missingness where listwise deletion will not bias model estimates (Little, et al., 2014). Data that is MAR refers to missingness that can be

predicted by some extraneous variable or set of variables that is associated with the missing values of the variable of interest, but that missingness does not fall on one side of the variables' scales more than the other (*ibid*). For example, self-employed people may be more likely to have missing data on net income than employed people because of the routinisation of payslips and end of year earnings forms in formal employment. When questions are well designed and piloted, MAR patterns in missing data are the most common cases of missing data, and introduced bias and lost power can be rectified using modern methods of missing data imputation and estimation (*ibid*). MNAR refers to missingness that is a result of respondents' levels of a particular variable, typically when very high or very low responses may be seen as undesirable, such as in the case of cigarette smoking or income levels (Allison, 2004; Little, et al., 2014). This can only be avoided with careful survey design and piloting, however, there is evidence that recent advancements in the treatment of missing data can recover some of the bias that MNAR can introduce (Collins, et al., 2001). One way in which types of missingness can be tested, in particular, whether missingness in one variable is associated with missingness in other variables, is by visualising the strength of the correlation between missing responses in each variable. Appendix 1.1 through 1.6 show the patterns of missingness dependency within the *Understanding Society* data. Interrogating this visualisation reveals that missingness is heavily dependent on whether a question is nested within an inventory of questions – refusal to answer one of the items likely results in refusal to answer any, or the interviewer may seek approval from the respondent about the entire series of questions as opposed to each individual one.

The history and future of techniques to deal with missing data is too detailed to outline here. Suffice to say that early forms of imputation, such as single imputation or mean-replacement, often did more harm than that mentioned by listwise deletion (Allison, 2001). Emerging forms of data imputation that are yet to be widely adopted include the use of classification and regression trees (CART) and multilevel structures to impute missing data based on localised models for highly clustered variables and observations which offer more accurate imputation when missing data does not meet parametric model assumptions, but these methods are extremely computationally intensive and require a certain sized sample to be reliable (Loh, et al., 2016). This research uses van Buuren & Groothuis-Oudshoorn's (2011) R package for multivariate imputation by chained equations (MICE). Simply put, MICE attempts to impute missing values in the data by cycling through predictions of the missing values of a given variable based on global models that use all other variables as predictors, or a selection of

variables chosen by the strength of their correlation coefficient (Azur, et al., 2011). The model for each variable is specified by the user, for example, a logistic regression model would be used to predict the missing values for a binary variable, and MICE typically performs well when the assumptions underpinning these models are met, especially when predictive mean matching (a semi-parametric method) is used (White, et al., 2011). The process is repeated for ten to thirty cycles and with several imputations of the data, usually three to five, at which point the mean and standard deviation for each of the imputed datasets is visually inspected to check for convergence in the missing data imputations (Azur, et al., 2011; van Buuren & Groothuis-Outshoorn, 2011).

While MICE creates several new datasets with imputed missing values, full information maximum likelihood (FIML) estimates model parameters (means, variances, etc.) using all information that is available in the dataset (Dong & Peng, 2013). Model estimates are usually shown to be unbiased when the underlying assumptions of normality are met, which should be a feature of any SEM models in any case, and these often have lower standard errors than equivalent MI approaches (Schafer & Graham, 2002). This approach typically works well for MCAR and MAR data, but there are weaknesses where certain models, especially in model building or diagnostic exercises, may have a small number of variables making estimation under FIML poor; in some cases specific variables may have high counts of missing values making model convergence (explained in the following section) fail frequently. Furthermore, FIML requires parameters to be estimated for each model run, whereas MI can be completed prior to running models. This makes FIML more time-intensive when running a large number of complex models, as is the case here.

As FIML is such a time-intensive solution to missing data, this leaves one further consideration for multiple imputation by chained equations: whether to impute estimates for a scale construct after fitting a model or whether to impute item estimates before constructing a scale. Gottschall, et al., (2012) compared bias and efficiency in item- versus scale-level multiple imputations in a planned-missingness design. Bias, in this instance, refers to the tendency for estimates to be consistently skewed away from the true estimate; efficiency, in contrast, refers to the error of the factor estimates – including their means, correlations, and regression coefficients. They found that no bias was introduced using the different approaches to scale score estimation, but that efficiency was substantially improved when missing data was imputed at the item-level. They note that accurate item-level imputation can be difficult with smaller samples and limited

survey items, however, as previously noted, *Understanding Society* has a very large sample size and a large number of variables. For this reason, missing data was imputed for items first and then scales were constructed from these items using factor analysis.

This provides an opportunity to explore the same models using different approaches to missing data to observe how the inference might change depending on how missing data is treated, and how sensitive different imputations of missing data might be. However, due to the complexity of the models, it is unlikely that any approach other than MICE will yield results, meaning that engagement with and understanding of the weaknesses of MICE is essential to this kind of research. Comparable models that show parameter estimates (the size of effects, variance, and model fit) for listwise deletion, five multivariate imputations of missing data by chained equations, and full information maximum likelihood adjusted models, are presented in this thesis to contribute to understanding how data missingness can effect inferences. In most cases, however, model estimation was impossible under listwise deletion because of sample sizes of zero and under FIML for the same reason. These treatment methods for missing data appears to be problematic when the amount of data that is missing is large.

This leaves the problem of deciding which MICE imputation to report, however, formulae for calculating pooled estimates of parameters for multiple imputed datasets are provided by Rubin (1987 in Dong & Peng, 2013) and are used throughout the findings section. The treatment of missing data could have been improved here. It is not explored what benefits there may be to imputing data when using a wide-format data structure as opposed to a long-format. Logically, imputation using wide-format may give better weight to time-invariant changes in variables, however, it may equally ascribe too much predictive power to these variables. Imputation that takes into account the clustering at person and household level may impute more accurate missing values, but the computing power required to impute these using MICE even as fixed effects on a dataset of 293,120 time-person-household observations with 339 variables is intensive, taking three weeks of continuous computation in this project, although FIML estimation methods are helpful for taking this clustering into account in multilevel models. CART regression may offer benefits over global regression prediction, but was not possible to run successfully on a dataset this complex. These weaknesses to the treatment of missing data should be taken into account when considering the final conclusions.

This research used the *R* MICE package to impute five datasets with missing values estimated using global regression models, where continuous variables' values are imputed using predictive mean matching, ordinal variables are predicted using proportional odds logistic regression, and binary variables are predicted using logistic regression. 20 iterations of each imputed data set were completed and the convergence of the mean and standard deviation of the missing values was inspected visually for each variable (shown in appendix 3). The outcome of these imputations is reported alongside the findings.

3.8 What will this combination of methods add to our understanding of poverty and how it should be measured?

Invariance testing represents an oft neglected but crucial process in developing theoretically sound measures of constructs over time and between groups. It involves a time-consuming comparison of fit-statistics between competing nested models to assess whether increasingly strict restraints on parameters produce significantly worse fitting models (van de Schoot, et al., 2012; Little, 2013). More simply put, the structure and parts of a good measure should remain sufficiently similar over time or for all groups in which it can be applied in order to be comparable. Invariance testing for different groups or times forces the model to be more and more identical in its structure until it is either shown to work well with a fixed structure regardless of group or time or collapses under the constraints. According to the quality profile for *Understanding Society*, "to [the authors'] knowledge no analyses of construct validity and reliability have as yet been undertaken in the *Understanding Society* context" (Lynn & Knies, 2016: 27-28). The findings presented here may therefore prove valuable for other researchers who wish to use reliable and valid measures that are identified in this context as dimensions of poverty.

The extensive invariance, reliability, and validity testing of multiple constructs in this complex model of multidimensional poverty is the primary contribution of this thesis to the literature. If measuring poverty in this way is shown to be time- and group-invariant, an argument can be made that the measurement of poverty using structural equation modelling frameworks is not only feasible but reliable and valid (Carmines & Zeller, 1979). Although reliability and validity are core concepts in sociological research methods training (see: Bryman, 2008, or any number of undergraduate research methods textbooks), such testing is not undertaken or reported as frequently in sociological studies as it is in psychological studies (de Souza, et al., 2017). As the field of psychometrics has been under greater pressure to demonstrate reliable and valid

measures of constructs, in part due to their application in clinical practice, sociology has fallen behind the curve in this respect. An additional reason that analyses of reliability and validity have not been adopted in sociology can be seen in the relative lack of deliberate survey design required to measure sociological concepts as factors, in favour of a focus on singular variables as opposed to sets of items deliberately designed to capture an underlying concept.

This is why, using secondary data, the underlying factors here and in Tomlinson, et al., (2008) are pieced together by various measures that were often not intended to measure a single underlying construct. The composition of constructs require some creative thinking and advanced models that can deal with a mixture of categorical, ordinal, and continuous variables. For example, a financial strain factor may be formed by whether someone has been unable to pay a council tax bill in the last twelve months (categorical), a person's subjective self-assessment of their financial health (ordinal), and household income (continuous). This approach makes reliability and validity analysis even more important in order to establish confidence in the measure but, frustratingly, means that many of the typical methods of reliability analysis are not able to be applied to assess this. For example, the most conventional form of assessing internal reliability, Cronbach's alpha (Cronbach, 1951), cannot be reasonably applied to this kind of structure of items. More flexible criteria for assessing reliability and validity are required. Factor analysis is 'explicitly designed to cope with' sets of items requiring such flexibility (Carmines & Zeller, 1979: 59)

Revisiting these concepts of reliability and validity is a reminder of why they are important from an epistemological and ontological sense – in terms of the systemic processes of knowing something to exist and how it can be shown to exist. Reliability can be broadly split into three types: external reliability (or test-retest reliability), interobserver reliability, and internal reliability. Validity can be assessed similarly through the lens of discriminant and divergent validity. A good construct should be similar to other constructs it is theorised to be similar to, but unique enough to be distinguishable from other things (Cambell & Fiske, 1959; Fornell & Larcker, 1981). Although the statistical assessment of these criteria have found great purchase and are now routinely used in psychology and its applied subdisciplines, they have yet to catch on in sociological research.

The purpose of invariance testing is to establish the exact degree of external reliability that exists in factors (van de Schoot, et al., 2012). Factors with items that are constrained to the

same factor loadings, intercepts, residual error variance, and, when appropriate, residual item autocorrelation (residual correlations for items between equal points in time), that continue to have good model fit, are shown to be stable – items can be interpreted in broadly the same way and relate to the same underlying factor regardless of temporal changes (Millsap, 2011; van de Schoot, et al., 2012).

This is akin to ensuring that the intervals on a ruler or a thermometer and its zero point still refer to the same things at all points: that you're measuring distances in centimeters and not inches, that zero refers to the freezing point of water, and so on. This method helps us test the item-specific reliability however, with regards to item test-retest reliability, much confidence rests with the Institute for Social and Economic Research (ISER) at the University of Essex, who overview the design and implementation of the *Understanding Society* survey. The *US* team first pilot all survey instruments and ensure that individual questions that are repeated through each wave are asked "in an identical form, with identical routing, and in as similar a context as possible" (Lynn & Knies, 2016: 21). In addition to this, interviewers are given extensive training prior to data collection to ensure that reliable data is collected.

The assumption of interobserver reliability, the idea that two or more raters will come to the same interpretations with identical information when assessing a response to a question (Cramer & Howitt, 2004), must also largely rest with the design processes at ISER. Although these can be assessed using statistics such as Cohen's *kappa* (Cohen, 1960) or the intra-class correlation coefficient (ICC) (Shrout & Fleiss, 1979), these are not an adequate basis from which to claim interobserver reliability compared to piloting, especially with mixed-methods data collection. The ICC is a statistic that also captures the extent of hierarchical multilevel effects within a model and is often used to decide whether a multilevel model should be adopted (Raykov, 2011; Robson & Pevalin, 2015); in a model with data as heavily nested as *Understanding Society*, it is not clear how much of a high ICC is attributable to group effects and how much is attributable to high interobserver reliability. Due to the large number of variables that are collected in the survey it is possible that interobserver reliability has not been established for every variable, but trust in ISER's instrument design and piloting is not likely to be misplaced.

Where the previous two definitons focus on whether instruments are replicable over time and between different individuals, the final one, internal reliability, is concerned with whether

several items purporting to measure the same underlying construct do so with a high level of agreement (Rossi, et al., 2013). Any one attempt to measure a construct will elicit subtley different responses or yield marginally different interpretations and capture certain parts of the whole of what is intended to be measured (see figure 12). The use of factor analysis and interrogation of fit statistics establishes whether there is internal reliability within a set of variables either explicitly designed to measure the same construct or that have been placed together under the same theoretical framework. If items are strongly related to one another a factor solution should be able to successfully predict the majority of individual responses to each item based only on how much variance they share with one another. If a factor does this well, the corresponding fit statistics will be high, because of the high internal reliability of the items. Therefore, when talking about a 'good' or 'well-fitting' model in factor analysis and SEM, we are also always talking about factors that have good internal reliability.

This is the natural extension to Bradshaw and Finch's (2003) observation that measures of poverty that capture different dimensions do not frequently overlap. They must either be considered independently, or one should be chosen, however, both approaches have weaknesses. If measures are considered independently there is a real danger of overestimating the prevalence of poverty, as the level of overlap may not be taken into consideration. Even if it is taken into consideration in the methodological framework, the degrees of similarity may not be communicated effectively. In contrast, if only one measure is chosen the true experience of poverty is not being accurately reflected in the metrics used to measure it. The approach suggested here aims to capture multiple, if not all, dimensions of poverty simultaneously by creating factors which simultaneously predict the commonalities between the dimensions that are available. With this, the consequence of the traditional approach of drawing a 'poverty line' becomes clear: 'no single poverty line can capture [deprivation in] all these dimensions' (Alcock, 2006: 80). The problem with measuring poverty therefore lies not within the construction of a single multidimensional metric, that much is possible and is shown here and has been shown before (Tomlinson, et al., 2008), but with the policy-driven requirement for any such metric to have a clear 'line' for the binary distinction between persons who are poor and persons who are not; this restriction is the key limiting factor in measuring poverty. Poverty is multidimensional and poverty is a spectrum, not a binary.

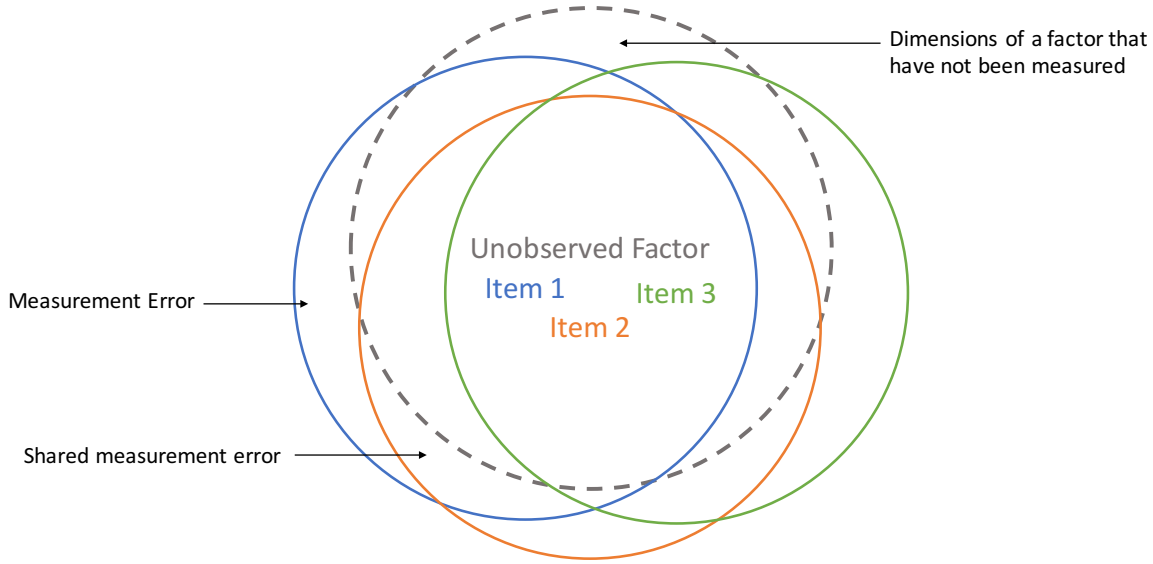


Figure 12: *The theoretical relationship between measurement items and an underlying construct. Items should be broadly in agreement with one another and have a lot of 'overlap' measuring the same things, but may also capture different constitient parts of the construct. All measurement items will have a degree of error associated with them, and sometimes this error will be shared, for example, if the error is for certain people to overestimate their responses in all items. There are also likely to be parts of the underlying factor that are not captured by the items, or are 'uncapturable' using this specific method.*

However, just because constructs and measurements are reliable does not mean that they reflect reality. It is possible to do a very good job of measuring completely the wrong thing. This is partially assessed through the testing of model fit, however, a stronger statistical claim can be made through the use of *construction/exploration* and *validation/confirmation* subsets. Data is separated into an exploratory subset and several confirmatory subsets randomly (Jöreskog, 1993). The proposed factor solution that has been informed by theory is tested and adjusted as required using the exploratory subset – if items do not fit well and have low factor loadings they may be removed. This process of adjusting models *ad hoc* runs the risk of good fit being attributable solely because of measurement error in the sample (Gelman & Loken, 2013).

The need to safeguard against modelling random noise in the data by using exploratory and validation subsets has already been discussed, but in addition to this the use of several imputations of missing data provides additional protection against this pitfall. Although each iteration of the imputations moves them towards convergence, there will still be unique random noise within each set of data. Similar to comparing model fit from MICE data to listwise deletion data and FIML estimation, statistics from different imputations can be compared to one another to test the rigour of a model. If models run on all imputations show similar estimates and model fit, this provides some encouragement that our proposed measure is good.

Furthermore, if estimates and model fit in validation data models for each imputation remain similar to the exploratory fit statistics from the first imputation of data, then an argument can be made that the models have not been constructed based on maximising explanations of random noise in the data, samples, or imputations.

Testing the final factor structure on a new random subset of data sidesteps the problem and shows whether improved fit is likely to be as a result of a better selection of items rather than random noise. All factors below were tested and adjusted first on the exploratory subset of data and later confirmed as genuinely good-fitting using a validation subset. Furthermore, confirmatory factor structures are then tested on each of the five imputations of missing data. If models were just fitting well because of random noise in the specific imputation they were constructed with the same model tested on a different imputation should produce significantly worse model fit. Although the results show some decrease in model fit in imputations two through five (appendix 4), this was negligible in all cases; usually a change of less than 0.001 in the CFI and TLI indices of model fit.

Another way of establishing validity, using the concepts of convergent and discriminant validity, is proposed by Fornell & Larcker (1981). *Discriminant validity* refers to the idea that, for a construct to be reflective of something that exists in the real world, it should be sufficiently different from something similar but theoretically unique (John & Benet-Martinez, 2000). If it is not, then there is no strong evidence that the two factors should be treated as different constructs either because they are actually the same thing, and the theory should be revised, or because our measurements are not adequately differentiating between them, and should be refined. This involves calculating the *average variance extracted* within a factor and comparing this statistic with squared correlations between similar constructs (Fornell & Larcker, 1981). For a construct to show discriminant validity the average variance extracted should exceed the highest squared between-factor correlation.

Convergent validity refers to the idea that a valid construct should be similar to other constructs that it is theoretically close to. This is typically just assessed by interrogating the correlations between constructs. However, with a sample size as large as the sample used here (over 293,000 in each imputation at the time-individual level), statistical significance tests between constructs carry virtually no substantive meaning, as a correlation of any size is likely to have a standard error so small that it will be statistically significant. Instead, the approach used here

has been that using a higher-order factor to simultaneously predict each lower level factor and then inspecting the fit statistics of this higher order factor provides a better indicator of convergent validity that is more robust to large sample-sizes. If factors show convergent validity, then all loadings for the higher-order factor should be "salient" – meaning all standardised factor loadings are greater than 0.3 (Brown, 2006: 130) – and the model fit should not drop significantly from a model where factor correlations are freely estimated.

Although this kind of sophistication may seem excessive, it is highly relevant to the substantive question of how to create theoretically informed sociological measures. It is a way to assess what is a fundamental part of Townsend (1979) and others' definitions of poverty as a dynamic concept over time and for different groups and contexts. If poverty means something different at different times, invariance testing is key to identifying where these differences exist, whether they are valid and reliable, and points of change where our indicators of poverty may need to be revised to fit with its shifting form. If poverty means something different for different groups of people, invariance testing is key to finding out for whom it differs and in what way. If what poverty means changes over time, our metrics should also change over time. However, identifying when poverty has changed sufficiently to warrant a reimagining of our instruments of measurement is, as of the time of writing, instigated more by the face validity of existing measures and levels of expert consensus, rather than by any kind of statistical touchstone.

3.9. Summary and reflections

This chapter has presented a series of methods that constitute a novel approach to the measurement of poverty using high-quality longitudinal secondary data. Several research questions have been identified and form the following findings chapters. *Understanding Society* is a national survey that has been developed and piloted thoroughly to reduce sample attrition and maximise response rates. It has a broad enough remit to be able to capture multiple dimensions of poverty identified in multidimensional theories. The logic underpinning the methods chosen reflects the aims of the research in their identification of underlying latent constructs that are multidimensional and difficult to operationalise. This makes SEM and FA valuable in investigating the research questions posited at the beginning of the chapter. Further, this chapter has identified developments in the methods used to account for important features of the data, such as its multilevel structure and patterns of missing data, even though not all solutions were able to be implemented. The use of MICE and FIML to handle data missingness counteracts potential bias that may be introduced by data 'not missing at random' patterns. The

principles informing the exploratory analysis of data and attempts to safeguard the researcher from following a 'garden of forking paths' have been made explicitly and, insofar as it's possible, I have tried to present an overview of these complex methods and their purposes in a way that non-technical audiences can engage with and critique.

I believe this demonstrates a significant development in the multidimensional measure of poverty, building on the original work of Tomlinson, et al. (2008). There is still work to be done and improvements that can be made. With regards to data, variables that constitute dimensions of poverty are still collected piecemeal and although some relate to well-developed sets of questionnaire items, especially in the field of subjective health and wellbeing, others, such as questions about housing quality or financial strain are not asked with the regularity that they might have if survey methodologists were explicitly designing surveys to capture indicators of multidimensional poverty. This is a downside of reliance on secondary data. Secondly, there are challenges establishing household-level identifiers where a push for a standardised consensually agreed method for tracking quasi-household membership over time would be valuable. Thirdly, there is work to be done in making procedures for imputing missing data in complex survey data more accessible for end-users of statistical packages and secondary data. These include questions about whether missing data should be imputed using wide- or long-format, and how to deal with imputing multilevel data using classification and regression trees when the number of clusters is very large. This is especially important if literature on bias that can be introduced when using listwise deletion is accurate. Finally, technological limitations in regards to computing power remain to be a notable hindrance for employing these more advanced, complex, methods.

The solutions used here are very unlikely to be the best solutions. On reflection, I believe that making these problems explicit and attempting to address them to the best of my ability is more valuable for advancing the field than ignoring them or trying to conceal their existence and impact, especially in a field where they are not often used. This is the reality that most users of applied statistics find themselves in. We must engage with the substantive topics of our inquiry, aim to use the best available methods, and be abreast of recent methodological advancements that we may not be in a position to contribute to or appreciate fully because of their complexity. We are beholden to the methodological innovators and their ability to communicate the application of their developments in a way we can partake in, and in software developers to create environments that we can engage with. Even then, it is unlikely that we will employ

these techniques perfectly, although some no doubt claim that they have. We simply do our best. When engaging with colleagues at the forefront of quantitative methodological advancement there is pressure to be perfect in our deployment of methods, or at least *to be seen to be* perfect. Methodological reflexivity is not something that quantitative researchers engage with often. I think this is a mistake. It encourages omission of the fact that things do go wrong or that we are unable to 'get something working' out of fear that the admission of the limitations of our research and abilities exposes too many weak points that may be exploited by those more methodologically adept than ourselves. We make excuses or obfuscations that can result in overconfidence in our findings. This chapter represents an honest documentation of *my* best effort, but the approach taken here is in no way the best that can be done.

Chapters Four, Five, & Six: Findings

4. Introduction to Findings

The following three chapters contain the substantive findings from the research and are organised as follows: this chapter considers the findings from the exploratory part of the analysis, including the identification of eleven distinct clusters of the selected variables. Each of these clusters is then discussed in reference to the literature to assert what they appear to be representing. The intersections between clusters are also considered. These intersections are interpreted as reflecting relationships between underlying latent factors, and are instructive in the design of the multidimensional construct. Lastly, the chapter considers critically the 'missing dimensions' from survey data, and the politics of large national surveys that might result in the underrepresentation of certain issues, or resultant design structures that limit analysis pertinent to marginalised groups.

Chapter five outlines the factor analysis results for each of the latent constructs identified from the cross-referencing of the literature with the exploratory analysis in chapter four. This includes whether the factors were stable over time and how well they can be captured using existing *Understanding Society* variables. Before presenting the factor formation of a multidimensional poverty construct, the validity of each of the independent dimensions is tested using the Fornell-Larcker criterion. The consequences of discriminant validity for some factors and a lack of discriminant validity for others is discussed before presenting the multidimensional poverty factor analysis. Finally, the theoretical implications of this factor analysis are discussed in reference to the work of Townsend (2010), Lister (2004), Walker (2014), and others.

Chapter six presents the results of multigroup invariance testing of the multidimensional poverty construct across the select few social divisions able to be derived from the *Understanding Society* data: household type, gender, and age/cohort. This section establishes whether the proposed measure can be applied equivalently to different social divisions, and identifies any potential weaknesses. Finally, the concluding chapter seven discusses the positioning of the contributions of the thesis within poverty studies literature, as well as within the broader methodological literature, the discipline of sociology, and the policy context. It considers the limitations of the research and potential ways of moving forwards with regards to the measurement of multidimensional poverty.

4.1 Exploratory Analysis Using Large Scale Correlation Plots with Hierarchical Clustering

Each construct's development was based firstly on Tomlinson, et al.'s, (2010) factor solution. Items that carried over from the *British Household Panel Survey* were identified and retained. Following this, all items from *Understanding Society* were screened for their applicability to theoretical dimensions of poverty informed by the literature review. Finally, a visualisation showing the nonparametric correlations between all variables of interest clustered using hierarchical clustering was used to identify potential factors (see figure 13). This plot was then inspected visually and annotated to easily identify potential items that would form factors. Each of these clusters of items is outlined below, with reference to their theoretical foundation from existing literature.

This type of visualisation has been used in other fields (for example, see: Liu, et al., 2012), and is sometimes applied in the social sciences, however, although it offers great benefits to the identification of potential factors in exploratory data analysis, its usage is not currently common practice. A correlation matrix without hierarchical clustering of variables is difficult to interpret, as patterns are not immediately visible. However, by producing a distance matrix, using hierarchical clustering and then sorting the axes by this matrix, potential groupings of variables become much more obvious: positive correlations clustered along the diagonal of the matrix imply first-order factors and clusters of correlations throughout the plot imply correlations between factors, which may also be considered higher-order factors. The creation of these large feature matrices draw upon the *corrplot* package by Wei, et al., (2017) as well as native functions in *R* (R Core Team, 2016) and in *tidyverse* packages *ggplot* (Wickham, 2009).

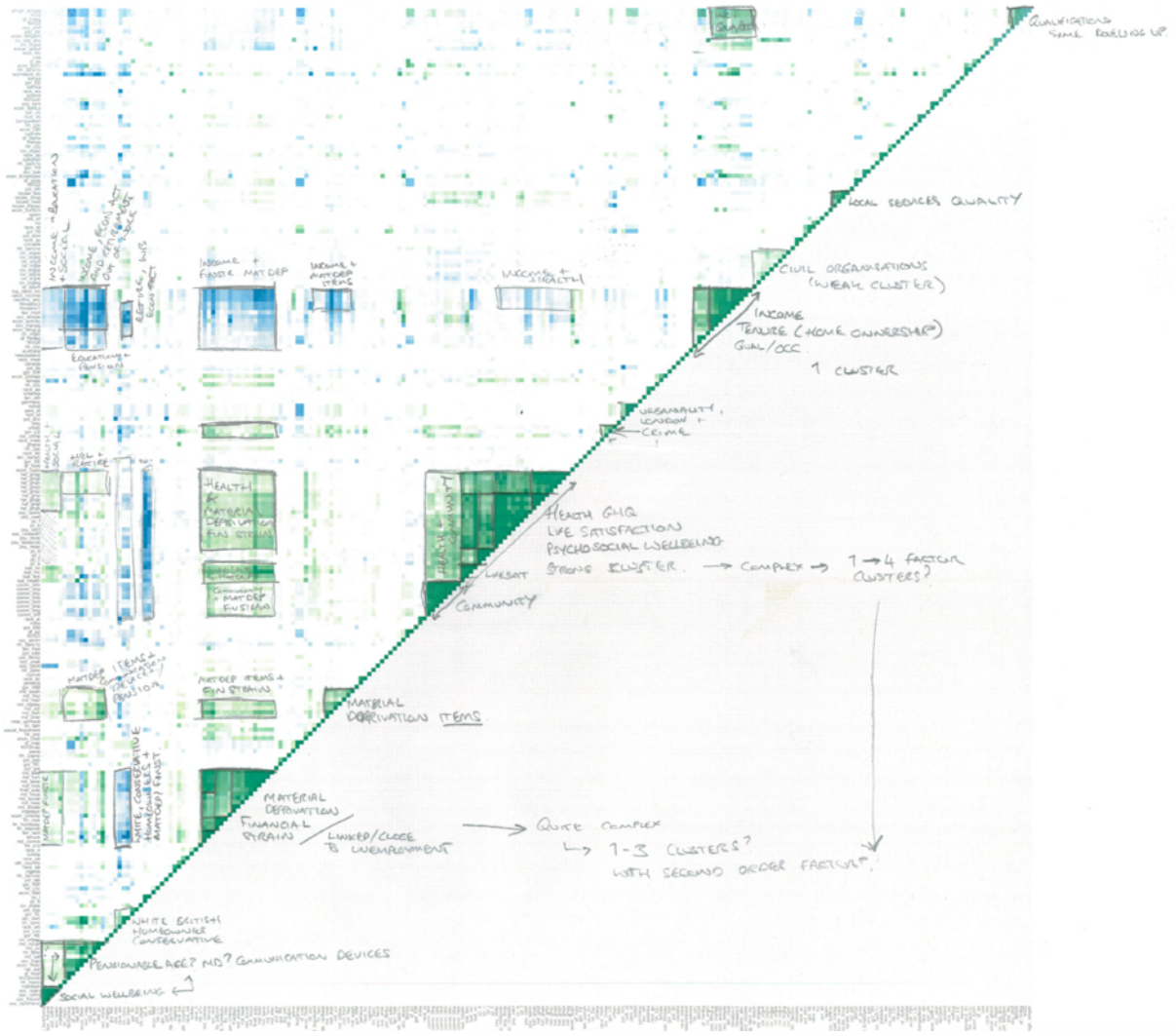


Figure 13: *Hierarchically clustered items in a large correlation matrix visualisation. This visualisation was used to identify potential factor solutions as well as visualise the potential relationships between them. This figure is reproduced at a larger printed resolution in the pdf version of the appendix as entry 2.2.*

There are at least two strong reasons why these kinds of visualisations are pertinent for sociologists. The first reason is a technical one. In the previous chapter the importance of exploratory data analysis using visual representations of data before relying on tests of statistical significance was discussed (Gelman & Loken, 2013). This method allows for potential factor structures to be identified without reliance on statistical tests and using a powerful method of pattern identification, the human eye. Secondly, sociological inquiry often deals with very complex maps of different factors and their interactions, but many forms of data visualisation are limited in how many relationships and patterns can be shown simultaneously. Sociologists may be interested in hundreds, if not thousands, of potential combinations of variables at a time. Having even a general visualisation of how all of these

factors relate helps avoid a similar kind of ‘garden of forking paths’ where a researcher may only follow up specific types of exploratory visualisations. This approach is much more manageable to use as a first step as compared to investigating the relationship between many thousands of possible ‘pairs’ of combinations and their interactions.

The hierarchically clustered correlation matrix plot showed several clear potential factors; the integration of the factors below and their relationships hopefully shows by example how useful the method is for systematically piecing together a coherent picture from a large amount of information. The identifiable clusters from the bottom-left to the top-right of the diagonal on figure 13 are listed below. Of these, some related to dimensions of poverty and others appeared to be linked to certain social divisions and their intersections with poverty, that are considered as they inform the discussion of chapter six, on multigroup invariance:

- Socio-relational wellbeing (social isolation), with regards to friendship and openness
- Pensioner communalities, needs specific to older age
- Retired White British homeowner couples, the least likely to be associated with higher levels of poverty
- Financial strain
- Material deprivation of consumable goods, activities, and services
- Material deprivation of commodities, including household goods
- Civic participation and community cohesion
- Life satisfaction
- General & Psychosocial Health
- Socioeconomic status
- Quality of local services

Each cluster was then assessed in reference to the strength of its intersectional relationship with the other clusters. To be a credible dimensional in a multidimensional poverty construct, factors should show at least some moderate relationship with each other. Because of the high sample size there is a risk that potential factors with only a very weak relationship to other factors in a second-order factor analysis can show up as statistically ‘significant’ while being substantively meaningless, therefore, this first ‘filter’ of visually based exploratory relevance is used as a selection criteria. This process demonstrates the methodological value of the approach that has been used.

4.1.1 Social-relational Wellbeing (Social Isolation): Friendship and openness

The first small cluster on the diagonal of the annotated correlation plot seems to show a cluster of variables related to social-relational wellbeing, specifically friendship. These include the variables *soc_open, soc_relfrd, soc_frdund,* and *soc_lackfriend* which were the recoded identifiers for *scfopenup, scfrely, scfundstnd*, and *scfrendany* in table 4 respectively. These questions related to whether participants felt like they had meaningful friendships, and whether they felt like they had friends they could 'open up' to about things that mattered to them, whether they felt their friends were reliable and understanding. These all seemed to be highly correlated, implying that a factor could be created that reflected broadly the quality of individual respondents' personal social support networks.

This proposed cluster seems to have some visual association with a financial strain factor, a general health factor, and an income and education cluster – or socioeconomic factor. Those with lower scores on the range of social wellbeing questionnaire items also seem to be experiencing more financial strain, material deprivation of ephemeral consumption, such as fuel, social outings, and holidays, and report poorer general health.

The direction of causality here is obviously not possible to gauge, but it may be that those with high levels of financial strain as less able to maintain and foster close friendships due to the prioritisation of their personal fiscal matters. Those unable to keep their homes in certain states of repair, or those unable to participate easily in social outings because of limited finances may also have difficulties maintaining close friendships. States of repair in this instance imply more than just livable conditions, it implies a state that a person does not feel shame in sharing with guests whom their opinion of them matters deeply. It follows that, in line with the arguments of Lister (2004) and Walker (2014), poverty is characterised by shame, it is unsurprising that greater feelings in shame may manifest in more difficulty maintaining social relationships that are crucial to wellbeing. Prolonged feelings of shame here, alongside real financial restrictions, may act as the mechanism by which poverty and deprivation damage the fibres of friendship ties.

The plot also seems to show that those with lower levels of friendship report worse general health; as these general health questions largely capture psychosocial health this may not be surprising, as those with weaker social ties may have fewer avenues to vent frustrations, feel

listened to and supported, and share in personal lives (Smart, 2007). Although this may be confounded by other factors, such as age, there is some evidence that social relationships do matter for the health of individuals over the life course. For example, Yang, et al. (2016) identified that a lack of social connections was associated with increased risks of health problems at critical stages in life – including increased risk of diabetes and inflammation that in many cases matched the impact of a lack of physical exercise.

Lastly, there is a negative correlation between a more vague socioeconomic class factor and social-relational wellbeing. This implies that repondents with lower income, education, and occupational status, have on average higher social-relational wellness compared to those on higher incomes, educational attainment, and occupational status. This aligns with observations of dynamics amongst working class populations where solidarity, community cohesion, and inter-personal support networks are highly valued (Roberts, 2001; Joyce, 1995). There has historically been greater community stability amongst working class populations in the United Kingdom, which has a positive effect on local friendship ties and rates of local social participation (Sampson, 1988). This seems to show some persistence of a community foundation based on close social ties amongst some sociodemographic groups, even in an era where some have argued that there has been a 'death of class' (Eidlin, 2014).

4.1.2 Pensioner Communalities: Qualities specific to those in older age

The next clear factor was largely unexpected and refers to common features that seem to centre around a person being of pensionable age. These included the renamed variables *md_nocar, md_pc, hh_fpnc, md_sat, occ_inapp, econ_Retir, hel_lsill, qf_noqf, NVQnone,* and *hh_mpnc*. In order, these referred to whether participants did not own a car, whether they did not own a personal computer or laptop, whether they lived in a household where the only person was of pensionable age and did not live with any children, whether they could not afford but wanted satellite television, whether their occupation was inapplicable and their economic status was retired, whether they had any longstanding illnesses, and whether they had no qualifications.

This specific clustering is important as it seems to imply that there may be an additional factor that is related to poverty and is specific to the experiences of older people at the time of the survey. Indeed, the specific importance of certain dimensions of social quality for older people, especially those associated with health and illness, has been highlighted in existing research and may be represented here (Foster, et al., 2018). Beyond this, however, it also shows how at

a grouped cohort level there are significant changes in what constitutes desirable material possessions, or things that certain generations are unlikely to have access to, with satellite television, personal computers, and vehicles being more desirable but more unaffordable or impractical. Their absence seems to be most acute amongst older people. We might therefore expect to see multidimensional poverty factors fit less well for older people than they do for other cohorts.

This older-age specific cluster of items also showed some association with other commodities that were not available, but not to as great a level as the specific ones already highlighted. Washing machines, freezers, music players, and microwaves, were also examples of items that were more likely to not be owned by people with this cluster of characteristics. Again, it is likely that the relationship here is not straightforward. Although older people have historically been more likely to be in poverty as measured by the HBAI measure (Joyce, 2014), and may be more likely to desire some items, especially due to illness of disability (Bartlett, et al., 2013), the items included in the material deprivation item set may not be pertinent to certain generations. Culturally, they may be less inclined to want media players or personal computers, in the first place, but may be severely materially deprived if they do not have access to certain things that may be standard for older people with higher incomes, especially at the point where age intersects with disability. These might be items such as stairlifts, accessible vehicles, or health monitors. To further complicate the picture, there has been concerns that questions related to material deprivation work well for families with children but are more often misinterpreted or illicit unusual responses from people in older age (McKay, 2008).

As would be expected, this factor is also strongly associated with the more vague socioeconomic status factor, which seems to indicate a specific cohort identity wherein people who fall within the higher range of scores in this factor are also more likely to have lower incomes, lower occupational status, and lower levels of qualifications. With a very different labour market, arguably growing democratisation of education, the decline of the UK manufacturing industry, the rise of highly-educated workforces, global competition, and a new class of precarious workers, it is unsuprising to see such an interaction between the two factors, especially when a large cohort with similar life experiences are grouped together within the vague categories of ‘retired’ (Crouch, 2016; Savage, et al., 2013, 2015).

4.1.3 The "Retired White British Homeowners"

Another unexpected but interesting cluster of a very small number of items was that of households that were of pensionable age, living with no children, who identified as 'White British' and owned their own home. This appeared to be a factor that had certain household contexts that maximised their income. Their housing tenure status implied that they had been relatively prosperous in their working age, their lack of dependent children implies greater disposable income, and their ethnicity implies an important lack of ethnicity-related barriers during their working lives (Payne, 2013). Although this is an overstatement, it is not hard to imagine that this factor represents a generation and specific intersection that many would refer to as 'baby boomers', many of whom benefitted from the labour market stability and welfare state conception after the Second World War, and subsequent fiscal policies that consolidated accumulated wealth introduced by later governments. For those who meet the triad of criteria, on average, they are in a position of lower financial strain and material deprivation, with somewhat better health. There are also artefacts that make the interpretation for this group difficult, for example, they are less likely to be in the 'mortgage' tenure group and are likely to have lower net incomes, however, this may be because they are likely to, at this point in life, own their homes outright and live in smaller households. This is shown by the grouping of ownership category of tenure being clustered within the items, and the fact that there is no correlation with the income variable that is equivalised for household size.

This raises a very interesting contrast with the factor discussed previously, about general widowed or single pensioner communalities. It may hint at a polarisation of people beyond pensionable age, and perhaps reflects the bluntness of grouping all those in older age within one nebulous category, where certain poverty measures may work better than others. The implication may be that the pattern between widowhood, separation, or singledom in pensionable age and poorer quality of life shows that, far from being a continuation of the wealth and prestige that a person had achieved by the end of their working life, there are substantial changes that occur at the stage of later ageing. Even more concerning is that there may be two or more distinct experiences once someone reaches retirement, contingent on their sociodemographic histories. Although not the subject of this thesis, it supports the call for a radical strategy for social policy on ageing that focuses on the whole life course (Walker, 2017).

4.1.4 Financial Strain: Paying the bills

The next major cluster of variables reflects financial strain. These variables are also in incredibly close proximity to material deprivation items, outlined below, which may indicate that the two may be effectively measuring the same underlying factor. This may mean that the two concepts, financial strain and material deprivation, are not sufficiently distinguished by the measurement items (something that a Fornell-Larcker test investigates in the following chapter). However, from only the visualisation there does seem to be an argument that the two can be treated as unique factors. The financial strain factor includes a cluster of the three variables *fin_ctax, fin_homepay* and *fin_billprob*. These items were also strongly negatively correlated with income, as would be expected from the findings of Tomlinson, et al. (2008), however, someone could be experiencing financial strain whilst not necessarily being on a low income. Nonetheless, a lower income appears to be commonly associated with financial strain. The tenure status of renting privately, *ten_rent*, also appears to cluster with these items, which may reflect the strain that people living in rented accomodation experience as the most common tenure position of those living in poverty (Crisp, et al., 2016).

As would be expected, there is some correlation between the financial strain items and the more permanent fixtures that make up a different form of material deprivation; deprivation of household goods that are much more changeable to societal norms but also last far longer than more transitory activities and services that require regular reacquisition. People living in financial strain also seem more likely to be in a position of being unable to afford items such as washing machines, freezers, tumble driers, and media players. However, these kinds of material goods are clearly clustered together separately, and therefore it may be that not all measures of material deprivation capture the same things. This goes to support the assertion made later that questions about material deprivation should clearly be split into 'consumable' material deprivation items, and more static household 'commodities' that are also very contextual.

Items associated with financial strain also seem to be closely related to items regarding life satisfaction, general and psychosocial health. In some ways, these can be pictured as some of the spokes coming from the central dimension of unacceptable financial and material hardship in Ruth Lister's symbolic-relational wheel of poverty (Lister, 2004). Poorer life satisfaction may also be indicative of internalised feelings of shame as a result of financial position, with the inability of being able to keep up with bills and achieve stability impacting peoples'

reflections of their own self-worth and self-efficacy, which supports a theoretical framework that takes the standpoint of poverty being the prevention of socio-relational needs that occur closely together with financial difficulties (Walker, 2014; Doyal & Gough, 1991).

Financial strain, as a factor, reflects a far more holistic evaluation of a lack of fiscal resources compared to using only income and assuming that low equivalised income will be synonymous with financial strain. This complexity within a metric is important and will become a greater necessity as childcare and housing costs in particular become more polarised. Household finances may come under greater and greater strain due rising costs of living, as had been shown by the JRF (2018), not necessarily because of changes in the distribution of income. This proposed factor would provide a metric that is sensitive to this and better identifies insufficient income levels.

4.1.5 Material Deprivation of Consumables: Activities, habits, services, & items that make up participation

Items that all cluster incredibly closely with one another relate to actitivies, services, and consumable goods that require regular replenishment or continuous funding. These included the recoded and relabelled variables *md_elec, md_furn, md_save, md_hols, finst_subj, md_hous, md_social, md_heat,* and *md_shoes.* Respectively, these refer to responses about whether people can afford, if they so desired: to replace major electrical goods if they were to break; whether they can afford to replace furniture if it breaks; whether they are able to put aside savings regularly; whether they can afford to go on a week-long holiday not staying with family or friends; their self-assessment of their financial situation (subjective financial situation); to keep their house in a decent state or repair; to feed friends or family once per month if they visit; to keep their home warm; and, to have two pairs of all-weather shoes for each member of the household. Many of these are activities, goods, and services that bare a strong connection to the idea of poverty as being the inability to participate in society to a level expected 'by virtue of their membership' (Townsend, 1979: 31). The fact that subjective assessments of household finances cluster so highly within these material deprivation items also seems to imply that the ability to achieve these things is most strongly linked to finances, supporting the basic assertion of Townsend that poverty is about participation in what is expected of the person.

The underlying cluster of these items also seem to be associated strongly with questions about life satisfaction and with health and psychosocial wellbeing items. This, again, supports the theoretical position and findings in qualitative research that relational-symbolic and psychosocial dimensions are critical to a holistic definition of poverty (Lister, 2004; Walker, 2014). Life satisfaction may act as an approximate reflection of shame when framed as an assessment of ones life (Walker, 2014), their achievements and level of self-actualisation, which is inherently relative to others. Psychosocial health indicators may in turn reflect the manifestations of various assaults on the dignity of people living in poverty, including how they are characterised in the media or in political discourse (Lister, 2004; Shildrick & MacDonald, 2013). These are particularly salient if one follows the line of the matrix up from the subjective financial assessment item – there are very clear, very high correlations between higher scores on a person's self-assessed financial situation (higher scores equal a worse subjective situation) and every item in the life satisfaction, psychological, and general health questionnaire items. In other words, if you feel like your financial situaion is bad you are also more likely to feel unhappy about your life and be exhibiting symptoms of poor general psychosocial wellbeing. Of course, high responses on indicators of poverty do not necessarily predispose someone to poor health and life satisfaction, but they are so strongly associated that it would be disingenuous to treat them as not intrinsic to the experience of poverty, as is recounted in many narratives of people with lived experience (ATD Fourth World, N.D.).

Of particular interest is that these items do not cluster together with other items that are traditionally grouped under the label of indicators of material deprivation, namely household goods such as media players, fridge freezers, or dishwashers. Therefore, any scales that seek to combine these items uncritically as measures of material deprivation, such as by simply adding up how many of these items a household is lacking, may not be capturing different kinds of material deprivation. By combining them uncritically, they may be creating a measure of something that rarely actually happens. Instead, this visualisation suggests that material deprivation should be separated into two distinct factors.

4.1.6 Material Deprivation (Commodities): Household Goods

Items pertaining to whether a household possesses certain goods and appliances make up another loosely collected cluster. These include the recoded variables *md_wash, md_frez, md_micr, md_cdplay, md_vcr, md_dish,* and *md_drier*, which refer to whether the household possessed a washing machine, a freezer, a microwave, a CD player, a video recorder or DVD

player, a dishwasher, or a tumble drier, respectively. Many of these items are clearly very strongly linked to a specific set of culturally and temporally sensitive desirable items, some items already reflect redundant appliances. For example, many entertainment systems have moved away from digital disc based media in favour of online streaming, or fulfil multiple functions as music players and digital content platforms for movies and/or video games.

Items that are highly specific to the context of a given society at a given point in time, items that are subject to great degrees of change in both their achievability and their desirability, are perhaps the ones that we would most expect to be unstable as a poverty dimension. The idea that these are unstable might lead to a conclusion that they are therefore not useful, or should not be included. This is not necessarily true. That line of thinking reflects a view of metrics as stable and unchanging and does not fit within approaches that attempt to measure dynamic concepts, it is a view that is fit within a very rigid worldview that presupposes that there are only underlying 'things' that are not subject to change or re-evaluation over time that are most often found in psychometric approaches. At worst, it disregards important contextual items to the experience of poverty just because they are difficult to integrate into a methodological framework. The ability to attain ownership of material possessions that are desirable in society, and provide many benefits to the performance of certain roles, **is** an important dimension of poverty.

The possession of items that facilitate participation in entertainment and culture; maintenance of the home; and the capacity to provide for those close to us - that people are widely expected to own - is undoubtedly important and outweighs the inconvenience of having to reassess occassionally what these items should be in line with specific cultural and societal changes. Their awkwardness should not prohibit them being used within metrics, however, their interpretation may differ. Increases or decreases in factor scores in this dimension are not so straightforwardly interpreted as increases or decreases in poverty, as they may also reflect changes in societal norms. The reassessment, reliability, and validity checking of these items is therefore crucial to identifying when they might need to be updated, and for identifying how applicable to the contemporary setting they may be. The questions themselves, however, might be improved by being worded in the way that the material deprivation questions related to consumable good, activities, and services have been worded by allowing respondents to clarify whether they lack the possession of these goods because they are unable to afford them or whether they lack them because they do not desire them; for example, if they do not enjoy the

activities associated with them, such as in the case of media players, or do not feel like they would benefit them, such as in the case of goods like microwaves or freezers.

Material deprivation of these kinds of goods seems to be associated with levels of consumable material deprivation, the collection of items described above, and with financial strain. More financial strain and material deprivation in the participation domain is associated with higher probabilities of not possessing these appliances, suggesting an important link to the experience of poverty more generally. Lower income and socioeconomic status in terms of education is also associated with a lower likelihood of possessing these items. However, a lack of these goods is not necessarily associated with lower reported health and psychosocial wellbeing. We might expect this because of the lack of distinction between forgoing such items for reasons of personal wants or preference and forgoing them because of insufficient income. The benefit of condensing these items into a higher order factor with other dimensions of poverty is that, because factor analysis only models commonalities between manifest variables, the higher-order factor should only contain contributions from this material deprivation dimension that are also associated with the other factors, such as financial strain and life satisfaction, with the residual variance being contained as an error term that does not contribute to scores in the final multidimensional poverty factor.

4.1.7 Civic Participation & Community Cohesion

Two specific clusters with perhaps unexpected outcomes that may challenge some narratives of poverty were identified in civic participation and community cohesion. Although items associated with civic participation appear to cluster together into their own underlying factor, including membership of organisations such as political parties, trade unions, environmental groups, school associations, tenants associations, religious and community organisations, volunteer groups, social and sports clubs, and/or gender-focused organisations (with examples of a Women's Institute group or feminist group from the *Understanding Society* coding), an underlying factor associated with membership of these groups does not seem to show any strong association with any of the other underlying factors related to poverty, or indeed to any other factors that are more broad and include health.

Similarly, participants were asked to respond to items that indicate a level of community cohesion and belonging, which all appear to cluster strongly together. These items included questions asking the extent to which they feel like they belong to their neighbourhood, that

their local friends mean a lot, that advice is attainable locally, that they can borrow things from their neighbours, that they are willing to improve their neighbourhood, that they are similar to others in their neighbourhood, and that they talk regularly to others living in the neighbourhood. The high correlations between each item in the cluster indicate a high level of reliability for a 'community cohesion/involvement' factor, however, the exploratory visualisation indicates that there only seems to be a strong argument for an association between this factor and the underlying life satisfaction and health, though not for the material deprivation or financial strain factors.

This is surprising given some existing literature. One interpretation of a study by Gaille, et al., (2003) in Eckhard (2018) seems to imply that existing research already shows a link between social isolation and poverty, with those socially isolated being more likely to be living in poverty. However, there is no evidence in the research to support such a strong claim. Eckhard (2018) goes on to show a statistically significant correlation between poverty and social isolation in German panel data after controlling for health, life events, and time, referring to this as a 'strong correlation' (Eckhard, 2018: 348). Although these are highly statistically significant, the argument that they are strong is somewhat misleading – using linear probability the increase in the percentage change in the likelihood of social isolation, as measured by living alone, having no one to turn to for long-term help, or having no one to talk to about thoughts and feelings, when a person is living in poverty (as measured under the OECD definition) is only between 1.3 and 1.8 per cent (*ibid:* 350). Although this is clearly an important increase in risk, especially at a macro population level, it hardly constitutes an fundamental dimension of multidimensional poverty or a strong correlation.

This is where poverty research and social policy commentary needs to be absolutely certain of the definitions and realities implied by certain terms and constructions of social problems. Social isolation, social cohesion, community cohesion, and social exclusion are often spoken of almost interchangeably and are frequently positioned as inevitable experiences of those in poverty, especially in the media (Stubbs, 2008). However, these are hugely different concepts with no clear definitions with regards to what level of cohesion and exclusion they are referring to. As discussed above when interrogating the 'social wellbeing' factor, when measured as the closeness of friendships the interpretation and outcomes will differ significantly from if it were measured as opportunities to meet new people, or diversity of friendship networks along gender, class, and ethnic divides. Similarly, community cohesion will differ significantly and

paradoxically depending on whether we are interested in social networks, neighbourhood involvement, engagement and a sense of ownership in town, city, or national events, or even non-spatial communities that exist online. An individual could, very reasonably, feel very disconnected from their immediate neighbourhood but deeply embedded in an online community. Another could very happily live alone with few local friends but have a thriving level of involvement with events in their city, or a dynamic life as a political activist at national events. In another sense, a person could feel socially excluded from society at large through damning narratives and discourse, but maintain networks of strong mutual interdependence with their local community.

In existing quantitative research and policy debates these distinctions are not made clearly enough: a vaguely defined concept is linked to any feasible measure uncritically because of inadequate operationalisation, as can often be the case in poverty measures (Lister, 2004). As a result, the social problem of isolation and community cohesion is often coupled to the social problem of poverty, despite contradictory evidence and there being no theoretical justification for this. As previously discussed (Roberts, 2001; Joyce, 1995; Jones, 2012; Sampson, 1988), people living in poverty actually have very strong connections in their communities in general, especially working class communities. It is more likely that stigmatising discourse around people in poverty has lead to a perspective of isolated, unlikeable, unpersonable individuals that has pervaded popular understanding (Shildrick & MacDonald, 2013; Shildrick, 2018). Anyone seeking to quantify poverty should be critical of any such uncritical conflation.

Because civic participation was included as a lower-order factor in Tomlinson, et al., (2008) it was carried forward to the exploratory and confirmatory factor analysis stage here. However, because the community cohesion cluster of items seemed to show so little visually obvious relationship with financial and material clusters of items, and does not have a strong qualitative and theoretical justification, it was not taken forward, as per the selection criteria outlined at the start of this chapter. This exploration does, however, contribute to understandings of what kinds of experiences do and do not seem to constitute dimensions of poverty – community cohesion does not seem to be a defining factor.

4.1.8 Life Satisfaction: Happiness and self-actualisation

Freedom from poverty implies not only an absence of destitution, achievement of material goods, liberation from ill health, and feeling heard, but also includes being afforded the

opportunity to achieve a level of self-actualisation and fulfilment of needs that we are told we should strive towards. This is at the heart of what makes poverty a barrier to achieving universal human needs such as those outlined in psychological literature (Maslow, 1943). Life satisfaction is the antithesis of shame and feelings of humiliation that form key relational-symbolic dimensions of poverty (Lister, 2004). Life satisfaction indicates a person's relational view of themselves within society, feelings of pride, success, or self-actualisation are comparative to how we perceive what these mean in wider society. The greater inequalities in society and the greater perceived feelings of inadequacy a person has, through stigmatising discourse or otherwise, the more likely a person is to feel dissatisfied with their achievements and position in life (Vecchio & Caprara, 2007; Ferrer-i-Carbonell & Ramos, 2012). This relationship is not trivial either, Graafland & Lous (2018: 14) found that income inequality is strongly associated (R = 0.61) with life satisfaction inequality. This factor is therefore crucial for doing justice to and gaining insight into the relational dimension of poverty.

Four life satisfaction items cluster strongly together and between psychosocial/general health and wellbeing items, and community cohesion items. These include *lsat_ov, lsat_inc, lsat_leis,* and *lsat_health*. In turn, they represent respondents' reported overall life satisfaction, their satisfaction with their income, their satisfaction with their leisure time and work/life balance, and their satisfaction with their health. The life satisfaction cluster of items appear to be most strongly associated with general health, psychological wellbeing, and community cohesion. This is unsurprising considering these factors make up many of the spokes of Lister's (2004) wheel of poverty, mutually reinforcing one another.

There is a well-established link between life satisfaction, health, and community cohesion. Greater life satisfaction is associated with more health-positive behaviours in young adulthood, lower prevalence of longstanding illness and poor general health throughout the life course, more frequent social interactions and support, fewer days of mental and physical distress each month, and lower incidence of health-risk behaviours such as heavy drinking, smoking, obesity, and physical inactivity (Grant, et al., 2009; Strine, et al., 2008). The causal links here are most likely to be complex and multidirectional. The experience of psychological distress, longstanding illness, and poor general health will likely result in a decrease in a person's life satisfaction. So too will a sense of disconnection or indifference towards a person's local community. However, feelings of inadequacy due to the perceived position of ones' worth relative to others in society may also have psychological implications and result in self-

imposed withdrawal from social life. Perhaps the best conclusion to consolidating this complex causality is to understand there is a broader underlying experience within which dissatisfaction in life, psychosocial poor health, and perceived feelings of isolation from neighbours and a lack of a sense of belonging coalesce into a compound malaise. Higher-order factor analysis can assist in focusing only on this compound malaise and its intersection with material deprivation and financial strain.

Indeed, life satisfaction items also share a high correlation with items from the consumable material deprivation cluster and the financial strain cluster. Respondents facing more financial strain, with fewer opportunities to participate in activities and attain societally expected goods and services, report lower levels of life-satisfaction. This relationship has been reported before when life satisfaction has been studied as independent from deprivation (Hick, 2016). In some ways, life satisfaction may be the closest existing indicator in secondary surveys that can be used to approximate feelings of shame (Walker, 2014).

4.1.9 General & Psychosocial Health: Poverty gets under the skin

Understanding Society contains two sets of questionnaire items designed to measure general wellbeing and psychosocial health – the twelve-item version of the General Health Questionnaire (GHQ), and the Warwick Edinburgh Mental Wellbeing Scale (WEMWBS). GHQ items 10 and 11 relating to self-confidence and self-worth were absent in the *Understanding Society* datasets used, meaning that the self-confidence factor used in Tomlinson, et al.'s, (2008) construction of a multidimensional poverty measure based on Shevlin & Adamson's (2005) factor structure could not be replicated. Instead, additional items from WEMWBS were included, despite them having more missing data, to capture additional dimensions of psychosocial health. Although psychosocial health is a critical dimension of poverty, especially within Lister's (2004) theoretical framework, it is important when considering the approach undertaken here to not weight the final model too heavily towards the psychological dimension by including an excess of factors. Therefore, condensing these factors into a single higher-order factor measuring general psychosocial distress, as was the method used in Tomlinson, et al. (2008), was necessary, and seems theoretically justified based on the clustering of items. However, this does create addition strain in the time required to fit the more complex model with three levels of factors.

Although this approach of condensing all items into a single higher order factor appears to ignore the complexities of psychosocial dimensions, there is significant evidence presented both here and elsewhere in a meta-analysis of various factor solutions for the GHQ12 that many of the proposed multi-factor structures are not sensible solutions, with a single psychological distress factor solution being more desirable (Gnambs & Staufenbiel, 2018). This is because the items share a very large amount of variance so any individual splitting of the items will be largely capturing the same underlying trait as any other split (*ibid*). This is problematic to any theory, as it implies that what are identified as unique dimensions may not actually be being sufficiently uniquely identified by our measurement instrument; or that the differentiation between different types of general psychological wellbeing may not be a valid taxonomy. If subsequent multiple factors that are argued to be distinct are then analysed independently of their relationship to one another the resulting analysis will suffer from a gross overemphasis of the size of any effects. This can happen even if variables are not so closely related that they cause statistical problems with multicollinearity.

This point can be illustrated by using a very simplified example. Imagine a researcher wanted to observe the impact of education on earnings. They measured earnings in three ways: hourly wage from highest paying employment, monthly gross pay, and annual income after taxes. These three outcomes are clearly inseparably linked to one another, even though they may differ slightly due to people having multiple jobs, or be reported in gross/net amounts. If the researcher produced three models predicting these by something like education they would probably find a positive relationship between education and each of the three measures of earnings. However, it would clearly be logically incorrect to suggest that education independently increased all three and compounded the benefits. Although the error is obvious in this very trivial example it can easily be overlooked when the nature of the relationship between multiple outcomes is not known. For example, if a mental health intervention was reported to reduce anxiety and depression, social dysfunction, and increase self-esteem, and the reader did not know the correlation between these scales, we would have no direct reason to believe that these were not largely distinct outcomes. If these scales all share over 90 per cent of their variance with one another the intervention is much less impressive than if they only shared 10 per cent of their variance.

The exploratory correlation plot exercise appears to support this idea of one large underlying factor as opposed to more than one, with all of the GHQ and WEMWBS items clustering

together with very strong correlations. It is possible that three potential unique factors could be interpreted from the visualisation, containing GHQ-7, GHQ-1, GHQ-8, GHQ-4, and GHQ-3 in the first; GHQ-9, GHQ-5, GHQ-6, and GHQ-2 in the second; and with all of the WEMWBS items, along with some questions from the youth panel that may be WEMWBS based, clustering together in a third potential factor. Again, there is no strong evidence that these clusters relate to three distinct underlying factors as opposed to one single one, and this is something that is best tested through assessing their discriminant validity (discussed later in the following chapter). However, the clusters of GHQ items here do repeat the same pattern of clustering shown in Shevlin & Adamson (2005) and replicated in Tomlinson, et al., (2008), with GHQ items 1, 3, 4, and 7 relating most closely to the concept of social dysfunction and items 2, 5, 6, and 9 relating most closely to the concept of anxiety and depression (Shevlin & Adamson, 2005). Further discussion of the validity of these metrics and factor solutions is not presented here, as there is a vast literature concerning the factorised structure of GHQ items and their reliability and validity (Gnambs & Staufenbiel, 2018). For the purpose of a multidimensional measure of poverty these psychosocial dimensions still need to be condensed into one factor to avoid weighting the final factor too heavily towards psychosocial health.

As previously discussed in the subsections regarding other factors, the health cluster of items share some correlation with the social wellbeing cluster of items, a substantial amount of correlation with material deprivation of household commodities items as well as with the financial strain cluster of items. Those showing higher levels of psychological distress also more frequently are experiencing financial strain, material deprivation, and slightly worse social wellbeing, supporting the idea that health should form part of a 'broader social scientific framework' (Lister, 2004: 36). Whereas poverty and health, especially mental health, are largely treated as cause and effect in literature with competing arguments about causality, these debates miss a third option in that they reflect, to a certain extent and with a sufficiently broad theory of poverty, a closely linked underlying experience that corresponds with internal unacceptable hardship. By this, I mean to suggest that people from disadvantaged socioeconomic positions become susceptible to greater health risks (Marmot & Wilkinson, 2006) and, given the disabling aspects of societal structures that exclude people with health differences or impairments (Goethals, et al., 2015; Beresford, 1996), those people face a double disadvantage of finding their socioeconomic position undermined further by the state of their health. The two may be mutually reinforcing, not strictly causal in one direction.

The decision of which direction causality takes is often reinforced by the position one takes in reference to policy reaction: should we change the unequal structures of society or change the individual? Often both are advocated for, but the epistemic authority of sociologists in public and social policy contexts is likely to be subordinate to that afforded to psychologists and public health scholars by virtue of their ties to medicine and positivist methodology, and for policy in the 21st Century to largely favour individualised approaches (Carter, 1998). Epistemic authority refers to actors conveying information that is seen as valid (Kruglanski, 1980). In other words, certain actors claims to knowledge are often considered more legitimate. Whether someone is judged to exhibit epistemologic authority depends on the culturally specific context of power relations between communicator and recipient of information and the trust inherent in the communicator's position (Hornikx, 2011), the extent to which the recipient already agrees with the information expressed (Bar-Tal, et al., 1991), and the degree of similarity in knowledge and discipline between the two actors (*ibid*).

What is also interesting in these cross-cluster relationships with health is the much weaker cluster association with items concerning socioeconomic status, education level, income, and occupational class. This is important considering how frequently income, occupational class, and education are used as proxies for poverty in studies that seek to better understand the link between income and health (Marmot & Wilkinson, 2006). Reviews of existing evidence highlight repeatedly the association between living or growing up in a low income household and experiencing worse mental health (Griggs & Walker, 2008). Pathways to poor health through material mechanisms and complex poverty pathways may be more elucidating than a narrow one-dimensional focus on income or socioeconomic markers (Benzeval, et al., 2014) Any neglect of the material mechanisms, especially those related to participation in societally normal activities, may severely understate the relationship between poverty and health, with consequences for policy.

4.1.10 (High) Socioeconomic Status

Though not a dimension of poverty, but a significant predictor, the exploratory visualisation indicates a cluster of items that seem to link somewhat untidily to an underlying socioeconomic class factor. The reason this cluster is somewhat untidy has to do with the coding of the items, where higher values indicate more privileged socioeconomic status; the variables were recoded to operate as observed variable controls, not as an independent latent variable, therefore there may be some differently coded items that did not 'cluster together' with the ones here. Within

this cluster of items are variables for income, home ownership, occupational class, highest educational qualification, and, interestingly, trade union membership.

The intersections between these variables and the dimensions of poverty clusters have been discussed in the overview of the previous clusters, but are recapped here. Income variables and, to a lesser extent, occupation and education variables are associated with financial strain and material deprivation. Lower income is associated with greater difficulty keeping up-to-date with bills, greater material deprivation of both consumable activities, services, and goods, which reflect participation, and lower rates of possession of household goods. Lower income and 'lower' occupational status and education also seems to show some small effect on general psychosocial health and life satisfaction, however, the association is very mixed depending on the specific items in these domains. For example, unsurprisingly, low income is associated with far worse satisfaction with income, somewhat worse life satisfaction 'overall', slightly worse satisfaction with health, but has no association with satisfaction with leisure time.

The grouping of items best reflects the conceptualisation of poverty within Spicker's (2007) clustering of concepts surrounding poverty, in particular the role of economic circumstances including economic class, resources, and social class, and the role that intergenerational transmission of inequalities and inequity play (Bishop & Rodríguez, 2018). Income, education, property ownership/wealth, and occupational class and patterns of economic activity reproduce one another and the resultant or preceding poverty within generations in society (Harper, et al., 2003; Carter & Barrett, 2006; Tomlinson & Walker, 2009; 2010; Field, 2010; Bradshaw, 2016; Zelinsky, et al., 2016), although by no means does this assert that children are fated to live out the same types of lives as their parents (Hills, 2015). However, despite social mobility being possible, it cannot be ignored that recurrent and perpetual poverty still significantly bias the trajectory of generations (Tomlinson & Walker, 2009; 2010), and therefore such patterns of interaction between the socioeconomic cluster and the multidimensional clusters of poverty is clearly visible.

4.1.11 Quality of Local Services: Environmental deprivation?

This research also aimed to investigate whether any evidence could be found in individual-level longitudinal data with relation to deprivation at the local-political and environmental level (Hastings, 2009). Four items queried how participants felt about the quality of their local services, and the reponses to these items cluster together in the exploratory correlation plot.

The items included in *Understanding Society* have been recoded as *locale_leis, locale_shop, locale_med,* and *locale_transp*. These refer to questions that ask the respondents to give an assessment of the quality of local services from a four-option scale ranging from poor to excellent. As is implied by their recoded names, the four questions ask for opinions on the quality of local leisure, shopping, medical, and transport services respectively.

The difficulty is not every household or individual will have had involvement with all possible services, making it difficult to quantify local area environmental deprivation using this metric. Most people will have some interaction with local health, leisure, transport, and shopping services on a fairly frequent basis, however, assessments of local schools will be exclusive to people who have remained in the same local area since they were schoolchildren, or people who have children of school age. It is possible that, using multilevel modelling techniques and aggregate ratings, they could be assessed at a higher Lower Super Output Area level or Middle Super Output Area level, but with the absence of neighbourhood, LSOA, or MSOA geographical identifiers that are exclusive to specially licensed versions of the dataset this is not possible.

The four services chosen were used because of their nearly universal applicability to all households. Regardless of individual characteristics or household type, most people and households will have had some exposure to these four types of services. But, within this lies a problem. These types of services are unlikely to be the types of services that are unequitably provided based on deprivation given findings from existing research. Targeted public services especially, as opposed to universal services, are more likely to be unequitable distributed due to the impetus to provide these to the most deserving and needing first and foremost (Mkandawire, 2005). However, universal services can be subject to rationing that leads to an unequitable distribution of service provision, especially in times of fiscal austerity (Klein, et al., 1996; Klein & Mayblin, 2012; Webb & Bywaters, 2018). For example, Hastings (2009) found that through mechanisms of rationing, where poorer communities were sometimes seen as less deserving of services than more affluent communities and therefore had their services 'downgraded' by those responsible for their delivery. Significant inequalities were found between the quality of local services provided to poorer communities and those provided to more affluent communities. These inequities also translate to the macro level of funding of services. In the case of children's services, Webb & Bywaters (2018) found that children in poorer local authorities were subject to far greater cuts in their services between 2010 and 2015

than children in less poor local authorities; Bailey, et al., (2015) also found this to be the case in cuts to services in general.

Local services can often be treated as static in conversations about levels of poverty. There is sometimes an assumption that services provided will broadly be of similar quality regardless of time, despite the kinds of pressures placed upon them. Accounting for this complexity means accepting that even if households' incomes and costs could remain stable, their quality of life, capabilities, and ability to participate in society, could be diminished in other ways by a reduction in the quality of local services. Poor quality schools result in worse educational outcomes and fewer opportunities; poor transport systems remove the capacity to maintain social connections and access opportunities; poor environmental services stifle recreation and create surroundings that induce feelings of shame. This makes them an important dimension that should be reflected in measurements of poverty, even if this may not be possible using existing survey data.

Unfortunately, the lack of additional items for the reasons given above may not adequately capture this relationship at a higher level. There did not appear to be obvious interactions between the local services cluster of items and any of the other clusters of items that relate to poverty. Of course, this may reflect that poverty, or at least poverty at the individual/household level, is not associated with the quality of local services, however, given the evidence that poorer neighbourhoods are more likely to have poorer services (Hastings, 2009) and the items chosen do not capture the same kinds of services highlighted in the literature as inequitably distributed, this lack of interaction is more likely to be a problem of measurement, not of theory. Future research might address this problem by modelling these variables as a dimension strictly at a higher conceptual level using neighbourhood identifiers. Such a methodological innovation will be vital for a truly multidimensional measure of poverty, but is beyond the scope of this thesis.

4.1.12 What's missing? Factors with no measurements

Of course, this exploratory analysis can tell us nothing about dimensions of poverty that have no approximate measurements in the *Understanding Society* survey. These clusters may give us a *more* multidimensional measure, but not a *complete* multidimensional one. There are dimensions of poverty from theoretical literature that are completely absent, although some of their variance may be captured by the dimensions that are measured. These include: equity and

quality of local services; dependency; housing (structural) conditions; poverty of time; denial of human rights; diminished citizenship; and lack of voice. Some of these dimensions, such as housing conditions and local environment, were identified as substantial dimensions in previous research (Tomlinson, et al., 2009). Others emerge consistently in more nuanced qualitative literature and literature with specific foci on social divisions and marginalised groups (Lister, 2004; Pantazis, et al. 2006; Alcock, 2006; Walker, 2014). Secondly, there are dimensions that are not measured explicitly, but are more likely to share a substantial amount of variance with the factors that can be derived. Many of these are likely to be captured within the life satisfaction and psychosocial health factors, and include shame and stigma; humiliation; disrespect; assault on dignity; lack of entitlement, and self-esteem; and powerlessness.

Missing factors that likely have low or unknown shared variance

The difficulty of accurately measuring quality of local services using the variables available in *Understanding Society* was discussed above. This potential factor was distinct from the others in this list due to the fact that it *is* measured, but without additional information cannot be isolated and turned into a metric in a meaningful way.

Dependency is a dimension of poverty highlighted in Lister's (2004) conceptualisation of poverty. A person is likely impoverished by the fact that, if the capacity to achieve something needs to be mediated through support from another, their sense of participation is likely to be diminished, and feelings of shame are likely to be heightened by their perceived or real lack of autonomy and independence. The measurement of dependency might be approximated through some more in-depth analysis of relationships to other members of the household, but in reality even this may fall short. For example, the Office for National Statistics produces data on the numbers of young adults aged 20 to 34 years that live with their parents (ONS, 2017). This criteria illustrating dependence might be replicable using the datasets' matrix of relationships between members of a household. Other criteria may also be developed to encapsulate whether someone requires continuous assistance in their daily life. However, such an attempt still may not accurately capture what we mean by dependence inflicted by poverty. A feeling of dependence is inferred by the above criteria, but whether it constitutes dependency is highly subjective. One potential candidate for the more direct measure of dependency may come from the Care Dependency Scale and its variants, a valid and reliable measure of dependence on

nursing related activities that measures dependency on others, particularly health professionals, for both physical and psychosocial activities (Dijkstra, et al., 2006; Piredda, et al., 2016).

Tomlinson, et al., (2008) were able to identify an underlying dimension of poverty which related to immediate environmental conditions concerning the quality of accommodation that households lived in. Whether a home has adequate space, facilities, and protection from the elements, including infestation, damp, and mould, relates to a set of questions that are no longer asked in the *Understanding Society* survey. The quality of housing in the UK has come under renewed scrutiny in the context of insufficient housebuilding under consecutive governments since 2010 and greater awareness of overcrowding and poor conditions in properties in London, with around three in ten people living in 'bad housing' in 2013 (Barnes, et al., 2013: 1). Bad housing here refers to housing that is overcrowded or fails to meet the Decent Homes Standard, and/or participants reported high levels of condensation, mould, or fungi (*ibid*: 4). This reflects a material dimension of poverty that may not always be captured by other material deprivation indicators, or by income, especially in regions such as London where wages may be high but affordable housing is not achievable even on a wage that is higher than the average for the country. Surveys seeking to collect data for a multidimensional poverty index could consider reincorporating variables that relate to the Decent Homes Standard, or use existing validated questions derived from the *English Housing Survey* (Ministry of Housing, Communities & Local Government, 2018).

Deprivation of time is a dimension of poverty that has not received a great deal of attention but may be increasingly crucial in the 21st century, considering the rise of precarious employment and zero-hour contracts in low-paid work (Standing, 2014; Newsome, et al., N.D.). Time deprivation and low income can be considered different manifestations of poverty, but the role of time is often ignored (Vickery, 1977; Burchardt, 2008): if the only means by which someone can escape income poverty is by contributing such a vast amount of their time to work and labour, can they really be said to not be living in poverty anymore? Such an unbalanced focus on work means that time for leisure, social activities, physical and mental recovery, and alleviating stress can be harmful for a person's mental wellbeing (Afonso, et al., 2017; Kuroda & Yamamoto, 2018). Metrics of poverty that focus solely on income ignore the fact that any movement away from income poverty may be a result of people moving into time poverty. In many cases time poverty and income poverty are often compounded. The prevailing narrative suggests that the income-poor may (at least) be time-rich and merely 'unwilling' to work,

which is highly reflective of historical narratives made to justify high levels of poverty as far back as the 18th and 19th century, that the poor in fortune were rich in vigor (Brisbane, 2012 [1840]). High and rising levels of 'in-work poverty' in the UK suggest otherwise (JRF, 2018). Left unmeasured, time poverty threatens to eclipse income poverty.

So far this discussion has only been limited to time poverty inflicted by a necessary excess of paid work, that could be easy to operationalise by using data on hours worked per week, but this would reflect a large underestimate of time poverty. Unpaid labour, such as time absorbed by adult or child caring responsibilities, constrains both peoples' capacity for activation in the labour market and places a restriction on their personal time, typically resulting in higher rates of income poverty (Barnardo's, 2012; Aldridge & Hughes, 2016). Because these caring responsibilities unevenly fall upon women, and the gender gap in labour market activation has closed far more than the gender gap in caring roles, they are at far greater risk of time poverty than men (Zilanawala, 2016). A consequence of this is that the dimensions of any higher-order construct of multidimensional poverty for women may be weighted very differently than one for men when time poverty is able to be measured. Time poverty may be far more central to the experience of poverty for women.

Denial of human rights and a diminished sense of citizenship is a dimension and experience of poverty that is also lacking in the multidimensional measure items. This is perhaps in part due to the difficulty of its operationalisation and the many forms it can take that can be hard to capture using survey methods. Human rights are rarely denied through an explicit altercation, with one party trying to exert their rights and the other denying them. Rather, denial of human rights tends to be more insidious. People are morely likely deprived of their rights due to obfuscation and misdirection than force; they are coerced into believing they do not have the rights given to them under law, especially when these relate to the right to dignity and respect (Honneth, 1995; Lister, 2015). Monitoring this kind of denial of human rights is difficult and not often undertaken in survey research as a person is unlikely to know whether they have been denied a certain right if they are not aware of the right in the first place. It may be possible to assess this if a series of questions were constructed that query whether someone has been subject to common denials of human rights, or to gauge their ability to assert their rights using hypotheticals. However, even this risks ignoring the more general experience of diminished citizenship.

Another difficult to capture dimension is that of voicelessness in politics and society. Voicelessness is derived from a two way relationship requiring a number of steps. To be measured it requires firstly knowledge of whether a person feels empowered and safe enough to speak. Secondly, it requires the knowledge of whether what was said was meaningfully heard by the intended recipient. This means a sense of being actively listened to, with the intended recipient open, interested, and serious about what was being expressed. Thirdly, we have to understand to what extent what was voiced and heard was then acted upon in a meaningful way. A deficit in any of these stages reflects a lack of respect and recognition resulting in voicelessness (Lister, 2006). Voicelessness might be captured by a real and perceived lack of representation of disadvantaged groups in decision-making, resulting in a sense of ennui about ones position and value in society, but this is very difficult to quantify (Fraser, 1997). Recent efforts to facilitate the amplification of the voice of people with lived experience of poverty by organisations such as ATD Fourth World have been central to countering this, and highly meaningful for those involved (ATD Fourth World, N.D.).

Missing factors that may be captured well by existing items

The remaining unobserved factors that feature heavily in theories based on the lived experiences of poverty can be argued to be captured to a certain extent within the psychosocial strain cluster of items. Nontheless, it is important to highlight that no specific collection of variables within the data appeared to represent them exactly. A measure of poverty that is truly multidimensional should seek to explicitly model these factors using reliable measurement. Their discriminant validity can then be tested against psychosocial dimensions based on the GHQ and WEMWBS to assess whether they are sufficiently unique concepts, or measured in a way that makes them distinct.

The consequences of shame and stigma, humiliation, disrespect, and assaults on dignity, are likely to be reflected to a certain extent in the psychosocial strain and life satisfaction dimensions of poverty: greater levels of anxiety and depression, diminished sense of control, and more social dysfunction (Velotti, et al., 2017). However, it should not be assumed that this will certainly be the case. Responses to stigma, humiliation, disrespect, and assaults on dignity are likely to be diverse and the psychosocial consequences sometimes associated with them should be not taken as an unproblematic approximation. Being subject to stigmatising environments and behaviours is not the same as the resultant psychological damage, which may also have been caused or compounded by a number of other factors that are not related to

poverty. This dimension of poverty can therefore only be seen through the lens of psychosocial strain and life satisfaction.

Another missing dimension, entitlement, is an extension of shame and goes beyond the experience of feeling and having shame elcitied by stigma and relates more closely to how poverty impinges a person's sense of self-worth – a sense of being deserving of not being ashamed. In other words, a sense of worth that you deserve dignity and respect by virtue of being human (Lister, 2015). Even those in societies with little stigma attached to the receipt of income support and benefits can suffer from a sense of lacking entitlement. This is documented succintly by three lone-parents receiving social assistance in Norway (Walker, 2014: 107):

> *"It's shame… one has to experience it to say it… I'm a burden on other people… Can anyone be proud of going to the social assistance office and asking for money? I don't think so*
>
> *… you start asking the sort of questions like 'Aren't I good enough?' You start to have doubts.*
>
> *… I don't really see myself as worthy enough… It has to do with self-confidence or self-image maybe. What is it [people] talk about?"*
>
> Walker (2014) *The Shame of Poverty*. 107

This may be a consequence of the internalisation of powerlessness that comes with poverty, a psychosocial need for oneself to feel valued and valuable by the viture of being human. This dimension is therefore closely related to, but distinct from, shame. It may be captured to some extent within the psychosocial dimensions and life satisfaction dimensions of the multidimensional measure, but some of its relational identity to wider society and welfare or other kinds of support is likely to go uncaptured.

Lastly, is a sense of powerlessness which characterises experiences of living in poverty. If lack of entitlement reflects the internalisation of powerlessness and stigma about a person's worth, powerlessness might reflect the enforced lack of capacity to escape being at risk of poverty. This is the sense of entrapment that encompasses poverty. It is evident not only in the dynamics of recurrent poverty (Tomlinson & Walker, 2010), but in the existence of a poverty premium (Tinson, et al., 2014; Davies, et al., 2016). Even when families escape poverty this can often be short lived, with recurrent spells of poverty compounding a sense of hopelessness. Households in poverty are subject to higher costs for essentials an estimated poverty premium of £490 per year (Davies, et al., 2016: 8). Faced with an increasingly punitive and coercive

welfare system and limited regulations around zero-hour contracts and precarious work places creates a greater share of power in the hands of employers and individualises risk, centering it on workers and families (Carter, 1998; Hewison, 2016; Reeve, 2017). These experiences may be captured passively by psychosocial strain and life satisfaction, but are not explicitly measured.

4.2 The politics of survey design

These dimensions are not missing randomly. The political nature of the funding and administration of large surveys should not be ignored. That is not to say that those commissioning, designing, and administering the survey deliberately choose to omit attempting to measure certain things. Rather, these things go unmeasured partly through a replication of decisions that downplay factors of importance to those with lived experience in favour of those that are important to policy actors: survey design is typically a very top-down process. Because surveys are partially about understanding society and 'things' (Jenkins, 2002) and partially about identifying and finding solutions for social problems (Becker, et al., 2012), certain actors have privilege and sway over their development. This includes those in government, civil service and societies, but also think-tanks and established experts. Longitudinal studies also follow a certain path dependency because of the need for consistency of variables over time – if these variables were decided at a time where the research zeitgeist was different, these will continue to take up space in the survey that could otherwise be used for new questions developed from the bottom-up.

It is important to acknowledge that power is not only held by actors either, but also by ideas. Within the context of previously described epistemic authority (Kruglanski, 1980) is the point that there are inequalities between the authority of different disciplines and topics, and many of these surveys are interdisciplinary in nature. Questions of public health, associated with medicine and a positivist ontological standpoint, and more loosely with macroeconomics, may carry more weight than, say, questions about happiness. Indeed, this may be why, when questions of happiness are raised and justified, it is often within a clinical or economic perspective, not a humanitarian or philosophical one. Similarly, medicine is often framed by the diagnosis of illnesses or diseases and their subsequent cures or prevention. The parallel is obvious when we consider that in social policy discourse the focus is typically on social problems, not social quality (Van Der Maesen & Walker, 2012). Despite the latter potentially having more longevity and utility, it does not fit the dominant epistemological narrative (Van

Der Maesen & Walker, 2012). A problem is synonymous with an illness, and a cure can be found for an illness. A moral responsibility for striving to better something that can be better is harder to fit within the same narrative. Without strong ideas, a narrative to support them, and actors in position of power to push them forward (Béland, 2005), they are unlikely to make the shortlist for a large survey questionnaire.

Survey design is therefore a game of prioritisation. The longer a questionnaire becomes, the higher the rate of attrition and non-response: space on the form is at a premium (Rolstad, et al., 2011). As with anything at a premium, it is power and influence that dictates who attains it, whether this is the power of the argument for a certain question to be included, or the influence of the person requesting the inclusion. This makes them a particularly difficult tool to use when trying to test theoretical constructs derived by working with people with lived experience. Establishing the reliability and validity of factors from existing variables, as imperfect as they may be, is a significant first step towards creating space and strong arguments for more measures relating to multiple dimensions of poverty in large surveys. If the quantitative measurement of poverty is to move beyond the 'semantic and statistical squabble' (Piachaud, 1987: 162) it is essential to make an effort to show that operationalised measures based on theories developed from the ground-up are possible and work well. Without this, quantitative definitions of poverty are fated to rotate around the same top-down, income-centric, oversimplified, measures that do little justice to the real experience of poverty.

4.3 Conclusions of Exploratory Analysis

This section has aimed to show that interrelated, underlying dimensions of poverty can be identified within variables in *Understanding Society*, including amongst many clusters of items which were not originally intended to measure latent constructs of poverty. Furthermore, it aimed to explore whether clusters of items could reasonably be explained by theories of poverty that have been developed from extensive qualitative research into the lived experience of poverty. The interactions between these underlying individual dimensions build evidence for the claim that the factors can be reasonably considered to be related to a higher order factor that refers to 'multidimensional poverty'. These clusters and their interactions with one another have been identified and the interpretations have been supported by literature that has been reviewed. Methodologically, this section has aimed to demonstrate that a visualisation based approach to exploratory factor analysis that utilises correlation plots and hierarchical clustering of items for large scale matrices is a powerful alternative or complimentary approach to

traditional EFA which fits within the wider movement in quantitative social research to develop working practices that avoid spurious conclusions based only on *p*-values or other criteria that is more sensitive to sample size than effect size (Gelman & Loken, 2013). Lastly, the political aspects of survey design and elements in theory that are missing from large surveys were considered, and arguments were made for their inclusion.

Twelve potential factors were identified from the clustering of items in the correlation plot. Three of these were not relevant to dimensions of poverty within theoretical literature, and are more likely to reflect differences between specific groups: these were socioeconomic status, pensioner communalities, and the clustering of a specific group variables that included retired couple home owning households. The specific reasons as to why these items may have clustered together and have strong relationships with poverty were considered in reference to literature on the social divisions and inequalities between groups and their risk of poverty. These were not taken forward to the stage of invariance testing as they were not deemed dimensions of poverty, *per se*.

Four further potential clusters were identified that, although related to one another, did not seem to be related to a higher-order poverty factor based on the visible interactions with other clusters of items. These related to environmental quality of local services, civic participation, community cohesion, and social isolation. Either the measurements used for these factors did not capture the specific dimensions of poverty they relate to, or theories of poverty that mention these kinds of dimensions should be revised. Local environment, civic participation, and social isolation feature in previous multidimensional models of poverty (Tomlinson, et al., 2008), so were retained to be assessed, and are reported in the following chapter. Community cohesion does not feature strongly in theories of poverty nor in previous attempts to create a multidimensional measure of poverty, neither is there a strong argument in the literature or exploratory visualisation to propose it as a dimension, so was not taken forward to invariance testing.

Five clusters of items seemed to reflect factors that both had strongly related items and showed significant amounts of interaction with other factors that were identified by the literature, indicating the existence of a higher-order construct that could reflect multidimensional poverty. These were named financial strain, material deprivation of commodities, material deprivation of consumables, life satisfaction, and psychosocial strain. All of these potential dimensions

bear a strong resemblence to dimensions in theories of poverty, accounts of lived experience, and recent work on powerlessness, shame, and social justice within the academic literature (Walker, 2014; Lister, 2015). Furthermore, the high correlations between these clusters of items suggests a potentially good-fitting higher-order factor can be constructed. There is some debate as to whether some of the clusters of items should be treated as more than one factor because of a lack of visual distinction. For example, psychosocial strain might be considered as one factor based on the correlation plot, but is frequently separated into three distinct factors in existing literature. Tests of discriminant validity are completed in the following section to consider this.

This analysis demonstrates that the psychosocial and relational consequences of poverty appear to be as important in the overall constitution of poverty, or at least some overarching experience, rather than simply being causes or consequences. Feelings of shame, stigma, and the subsequent psychological toll these exert on people are so acutely felt in qualitative explorations of what living in poverty feels like (Walker, 2014; ATD, N.D.; JRF, 2017; Shildrick, 2018), but are conspicuously absent in poverty metrics, despite the tools and data being available. As will be shown, these factors can be included in a multidimensional measure, but the political will is often not there to implement them, something which has recently been challenged by suggestions from the Social Metrics Commission (2018). This exploration shows evidence that these should not be so easily dismissed in favour of income focused metrics. Finding a way to integrate these experiences is essential.

This section therefore forms a strong foundation on which to build a multidimensional poverty index using factor analysis. When existing questionnaire instruments for factors do not exist and there are many possible inclusions for a complex model the hierarchically clustered correlation plot demonstrates a quite intuitive way of taking a very large number of items and deriving which may form realistic factors. Meanwhile, it avoids the pitfalls of relying on significance tests (Gelman & Loken, 2013). To repeat the same identification of factors using iterative exploratory factor analysis would not only be very time consuming, involving the culling of over a hundred variables by the end, but would also be computationally difficult. This is largely because of the large number of different types of variables that are simultaneously predicted. Some are binary, and require a logit or probit link, some are ordinal and require a probit link, others are non-normally distributed. The combination of a single non-

parametric correlation matrix converted into a colour spectrum, a distance matrix to cluster items on to aid pattern recognition, works very well.

Nonetheless, this exercise and the end results are far from flawless. The lack of measurements for critical dimensions of poverty is largely due to the restrictions of having to piece together dimensions that fit theory from existing data and existing instruments. Any future efforts to develop a truly multidimensional measure of poverty might push for the development and implementation of a comprehensive index of items that can be administered easily within large survey designs, similar to how the General Health Questionnaire or the Warwick-Edinburgh Mental Wellbeing Scale have been developed by psychologists. However, space on the questionnaire of a large survey is at a premium, time is precious and inclusion of items is political. In order to routinely collect the data needed for a multidimensional poverty index sociologists need to engage with the tools of invariance, reliability, and validity testing that demonstrate that such survey instruments are worthy of regular inclusion. This forms the central contribution of the following chapter.

Chapter Five: A Valid and Reliable Measure of Multidimensional Poverty?

This chapter begins by reviewing the importance of invariance testing from a policy perspective, answering the question: why is it important that our measures can be demonstrated to be reliable and valid? This is framed within Tony Atkinson (1985) and Peter Townsend's (2010 [2002]) critical discussions about the role that poverty metrics play and their inherent political nature. Further, it argues that Townsend's (1979) idea of poverty as a dynamic construct necessitates such a formalised procedure by which to assess its changing shape. A more complex methodology is necessary not only because it provides a more accurate measure of poverty, but because it moves us away from circular reimaginings that can be dictated by politics, not by methodological quality.

Within this context, the proposed novel method for assessing measurement invariance – that is, the stability of a measure between different groups or, more essentially here, over time – of dynamic factors is outlined. The results of applying this method to each of the factor analyses of the dimensions of poverty identified in the preceding chapter are then presented, followed by the application of a validity testing method to determine whether these ontologically distinct factors are sufficiently discriminated by their measurement variables. Finally, a poverty measure constructed using these methods is presented and its implications are discussed.

5.1.1 Constructing and testing the invariance of dynamic concepts

Existing research has already shown that the construction of a multidimensional poverty measure using methods found within a structural equation modelling framework is possible (Tomlinson, et al., 2008). By using the latent variables as simultaneous predictors of multiple outcomes it is possible to incorporate and weight proportionally both monetary and non-monetary factors with the use of a single scale. This frees poverty metrics from being confined to the context of baskets of goods that can be reduced to income, allowing measurements to capture more of the key sociorelational and sociopsychological dimensions that are frequently reported in theories of poverty and narratives from lived experience. However, a robust testing of the reliability and validity of this multidimensional measure has not yet been undertaken.

One reason for the success and widespread adoption of many psychometric measures has been on their creators' extensive testing of reliability and validity using consistent criteria. While significant progress has been made towards developing and arguing for more multidimensional measures, such as in work by ATD Fourth World (N.D.) globally, or by the Social Metrics

Commission (2018) nationally, arguments in support of these measures still rely largely on face validity created by consensus and in cooperation with those with lived experience. This is obviously a vital and important step in developing any indicator or measure, but only goes so far. As a result, measures need only prove face and content validity by consulting expert and layperson assessments on the measure (DeVon, et al., 2007) to be seen as credible alternatives. In other words, if a measure is seen as sensible it is seen as equally valid. This may explain such reluctance to move away from easily implemented but blunt measures like the HBAI metric.

Poverty, however, is political. If a poverty measure claims to better represent poverty due to it matching more closely with a public and expert perception – having high face validity – it is still likely to be rejected by those who disagree. If a new poverty measure finds that poverty is a much smaller problem than previously thought, or has been following a stronger downward trajectory even under austerity, proponents of more neoliberal or conservative political inclinations are likely to use this as support for the argument that poverty does not exist in the UK – in the same way that some government ministers responded to the UN special rapporteur's report (BBC, 2018). At the same time, those critical of neoliberal and conservative policies may be more likely to reject this measure as methodologically flawed, without any explicit reference to a methodology. Similarly, if a poverty metric claims very high and increasing levels of poverty, as the JRF's Minimum Income Standard has, it is likely to be met with skepticism and criticism by some, especially when items have been qualitatively derived (Snowdon, 2014). As Atkinson (1985: 160-17) writes: "it would patently be preferable to specify a single poverty standard and hence obtain a clearcut measure of the extent of poverty. This would however represent an 'all or nothing' approach, since those who disagree with the standard are likely to reject the findings out of hand". Figure 14 shows the way that comparing only four measures - the HBAI, the JRF's Minimum Income Standard, and Tomlinson, et al.'s, (2008) multidimensional poverty measure – show drastically differing narratives of poverty in the UK. The main distinction between their favourability is likely to come from the political interpretation of their *narrative*, not from an assessment of their methodological robustness.

Debates could, and likely will, continue to run in these circles, due to the underlying political standpoints. One way of fracturing, if not breaking, this impasse is by adopting very concise, objectively applied, methods to test model fit and loading estimates, measurement invariance, reliability, and discriminant validity. These provide information on how good the underlying

theory behind the construction of our measurement is, and whether we can reasonably support the inclusion of all items or dimensions selected based on the prevalence of them commonly occuring together. If items can be shown with methodological objectivity to be logically linked together by some underlying construct that can be reasonably interpreted as poverty it is hard to argue that these dimensions do not belong in a poverty metric. If a measure can be shown to be salient through time and is reliable in its estimates, through the use of time-group invariance testing and exploratory/validation subsetting of data, it is hard to argue that the more complex measure has no potential longevity or application. Lastly, if dimensions can be shown to be linked to one another in ways that theory suggests, but are sufficiently distinct for them to be identified as legitimate facets of experience in their own right, an argument for validity can be made that extends beyond face validity – that different dimensions really do reflect different aspects of poverty, they are not merely different ways of measuring the same thing.

Invariance testing typically assumes that constructs are static in time and that being able to show this is a good thing. However, poverty is described by Townsend as a dynamic concept (Townsend, 1979). We should expect the composition of poverty and the features that make up how it is experienced to change with society over time. Indeed, if a metric did not reflect or have built into it adequate checks to see how well it fits within the shifting state of the concept it seeks to measure, it should be seen as a cause for concern, not celebration. Our concepts and ideas shift with time and context. In the same way that income-based metrics adjust for the real value of money over time, multidimensional measures also need to be able to adjust appropriately for the changing meanings or importance of dimensions.

Just because a measure is conceptualised as dynamic does not mean that there are not periods of relative stability between which meaningful comparisons can be made. If we know our measure is dynamic and is expected to change over time, it would make equally little sense to throw out a more complex measure just because it cannot be used for all time in all contexts. Surprising as it may be, invariance testing is not built around the idea that breakpoints where concepts change significantly enough to warrant changes in its measurement to bring it back in line with the wider context exist. This is partly because of its genesis in psychological research. This is an oversimplification, but features that can be measured in relation to the mind, emotion, and cognition, are theorised to be relatively stable over time in a societal sense (although not necessarily within an individual as they age) (Millsap, 2011). As such, it was

necessary to adapt the typical process of invariance testing by introducing subsets of fixed parameters based on different lengths of time.

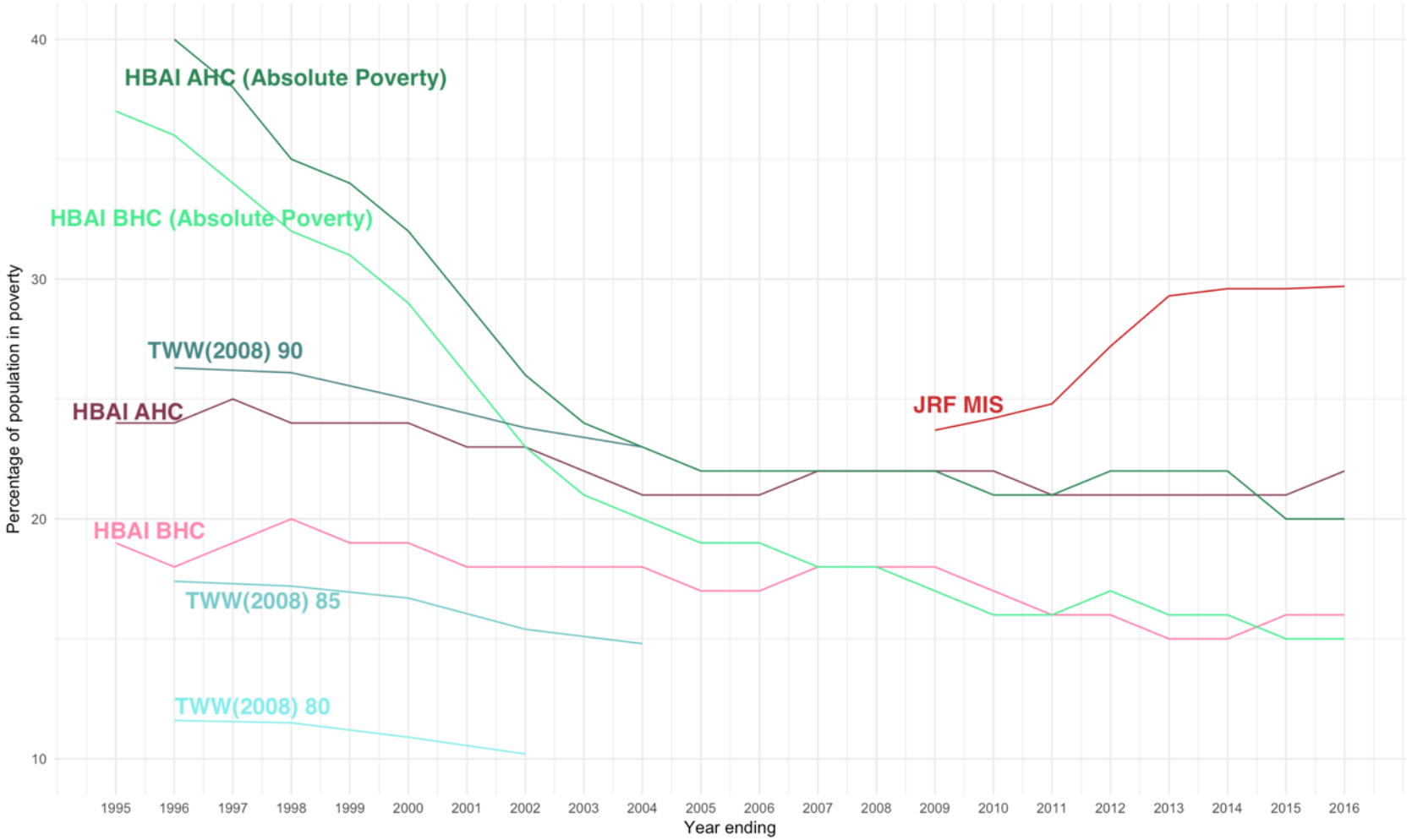


Figure 14: *Trajectories of poverty rates under different measures 1995-2016. Measures of poverty are far from in agreement when it comes to what the level of poverty is in the UK, and the choice of measure is largely dependent on subjective beliefs about the claims of face validity from each.*

5.1.2 Defining measurement invariance and discriminant validity for dynamic constructs

The terms used for testing measurement invariance are taken from psychometric literature (van de Schoot, et al. 2012; Millsap, 2011). Each level imposes additional restrictions. Briefly described, configural invariance refers to whether latent constructs can be 'made up of' the same observed variables regardless of time or group. Metric invariance refers to whether each of these observed variables can be weighted the same in the overall composition of the factor regardless of time or group. Scalar invariance, which is not used often here as it only refers to continuous observed variables, refers to whether people in different time points or different groups share the same intercept with the latent construct. An example helps illustrate this. Imagine that people living in London have, on average, higher incomes before housing costs regardless of the level of financial strain they are experiencing compared to people living outside of London. If this was the case, there would be quite a noticeable difference between the expected income of those living in London and those not living in London even when their score derived from the 'financial strain' was the same, say, zero. In this case, the construct of 'financial strain' does not have scalar invariance, because the way that income works as an observed variable that shares a linear relationship with the latent construct is different for different groups. Lastly, full invariance refers to whether all of the above criteria is met but, in addition, the residual errors in the predictions for each observed variable based on their derived score on the latent construct are the same regardless of group or time point. In other words, full invariance refers to when the underlying construct does a consistently good job of predicting all of its observed variables regardless of when in time or for which group they were measured.

Generic analysis scripts were written for seven competing types of factor invariance over time based on criteria from van de Schoot, et al., (2012). As each model produces a large amount of output detailing all parameters, a bespoke function was written that generates models testing several kinds of invariance and extracts only the statistics of interest from the text output so they could be transcribed into the tables throughout appendix 4. As this is a novel application of invariance testing the process is reported here as opposed to in the methodology section. This method here is used with respect to data having six time points, as this was the number of waves in *Understanding Society* at the time of the analysis. These criteria included:

- Configural invariance

    - This refers to whether factors can be said to be constructed with the same items regardless of time point or group. This implies that our metric can be measured using the same indicators regardless of time or group membership.
    - This criteria can be said to be met if the fit statistics reported meet the requirements for a 'good' fitting model (Kenny, 2014)
- Six-wave metric invariance
    - This stage tests whether factors can be said to have identical factor loadings for each item throughout all six time points. This implies that not only does our metric have the same configuration of items regardless of time or group membership, but that the items relate to the underlying scale in the same way. In other words, that the meaning of the latent construct is the same (Millsap, 2011).
    - This criteria can be said to be met if the fit statistics compared to the configural invariance model's fit statistics remain 'good'
- Two-by-three wave metric invariance
    - This stage tests whether factors can be said to have identical factor loadings for each item at fixed three-wave long time points (wave one to three and wave four to six)
    - This criteria can be said to be met if the fit statistics compared to the configural invariance model's fit statistics remain 'good'.
    - This level of invariance is only desirable if six-wave metric invariance has poor fit.
- Three-by-two wave metric invariance
    - This stage tests whether factors can be said to have identical factor loadings for each item at fixed two-wave long time points (wave one and two, waves three and four, and waves five and six).
    - This criteria can be said to be met if the fit statistics compared to the configural invariance model's fit statistics remain 'good'
    - This level of invariance is only desirable if two-by-three-wave metric invariance has poor fit.
- Six-wave full invariance
    - In addition to the previous criteria, full invariance tests whether the latent construct can be measured identically regardless of time point or group

membership (Millsap, 2011). This includes whether levels of underlying items have the same meaning and are explained to the same extent. In other words, the meaning and the scale of responses is the same regardless of time or group.
    - This criteria can be said to be met if the fit statistics compared to the metric invariance model's fit statistics remain 'good'
- Two-by-three wave full invariance
    - In addition to the previous criteria, full invariance tests whether the latent construct can be measured identically at two fixed three-wave long intervals of time points (wave one to three and wave four to six).
    - This criteria can be said to be met if the fit statistics compared to the metric invariance model's fit statistics remain 'good'
    - This criteria is only desirable if six-wave full invariance has poor fit.
- Three-by-two wave full invariance
    - In addition to the previous criteria, full invariance tests whether the latent construct can be measured identically at three fixed two-wave long intervals of time points (wave one and two, waves three and four, and waves five and six)
    - This criteria can be said to be met if the fit statistics compared to the metric invariance model's fit statistics remain 'good'.
    - This criteria is only desirable if two-by-three-wave full invariance has poor fit.

To illustrate this with an example, take a factor for financial strain as we have above, as characterised by the ability to not miss any housing payments, energy payments, or council tax payments, plus a subjective measure of financial wellbeing and the household income. Imagine that we asked these questions repeatedly over a six year period. But now imagine that, in the middle of that six year period, council tax had been abolished for families with earnings lower than the first income tax bracket. From that moment on, the inability to keep up with council tax payments would no longer share much of an association with many of the other factors. In technical terms, its factor loading would decrease substantially. The nature of financial strain therefore has changed, and a method needs to be used to identify that change. Without the above method, we would conclude that financial strain is not a stable construct, or that the ability to keep up with council tax was not a good measure of financial strain despite that being the case for half of the time. With the above method, we can conclude instead that financial strain is a stable construct during the first three waves with the inclusion of council tax

payments; but is also a stable construct in the latter three waves with the removal of council tax payments. Financial strain is a dynamic construct, the concept itself changes *as the financial expectations and obligations in society change*. What we would expect to see following the steps above is that the factor does not show six-wave full or metric invariance, but it would likely show three-by-two wave metric, or possibly even full, invariance. Each of these groupings – three-by-two, two-by-three, etc. – can be thought of as 'segments' of time that can be calculated from the total number of repeated surveys. The full span of time can be broken down into segments, and each of these segments can then be tested for measurement invariance.

Anything less than a three-by-two wave metric model in a set of longitudinal data with six time points is equal to the configural model. This omnibus of tests therefore identifies not only whether our factors are stable over time, but at which points they change. This is a key methodological development that reflects Peter Townsend's (1979) theory of poverty. The definition should change over time, therefore the methodological process should reflect this flexibility of identifying when the configurations of multidimensional metrics should be changed to reflect changes in society. Such a consideration is not typically made in recent literature on invariance testing, presumably because of the nature of the constructs it has been developed to test the stability of. The most desirable outcome is for them to be stable regardless of external societal change. This is often the case for psychometric factors such as those designed to measure mental health, narcissism, psychosocial wellbeing, and so on; they are not assumed to change relative to external changes, except perhaps over exceptionally long periods of time that result in evolutionary change. In contrast, the systematic investigation used here may be a helpful template for metrics that are more sensitive to societal change, or are based on relativist assumptions: metrics that are only meaningful within their broader contexts. This methodological framework might therefore be applied to any number of sociological investigations.

Because of this, the results of invariance tests can be split into two outcomes, one which always reflects the inadequacy of the metric and one which may infer weaknesses in the metric but may also simply reflect social change. I refer to the former as the traditional measurement invariance and the latter as *segmented stability over time*. Table 5 shows a general way of thinking about outcomes of this style of invariance testing. Measurement invariance is the type of invariance that is of primary importance, and adequate measurement invariance should always be retained over 'better' segmented stability over time. For example, if a model with

two segments of equal length ($k = 2$) meets the criteria for full invariance it should be chosen over a model that has only scalar invariance when all waves have the same restrictions (full wave stability). The logic behind this choice is that measurement invariance reflects a technical requirement for the consistent interpretation and meaning of underlying factors when measured longitudinally, but 'worse' segmented stability over time can indicate *either* a problem with underlying items *or* the changing nature of the meaning of a factor over time. Therefore, it is better to prioritise the former.

Table 5: Table showing the generalised interpretation of different invariance criteria

Importance	Primary	Secondary
	Measurement invariance (van de Schoot, et al., 2012)	**Segmented stability over time**
Better	Full invariance	Full (n-wave) stability
	Scalar invariance	$\frac{n}{(\frac{n}{k})}$ -by-k wave invariance where $k = 2$*
	Metric invariance	$\frac{n}{(\frac{n}{k})}$ -by-k wave invariance where $k = 3$*
		… $k \rightarrow n$
Worse	Configural invariance	Configural stability ($k = n$)

* Where n is equal to the total number of waves and k is equal to the number of equal sized segments of time.

As previously mentioned, usually an additional state of variance is also reported, scalar invariance (van de Schoot, et al., 2012). This is included in the table above for general purposes, but because most factors in this study used items that were measured on an ordinal or categorical scale their predictions are based on their categorical thresholds, not on their means. This means that their relationship to the latent variable is not expressed in a continuous linear model, therefore the intercepts do not have the same meaning and are predetermined to be equal between time and groups anyway. In more technical terms, their thresholds are already fixed to 1, because an increase in 1 on any of the variables refers to a change in category, e.g. from 'did not miss any bill payments' to 'did miss bill payments'. They also do not have variance parameters since categories have fixed meanings. These types of variables were the most commonly used, with the exception of income. As a consequence in the vast majority of models (over 95% of those in the appendices) a strict invariance model would have exactly the same structure as a full invariance model and invariance testing progressed from metric to full invariance directly. This is why scalar invariance is referenced here, but is not reported in the appendices.

Chi-squared difference tests are traditionally used for reporting change in model fit in invariance testing, however, although they are reported in the appendices it is clear that they cannot be relied upon because of the large sample size (Kenny, 2014). This is a weakness of the current methods when using SEM with large samples in complex survey data. Furthermore, it was found that very often when testing many groups with very large sample sizes fit statistics actually improved after imposing invariance constraints on the models. This may be because of penalties for parameters that often feature in the calculation of fit indices, where models are 'punished' for their lack of parsimony. The technical reasons for this are beyond my mathematical knowledge; all I can report is that it is something that happens. These issues are of substantive importance however, in that if it is common practice for models to be created that do not fix parameters and loadings to be equal across time, perfectly valid, even good, models may be being rejected due to their poor fit despite this being an artefact of the calculation of the fit indices and the role of large sample sizes and inadequate invariance constraints. Examples of this are shown throughout the reporting of the results for each dimension, and in the appendices.

Lastly, the Average Variance Extracted for each dimension is reported, which can be used to assess its discriminant validity (Fornell & Larcker, 1981). This value can be interpreted in the same way as an R^2 statistic. Values close to zero indicate a very low amount of variance extracted, values between 0.2 and 0.4 suggest a weak amount of variance extracted, values between 0.4 and 0.6 might indicate a moderate amount of variance extracted, values between 0.6 and 0.8 can be considered a high level of variance extracted, and values greater than 0.8 indicate a very high amount of variance extracted.

The Fornell-Larcker criterion indicates whether we can strongly assert that the dimensions we are trying to measure can really be considered independently identified using our measures. To use an analogy, imagine that you are stood on a hill and interested in identifying something on the horizon. You take a telescope and use this to aid your observation of the object on the hill. If your telescope is very good and gives you a clear view of what lies on the horizon, you will be able to identify very strongly what the object on the horizon is: how many sides it has, how many dimensions, and what makes it different to similar objects around it. If your telescope is very weak, it doesn't give you a close enough view, or the lens is smudged, you may be able to have an indication of what the object on the horizon is, maybe even parts of its shape, but

you might not be able to pick out the kind of details that distinguish it from the other objects surrounding it. In this analogy the object is the construct you wish to observe and the telescope is your tools of observation, or measurement. If the object is very large and very different from things that are similar to it, or if your observation is very concise and clear, you are able to confidently define what it is and what makes it different in relation to other, similar items.

5.2 Analysis and consequences of missingness and data imputation

Before assessing invariance, it is important to inspect the summary statistics for data derived from multiple imputation, including highlighting where it seemed successful and points where it may not have been. Multiple imputation of missing data requires all missing estimates to be stable – this is referred to as convergence. In non-technical terms, the missing data estimator creates its first 'best guess' of the values of missing data. This is the first 'iteration'. It then refines its guess in the second iteration using its new guesses for extra information. It does this again and again until the researcher is able to determine that its estimates, or guesses, are not changing substantially any longer and are just 'bouncing' between a plausible range of guesses that are at the limits of its ability to estimate. To determine whether this has happened the researcher usually consults graphs showing some summary statistics for the variables with missing data, these graphs are reproduced in the appendices and an example is shown below. Figure 15 shows where some variables seem to have converged (left) and where others have not (right). Variables where imputations have not converged may be problematic, and results from models that suggest a lack of convergence in their imputation should be treated with adequate caution until full data with no missingness or data with imputation convergence can be obtained and the model re-tested.

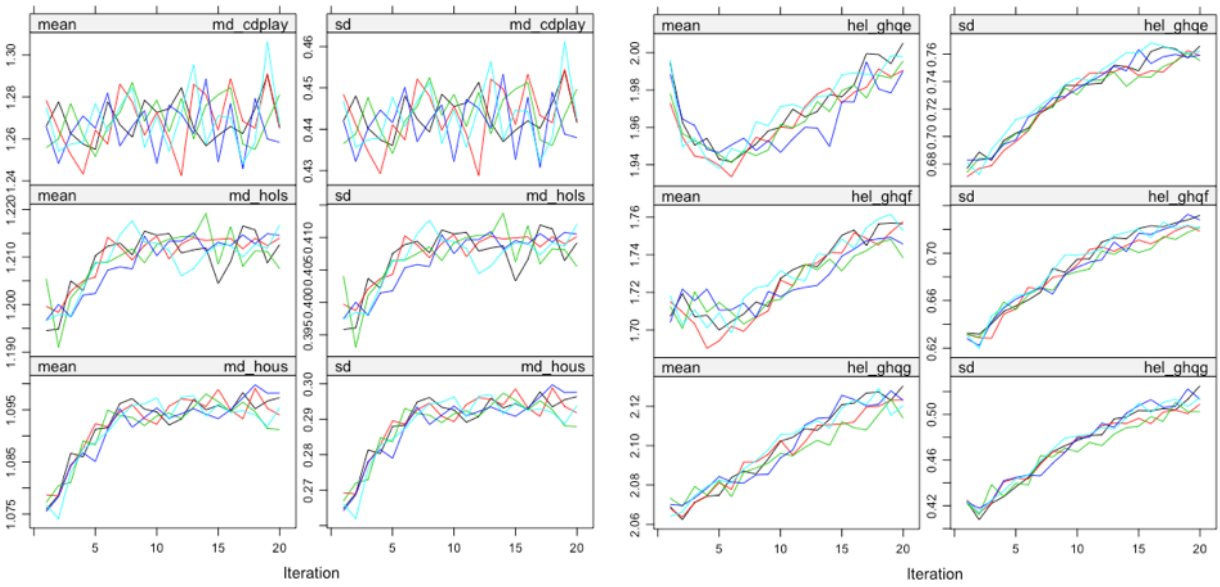


Figure 15: *Example of multiple imputation by chained equation mean and standard deviation convergence after twenty iterations using the MICE package. The left hand plot shows imputation estimates appear to have converged, the right hand plot shows variables where this may not have been the case. MICE convergence for all variables can be seen in appendix 3.*

Item sets from two of the previously identified clusters showed no strong evidence of summary statistics convergence after twenty iterations of five multiple imputations. These were questions relating to the General Health Questionnaire, prefixed with *hel_*, and questions relating to social isolation from personal friendships and relationships, prefixed with *soc_*. Inspecting these plots (appendix 3) shows that means and standard deviations for the imputations of GHQ items continued to increase with no sign of stabilising, however, the difference between the mean estimates in iteration one and iteration twenty appeared to be very small, changing by less than 0.05 on a four-point scale. Similarly, for questions associated with the social isolation dimension that related to personal friendships and relationships there was a strong case of non-convergence, although the difference between the summary statistics in iteration one and iteration twenty was also very small, less than 0.1 in many cases on a five-point scale.

Items that related to the life satisfaction cluster, prefixed by *lsat_*, and items from the WEBWM scale, prefixed with *psy_*, also showed some evidence of non-convergence. For life satisfaction, there was some evidence of convergence around the means of items, but less evidence when observing changes in standard deviations. This suggests that imputed data for life satisfaction

questions began with little deviation but continued to become more dispersed without stabilising with subsequent iterations. It is unclear what may have caused this, but the change in standard deviation estimates between iteration one and iteration twenty is again very small, around 0.06 on a seven-point scale for all items. Items from the WEBWMS scale began to converge by iteration twenty but there could be differing interpretations of whether this was the case, an additional five iterations may have made this more clear.

This is likely because of high multicollinearity between questions items, especially for the GHQ scale, and the high levels of non-random missingness in these item sets (appendix 1). In this case, what happens is a near-infinite possible number of combinations for imputations in these items exist, because they are so closely related to one another it is not possible to assign their specific values with confidence. Other GHQ questions and their most recent imputed values are used in each iteration to compute new values for the multiple imputation and it is likely that they created a loop in which any increases in one item resulted in artificial increases in another, which resulted in artificial increased predictions in a third, and so on, continuously. For this reason, it may be better to exclude highly correlated variables when specifying models for multiple imputation by setting some criteria that predictors must have correlations between 0.2 and 0.8, for instance, rather than by just having a lower threshold. An alternative would be to run imputation for factor scores and construct factors first, as opposed to imputing items directly, as factors combine items with high commonalities. However, this is less preferable for two reasons outlined in the methods chapter. The first reason is that Gottschall, et al., (2012) found that factor estimates were more efficient when item-level imputation was used, rather than scale-level factor score imputation. The second reason, as outlined when discussing the implications on missingness, is that models may fail to converge or be impossible without some form of missing data treatment so factor scores would not be possible to derive before imputation. Furthermore, even when it is possible, the use of listwise deletion frequently introduces bias into estimates and imputation of missing values has been shown to always be preferable to listwise or pairwise deletion (Enders & Bandalos, 2001; Acock, 2005).

The problem of biased imputations imposed by multicollinearity in MICE seems to become more substantial the more iterations that are performed. This is counterintuitive to the logic of MICE, where higher numbers of imputations usually result in more accurate estimates after they stabilise. On the contrary, it appears to be the case that if the iterations were able to be repeated indefinitely with highly correlated variables all imputations for missing data would

eventually hit the ceiling (if highly positively correlation) or floor (if highly negatively correlated) value of the scale. A compromise used here has been to allow iterations to run for as long as possible to give confidence that variables without multicollinearity problems have converged, but to run no more iterations beyond this. This is far from perfect but the exploration of multiple different approaches needed to be limited because of the substantial amount of time that MICE takes for imputing values for so many variables with such a large sample size.

Nevertheless, even if this problem is apparent it means that factors should still be largely representative of their underlying concept, but that the interpretation of the weighting of specific items, and any predictions made using these, should be treated with caution because their imputations may not be accurate. However, because the focus of this thesis is on underlying factors, not predicting discrete variables, multicollinearity in the models that multiple imputation is based on is less of a concern, just as long as the iterations have not been allowed to run on so long that the problematic predictions stabilise at the maximum value. It is worth acknowledging, though, that the imputed dataset has limitations for other types of analysis. However, for this study a benefit of latent variables in SEM frameworks is that we are more interested in the predicted values of unobserved constructs than the specific values of observed ones.

Other items that were used in the final models of multidimensional poverty appeared to show summary statistic convergence after twenty iterations, based on a visual inspection. One original methodological aim was to observe whether any obvious biases seemed to exist depending on whether models used listwise deletion, multiple imputation by chained equations (MICE), or full information maximum likelihood (FIML). However, only one set of variables relating to material deprivation of household goods and commodities had enough complete cases for the model to converge for listwise deletion and FIML. In this one example (appendix 4.3.1), listwise deletion and FIML yielded marginally better fit statistics than the MICE imputations (listwise = ~+.004 CFI, ~+.005 TLI, ~ -.02 RMSEA, ~ -1 WRMR; FIML = ~ +.001 CFI, ~ +.001 TLI, ~ -.001 RMSEA, ~ -0.3 WRMR). In the context of fit statistics these increases are negligible (Kenny, 2014). The factor loadings derived from FIML were virtually identical to the factor loadings from any of the imputed data, with no deviations greater than ±0.01. Listwise deletion also resulted in virtually identical factor loading estimates, with deviations no greater than ±.015. From this limited evidence it does not appear to be the case that factor loading estimates or model fit differs substantially depending on the treatment of

missing data, although this does not mean that the estimates are not different to what they would have been if there was no missing data.

The implications of the analysis of missing data and imputation methods seem to suggest that multiple imputation by chained equations is a good, but not unproblematic, way to undertake scoping research into 'emergent' factors in complex survey data. Emergent factors here refers to the idea that the factors that have been developed are derived from clusters of variables that were largely not designed to capture one-single factor in the design of the survey. Missing data is likely to be a common problem when trying to model complex emergent factors using multiple variables because these variables are often missing from specific waves, as there is no design impetus for them to all be measured at once so that a factor can be derived. Any conclusions of predictions using these kinds of factors therefore should be treated with caution. Imputed data where missingness is systematically based on survey design (with rotating modules of questions) should be interpreted more as proof of what is possible, with associated potentialities, than a finished tool in itself.

Because multiple imputation is regression based, imputed data will inevitably centre more around the mean, likely leading to more conservative estimates that may underestimate the error of what is being measured. The more data that is missing the more underestimated any errors are likely to be. The following findings should therefore be treated with caution, but achieve their aim of showing that a truly multidimensional, reliable, and valid quantitative measure of poverty that reflects theory much more closely than many existing measures is possible, even if the data required to have confidence in it is considerably lacking. This demonstration of feasibility might be used as a catalyst for surveys to formalise more regular collection of data that makes up this 'emergent' multidimensional poverty factor. Following these prerequisite requirements to deal with missing data and develop a methodology to test the invariance of dynamic constructs it was possible to test each dimension of poverty for validity and reliability with some level of confidence.

5.3 Dimensions of Poverty

The following section presents the factor analyses of each dimension of poverty as identified, in line with the methodology outlined in Chapter 3. The factors are introduced and the relevant variables used to measure them are described briefly. Any differences between the original

factor structure, modelled using an exploratory subset of data, and the final factor solution, modelled on four confirmatory subsets and on all data, are stated. Next, the model fit statistics for different types of invariance measures are discussed, and the level of measurement invariance met and the overall fit of the model is explained. Factor loadings are discussed briefly in reference to the literature where they appear to show some interesting variations, and any particularly weak factor loadings are shown.

In total, sets of exploratory and confirmatory factor analysis, measurement invariance, and AVE calculations were run on twelve different potential dimensions:

- Financial strain
- Material deprivation, which was split into the following two factors:
    - Material deprivation of consumables
    - Material deprivation of commodities
- Financial pressure, replicating Tomlinson, et al. (2009), a second order factor based on the above three factors
- Social isolation
- Civic participation
- Life satisfaction
- The following three sub-dimensions of psychosocial strain:
    - Anxiety/Depression
    - Social Dysfunction
    - Control
- Psychosocial strain
- Environment (local services quality)

5.3.1 Financial Strain

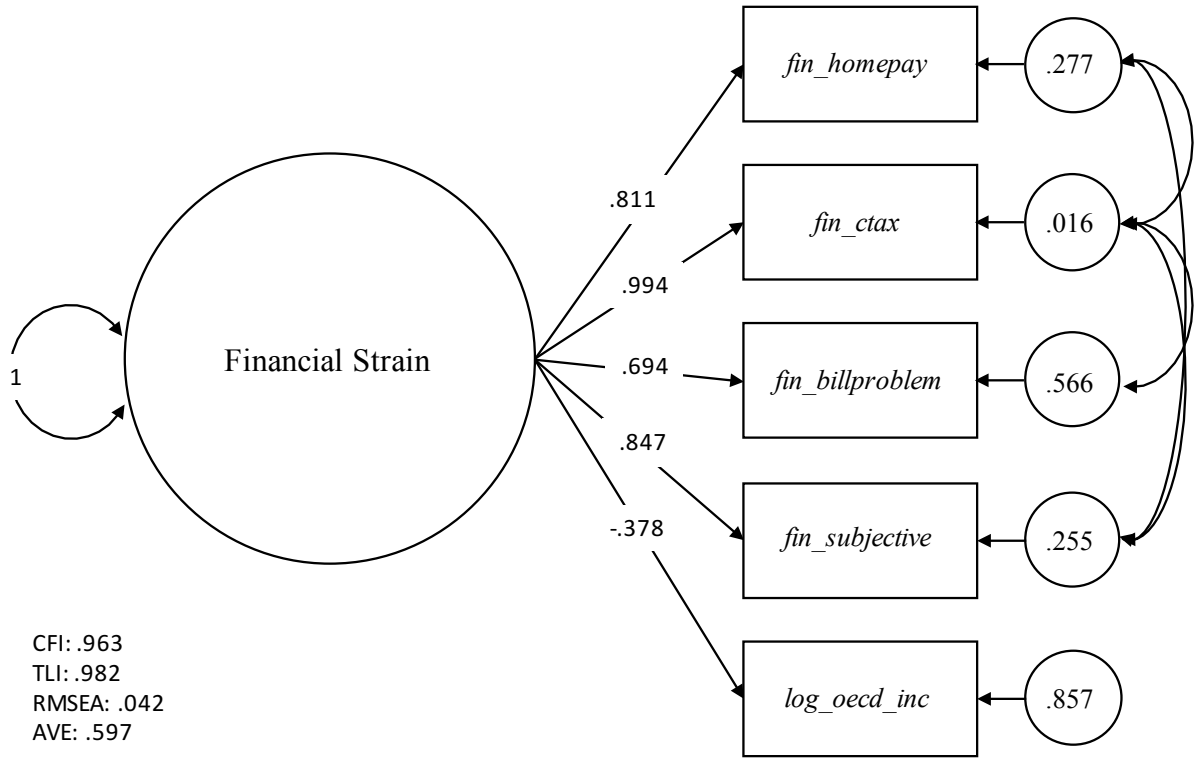


Figure 16: *Financial Strain CFA*

Financial strain represents the core of poverty, a state of hardship where a person or household lacks the financial resources to regularly make payments that are for essentials such as housing and the accompanying responsibilities. Financial strain is characterised by more and more frequent loss of ability to keep up with payments on council tax, rent or mortgage, and energy bills. It is also characterised by lower adjusted income, and assessed partly by the person's own subjective sense of their financial situation, which appears to act as a good barometer for how well they are coping financially.

This dimension was modelled firstly using an exploratory subset of data, where additional paths were specified for the residual error correlations between four of the variables. The factor structure was then confirmed to be reliable when modelled on a confirmatory subset of data. All imputations of the data gave very similar estimates (appendix 4). Iterative modelling with different constraints on invariance showed that the financial strain model fit the criteria for full invariance over time as specified in van de Schoot, et al. (2012) and outlined in section 6.1.2 (CFI = .963, TLI = .982, RMSEA = .042). This means that the factor is likely to be representative of a true underlying construct, not just a product of random noise, and can be interpreted in an identical way regardless of which of the six-waves it is applied to. The model

fit for the financial strain dimension was very good by the criteria specified by Kenny (2014), with a CFI and TLI in excess of .95, and an RMSEA of less than .05. This indicates that the underlying construct seems to be well approximated by the chosen variables. Furthermore, the model has a moderate to strong AVE.

Trouble keeping up with council tax bills has the single highest loading of .994, meaning almost all of its variance can be explained using factor scores from the underlying financial strain factor. This is most likely because if a person is having problems with their ability to pay bills and keep up with rent/mortgage payments they are very likely to have defaulted on their council tax bills prior to reaching this point. A person's subjective assessment of their financial situation was the second highest loading item, with a factor loading of .847, indicating that the assessments given by respondents appear to be fairly accurate when it comes to indicating prevalence of the other variables. Difficulty with rent or mortgage payments followed, with a factor loading of .811. Problems with paying energy bills had a lower loading at .694, which may be because the consequences of missing energy bill payments are more immediately enforced than the consequences of missing rent or mortgage payments, especially in the social rented market. It is therefore likely that these are maintained as a priority, as electricity, gas, and water are so essential. Lastly, the natural log of household income adjusted for household size had the least strong loading, and in a negative direction because of its coding. The reason for this smaller strength is likely to be found in differences between required income levels to meet housing costs in different areas. The level of income required to keep up with bills is likely to be different depending on where in the UK a person lives, and whether they live in an urban or rural area. This also highlights that even if income is equivalised and transformed to a normal distribution, it is not necessarily a good indicator of underlying financial strain, as criticisms of income-based metrics have shown before (Tomlinson, et al., 2010).

The financial strain dimension therefore appears to be a good-fitting and time-invariant construct when approximated using these variables. The pattern the loading of items has taken on seems to suggest that households under heavier financial strain are more likely to miss payments on council tax first and housing second, in favour of keeping the lights on and the boiler running. The extent to which income should be found in this dimension, or whether it is superfluous and too hard to adjust into a universally meaningful scale is debatable, as it does not add much to the variance extracted and the model fit is likely to improve were it to be removed.

5.3.2 Material Deprivation

Material deprivation was first tested using confirmatory factor analysis on an exploratory subset of data as a single factor containing all the items specified prior in the methodology chapter. This first arrangement of items led to a model with very poor fit (CFI = .869, TLI = .853, RMSEA = .077). A two-factor solution was then tested on the exploratory subset, splitting material deprivation items into the two clusters that were identified in the exploratory hierarchically clustered correlation plot phase: one broadly based on consumption and one broadly based on household commodities. This solution had very good fit statistics (CFI = .971, TLI = .967, RMSEA = .036). The two-factor solution for material deprivation was retained and each factor was then tested for measurement-invariance and discriminant validity separately.

5.3.3 Material Deprivation (Consumption)

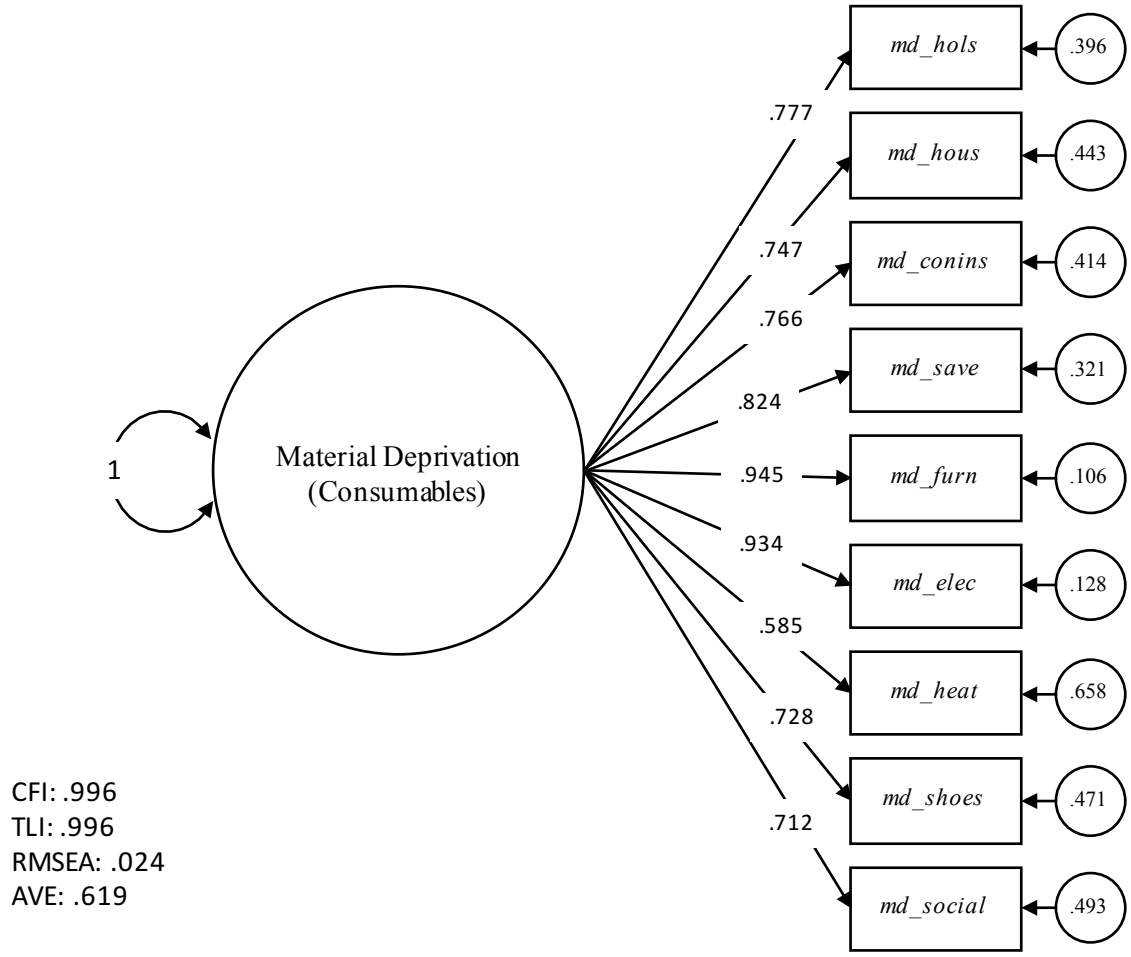


Figure 17: *Material Deprivation (Consumables) CFA*

Material deprivation of consumables is conceptualised as an underlying factor that most strongly represents the idea of relative deprivation, and indicators centre around activities,

goods, and services that represent current expectations of what is required to 'play the roles, participate in the relationships, and follow the customary behaviour' of the society that people live in (Townsend, 1979: 31). The latent variable predicts responses of 'we would like this, but can't afford it', to items that include whether the household are able to take a holiday for the duration of one-week, once a year, without staying with friends or family; whether they are able to keep their home in a decent state of repair; whether they are able to have contents insurance; whether they are able to save regularly; whether they are able to replace furnishings if they were to break; whether they can replace major electrical goods (electric ovens or fridge freezers) if they were to break; whether they can keep their home adequately warm; whether they are able to outfit each member of the family with two pairs of all-weather shoes; and, whether they can feed visitors at least once per month. These factors therefore represent various forms of participation, both social and cultural, as well as security, both material and financial, and comfort.

Under the assumption that all of these items can be predicted by a single underlying factor, the model fits exceptionally well, even under the constraints of full invariance (CFI = .996, TLI = .996, RMSEA = .024, appendix 4.2.1). Lower social and cultural participation, material and financial security and comfort can all be predicted very accurately by scores on one underlying construct. The construct also seemed to show a strong likelihood of measurement validity given its high AVE value (.619). All imputations show similar factor loadings and model fit, and model fit from the exploratory data subset was consistent with the model fit using the confirmatory subset and the full sample. This was to be expected given the strong cluster structure evident in the exploratory data analysis.

When inspecting items in order of their loadings a pattern emerges of certain goods having a much stronger predicted relationship with the underlying construct relative to other items. For instance, higher scores of relative material deprivation are extremely strongly associated with the inability to afford to replace furnishings and major electrical goods, followed by the ability to save regularly, followed by the ability to afford all-weather shoes, home repairs, contents insurance, and holidays, and lastly by the ability to heat the home adequately. This may reflect a certain prioritisation exercise that families facing relative deprivation generally follow. Firstly, and very frequently, the loss of participation in the security of major electrical goods and furnishings is accepted. This may be considered a low-level loss of participation because of the number of alternatives available – launderettes, non-perishable foods, or other forms of

entertainment - or otherwise because these items represent large investments that must be saved up for first, so cannot be replaced without deliberate effort to do so. The ability to save is then most commonly affected, perhaps in favour of averting other lower-loading factors. This may indeed be the reason why they have lower loadings, because the ability to save becomes impossible due to their replacement. These other factors are typically cheaper than major electrical goods but also, in many cases, more essential. Cheaper goods may be easier to budget for, but in addition the stigma and shame that may be associated with not being able to provide the children of the home with shoes, with not being able to provide a meal for a guest once a month, and not being able to partake in the social and cultural act of taking holidays, may be much greater and therefore more desirable to avoid than having a non-functioning electrical good or damaged piece of furniture.

5.3.4 Material Deprivation (Commodities)

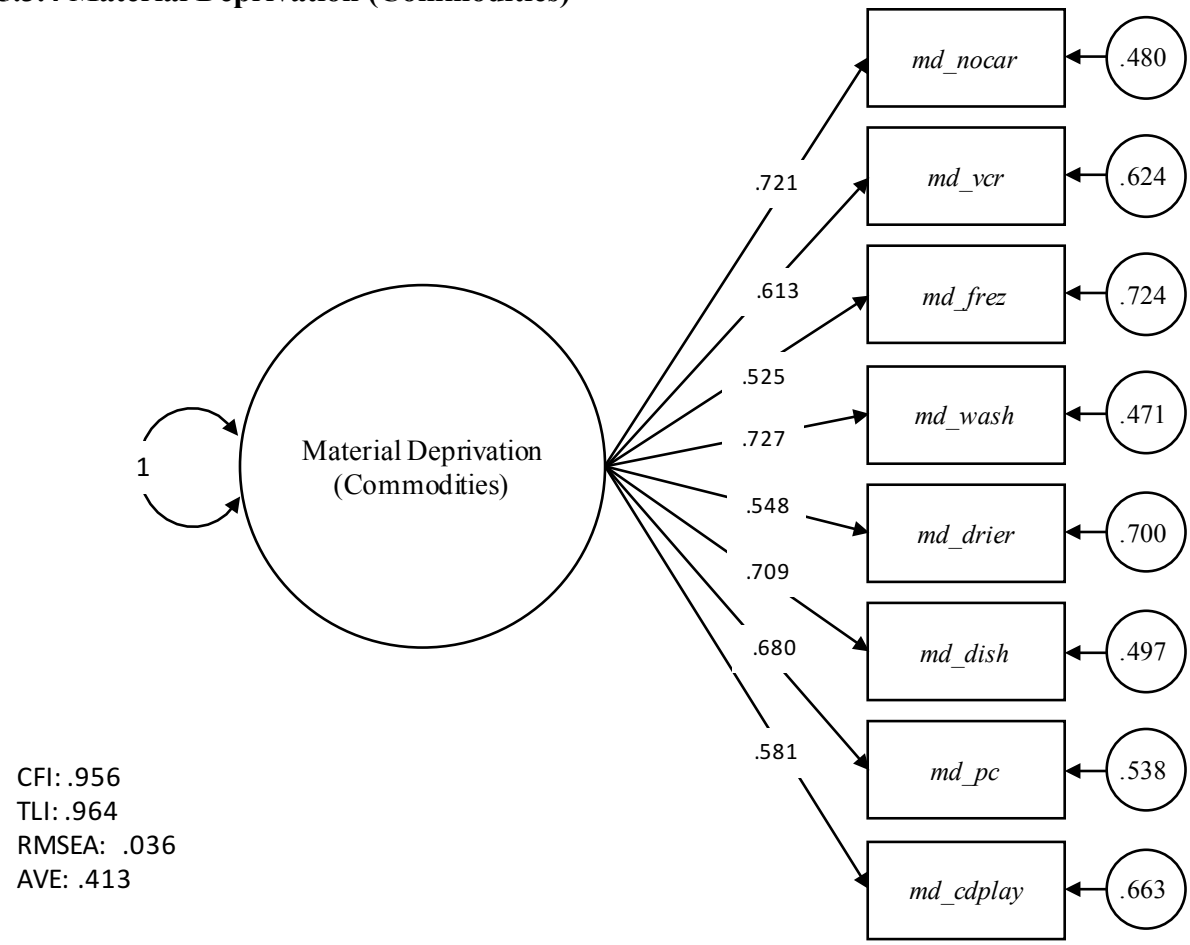


Figure 18: *Material Deprivation (Commodities) CFA*

Material deprivation of commodities, or material deprivation that could be considered more closely associated with absolute material deprivation within the specific context of the society, was initially constructed using the exploratory data subset with the inclusion of two additional

items: whether the household had satellite television and whether the household owned a microwave. These items related poorly with the others, and this was reflected in relatively poor fit statistics (CFI = .936, TLI = .918, RMSEA = .045). After removing these items and testing the new solution on a confirmatory subset of the data the model fit improved to the point where it could be considered good (CFI = .951, TLI = .932, RMSEA = .049).

There may be an empirical justification for the removal of these items. Satellite television no longer has the same societal value or desirability as it once had, with internet-based digital streaming platforms gaining popularity delivering arguably better customer experience and cheaper pricing (Sweney, 2018). Similarly, there have been complex changes in the desirability and utility of microwave ovens driven largely by food preferences (Cade, et al., 1999; Ferdman, 2014; Remnant & Adams, 2015; Foster & Lunn, 2007). This cultural change may reflect that certain goods are no longer strong indicators of material deprivation, and reflect differences in individual choices instead.

Of the items that did have salient loadings, it is possible that some of these may lose their value as indicators of material deprivation, but as of these findings they have measurement invariance for the six waves of the *Understanding Society* survey. Material deprivation of commodities is characterised by households lacking any combinations of a washing machine, tumble drier, dishwasher, car, personal computer, video and/or audio playback system, and personal computer. Households lacking any one particular item from this list are not necessarily in poverty, and an arbitrary count of how many they lack may not adequately capture whether they are missing because of poverty or because of personal preference. However, this type of factoring method scales the probabilities of these items not being owned based on their commonalities. Although this is still not perfectly capturing material deprivation it does a better job than weighting all of these items equally. For example, the reasons for someone not owning a freezer are probably very different for a person who does not own a dishwasher. The first may be forced to forgo a freezer for financial reasons, the second may not have a large or permanent enough living space, or different lifestyle preferences.

Using all items, including the ownership of modern essentials such as washing machines, this draws the factor slightly more towards representing poverty-induced material deprivation and away from lifestyle based factors. Its limited success is reflected in the later higher-order models (6.3.5 and 6.5). Largely the difficulty in adequately capturing this kind of commodity

deprivation is due to the lack of data collected on the reasonings behind peoples' decisions or situations that result in them not owning these items. Whereas the items for the previous factor have a specific response category for people who 'would like the item, but cannot afford it' the items in this factor do not have such a specific response: the object is recorded as a binary 'owns' or 'does not own'. An expansion of this category to include response categories for different reasons for non-ownership would likely increase the number of research questions that can be answered using these variables. This lack of distinction is likely what resulted in the final factor having a moderate-to-weak AVE statistic, despite good fit statistics. This factor also provides a good case study example of how AVE can add additional information about the interpretation of the quality of a factor, as the model fit can only tell a part of the story about how good an underlying construct is.

5.3.5 Financial Pressure (Second Order Factor)

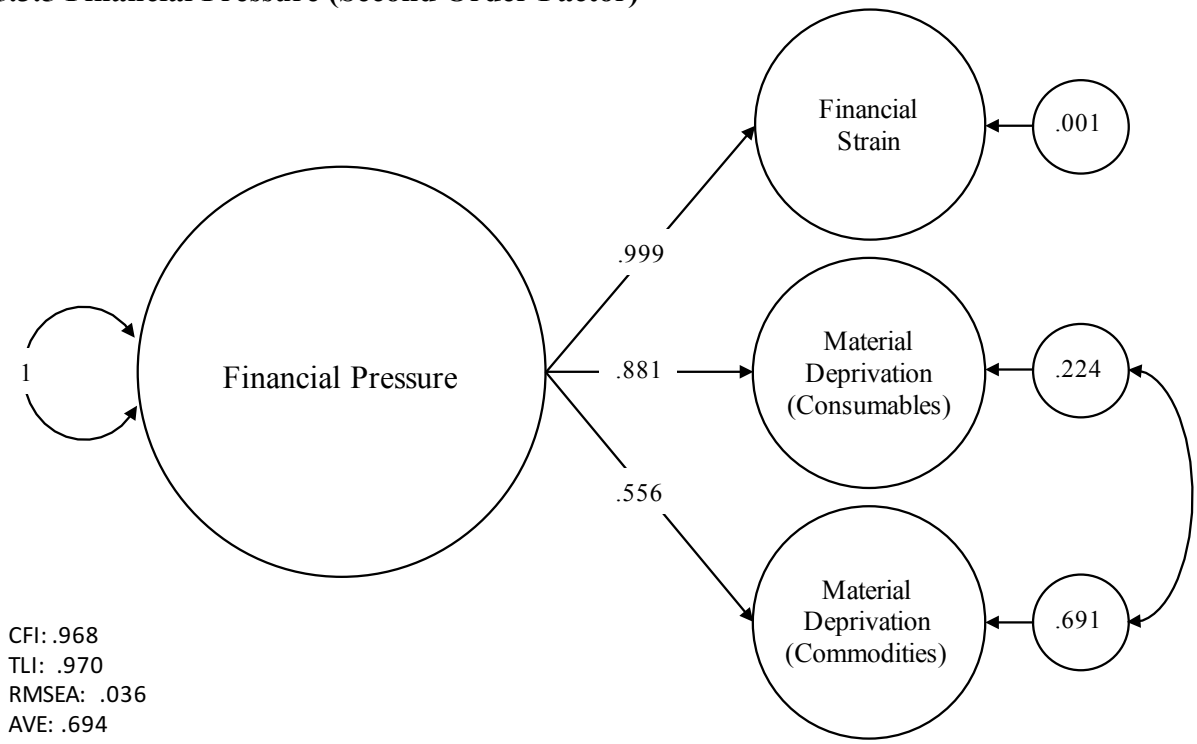


Figure 19: *Financial Pressure (Second order) CFA*

The three preceding factors were then loaded onto a higher-order factor for financial pressure, mirroring the approach taken by Tomlinson, et al., (2008). This second-order factor had good model fit in the exploratory subset only after adding a path to estimate the correlation between the two material deprivation factors. The reason for poor fit is likely due to the lower salience of the material deprivation (commodities) factor and limited number of items. The fact that this second-order factor only had two high salience manifest variables and required an estimation

of the correlation between errors later resulted in the third-order factor being unstable, especially after the removal of low-salience factors from the first rendition of the factor. This is explained in more detail later. Although it was not used in the final multidimensional poverty index, the factor is reported here for those who may want details of the invariance, reliability, and validity of a reduced but multidimensional material and financial pressure measure – such as for those wishing to model only the material and financial dimensions of poverty. The structure of the model was replicated with good fit statistics in the confirmatory subset of the data and subsequently tested for measurement invariance with the full sample using all imputations. The financial pressure factor showed good model fit and a strong AVE, as well as full measurement invariance over time (appendix 4.12.1).

These factors together show a far more nuanced approach to measuring financial security, strain, and material deprivation. They also show that financial strain and material deprivation have an integral connection. There can, on occasion, be an emphasis in policy direction on focusing on material deprivation that neglects this financial dimension – for example, by stating that the problem with those in poverty is that they lack the ability to budget and save - but this ignores that fact that financial strain and material deprivation are so closely linked that they can be approximated very well using the same underlying factor. Why then are policy interventions designed to reduce financial strain and material deprivation not treated with equal merit? It may be that the false narrative that people in poverty cannot be trusted to spend money wisely is all-pervasive – that the 'underserving poor' aren't responsible in their use of resources and must be instructed on acceptable types of spending (Alcock, 2006) - despite the apparent inextricable relationship shown here. People who have their financial strain reduced also have associated improvements in their material wellbeing and security and are subsequently able to participate in the valued roles expected of them in society. Increasing income results in better material conditions for the great majority of households.

5.3.6 Social Isolation (Relationships)

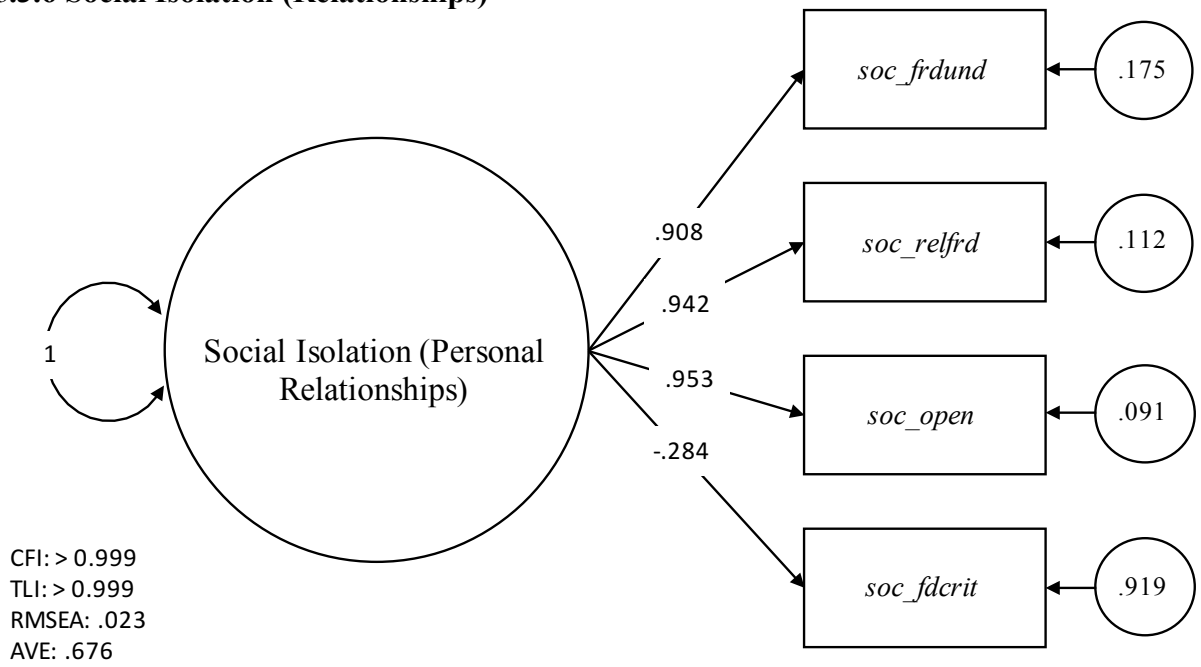


Figure 20: *Social Isolation (Personal Relationships) CFA*

The fit statistics for the social isolation (quality of personal relationships dimension) factor were extremely good (CFI > .999, TLI > .999, RMSEA = .023), suggesting that a single factor is able to replicate the covariance matrix almost perfectly. This is likely due to the very high loadings on each measurement item and the fact that the factor used a logit link. This means that the model predicted exact response categories for individuals as opposed to a value on a continuous scale that is not actually a valid response, reducing the amount of model prediction error. The factor showed measurement invariance over time and a strong AVE value, which meant it could be included as a dimension in a higher order factor.

The extent to which respondents felt like their friends were understanding of them (*soc_frdund*), whether they felt like they could rely on their friends (*soc_relfrd*), and whether they could be open around their friends were all strongly correlated (*soc_open*) and this is reflected in their factor loadings. An interesting but tangential finding was related to whether a respondent reported that their friends did not criticise them regularly (*soc_fdcrt*). This was negatively associated with the latent variable solution. Based on the coding of the item, this indicates that people with more open, more reliable, and more understanding personal relationships also have relationships where their friends are more likely to criticise them. It appears that a dynamic of openness and trust between friends fosters a greater capacity to air grievances directly.

5.3.7 Civic Participation

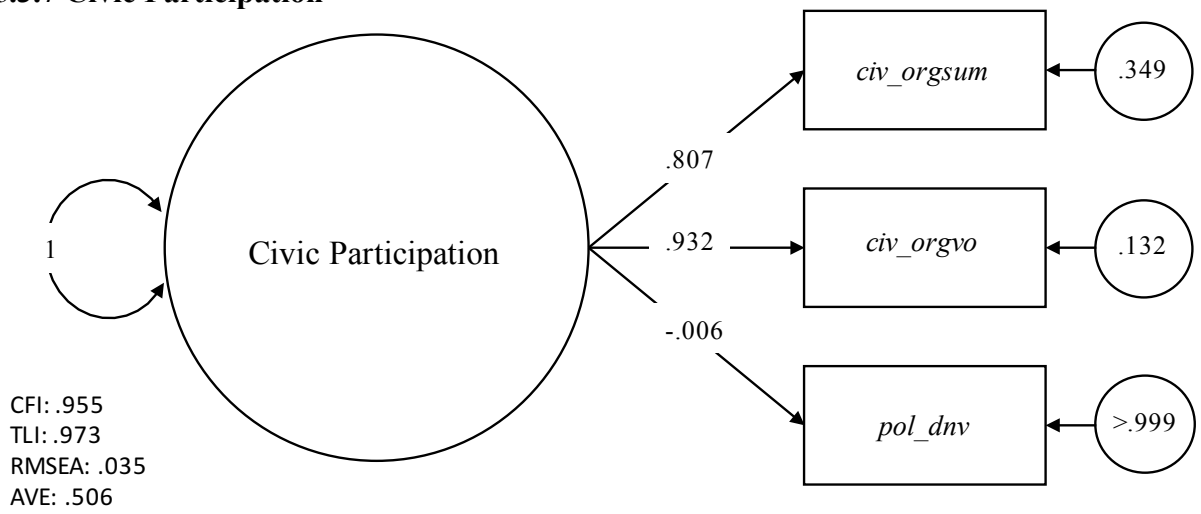


Figure 21: *Civic Participation CFA*

Civic participation was a factor that was virtually impossible to model. Despite the final fit statistics being good and the factor being time invariant (CFI = .955, TLI = .973, RMSEA = .035), one of the three indicator variables had a very low standardised factor loading of -.006, which was not statistically significant even with the sample size of over 290,000. By convention, variables with factor loadings of less than 0.3 should typically be removed and the model reassessed, as it is clear that they are not associated with the underlying construct (Byrne, 2009). Removing this manifest variable would have resulted in a factor that was based on only two variables, which is unlikely to contain enough information to model any underlying construct. For this reason the proposed factor was discarded and not used in any further modelling. This exercise demonstrates that it is important to inspect factor loadings and the basic requirements of models in factor analysis. A moderate AVE, full invariance, and very good fit statistics, does not necessarily make it a good factor (Kenny, 2014).

5.3.8 Life Satisfaction

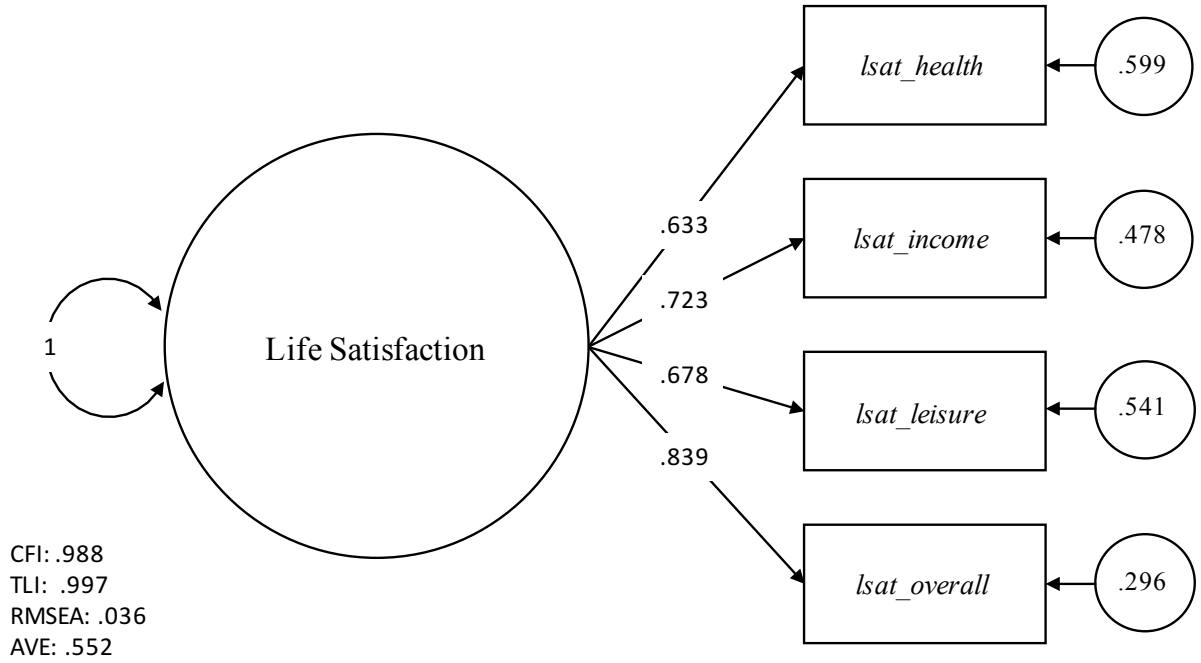


Figure 22: *Life Satisfaction CFA*

The life satisfaction factor had exceptionally good fit statistics (CFI = .988, TLI = .997, RMSEA = .036), had full measurement invariance over all six waves (appendix 4.9.1), with almost identical factor loadings on all imputations. Furthermore, life satisfaction had a moderate-to-strong AVE of .552. The highest loading factor was overall life satisfaction, followed by satisfaction with income, satisfaction with leisure time, and lastly, satisfaction with health. This indicates that people have somewhat converging priorities when reflecting on their life satisfaction; in other words, people tend to value leisure, income, and health collectively, rather than certain people prioritising specific outcomes. If satisfaction with income, health, and leisure were not related – if different people prioritised different things - their factor loadings would be far lower.

5.3.9 Psychosocial Strain – Anxiety/Depression

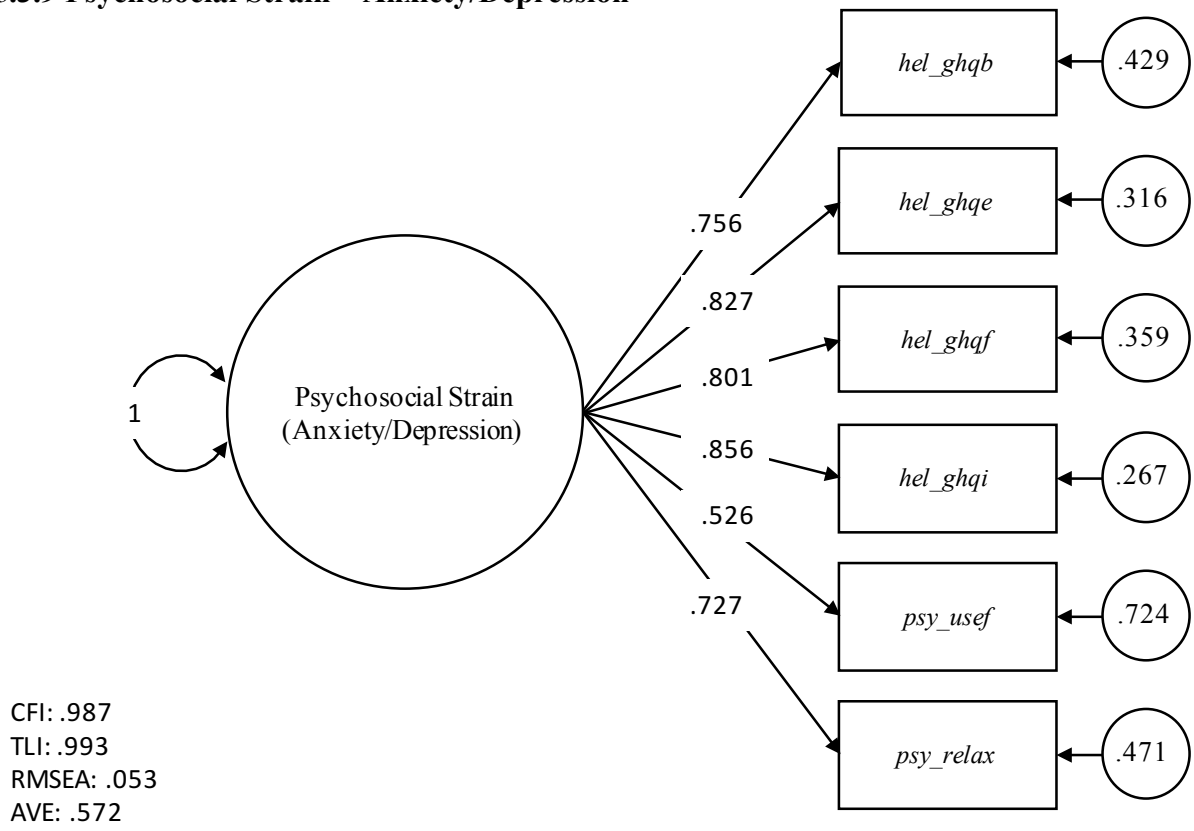


Figure 23: *Psychosocial Stain (Anxiety/Depression) CFA*

The first lower order factor for psychosocial strain, using questions relating to anxiety and depression as proposed by Shevlin & Adamson (2005), had very high model fit statistics (CFI = .987, TLI = .993, RMSEA = .053). The three-factor solution for psychosocial dimensions was explored first using a model containing all items and one factor and later using a three-factor solution. The three-factor solution resulted in better fit statistics and, due to its convention in the literature, was therefore retained. However, its validity is interrogated later in this section. Because this is primarily a psychometrics problem it is not the focus of this thesis.

The anxiety and depression factor showed full measurement invariance and near identical factor loadings on all five imputations of missing data. It also had a moderate-to-strong AVE statistic, implying a high amount of discriminant validity from other constructs with correlations smaller than .750. This is despite the inclusion of an additional two items from the WEMWBS scale which appeared to fit best in exploratory factor analysis. Although the AVE may have been slightly higher with these items removed, it is unlikely to change any substantive interpretation. In the context of multidimensional poverty, this factor may capture

some of the manifestations of continual attacks on a person's dignity and stigma (Lister, 2004; Walker, 2014), driven by a heightened sense of shame and diminished sense of self-esteem, as well as by more poverty-induced physiological drivers such as stress (Evans & English, 2003; Kasper, et al., 2008).

5.3.10 Psychosocial Strain – Social Dysfunction

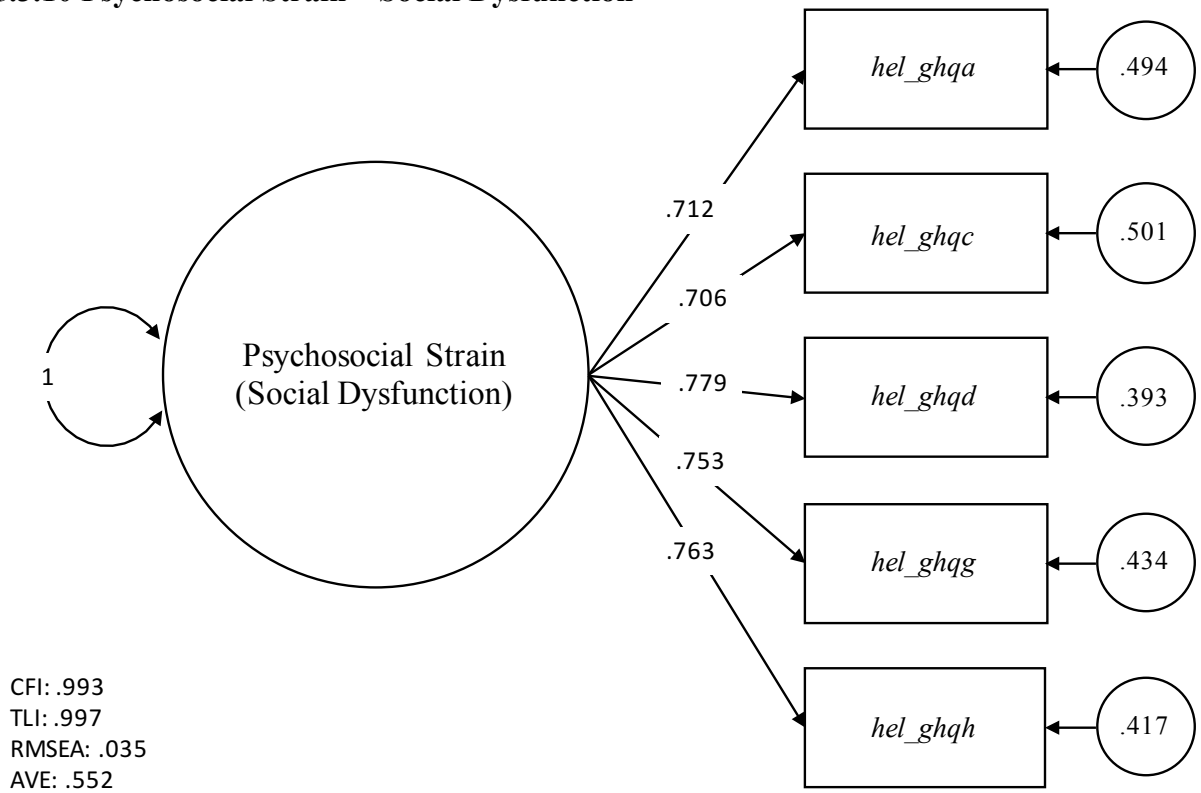


Figure 24: *Psychosocial Stain (Social Dysfunction) CFA*

The second dimension attempted to capture an underlying construct of psychosocial strain that has been classified in the literature as social dysfunction. This dimension is pertinent to poverty as it may capture some of the manifestations of social isolation not otherwise directly measured. Again, the model showed very high fit statistics (CFI = .993, TLI = .997, RMSEA = .035) and a moderate-to-strong AVE value. All items loaded onto the factor approximately equally. The factor also had full measurement invariance over all six waves and nearly identical factor loadings on all five imputations. Social dysfunction in this regard may represent a psychological manifestation of the problems associated with not being able to participate in the social and cultural dimensions expected by membership of society.

5.3.11 Psychosocial Strain – Control

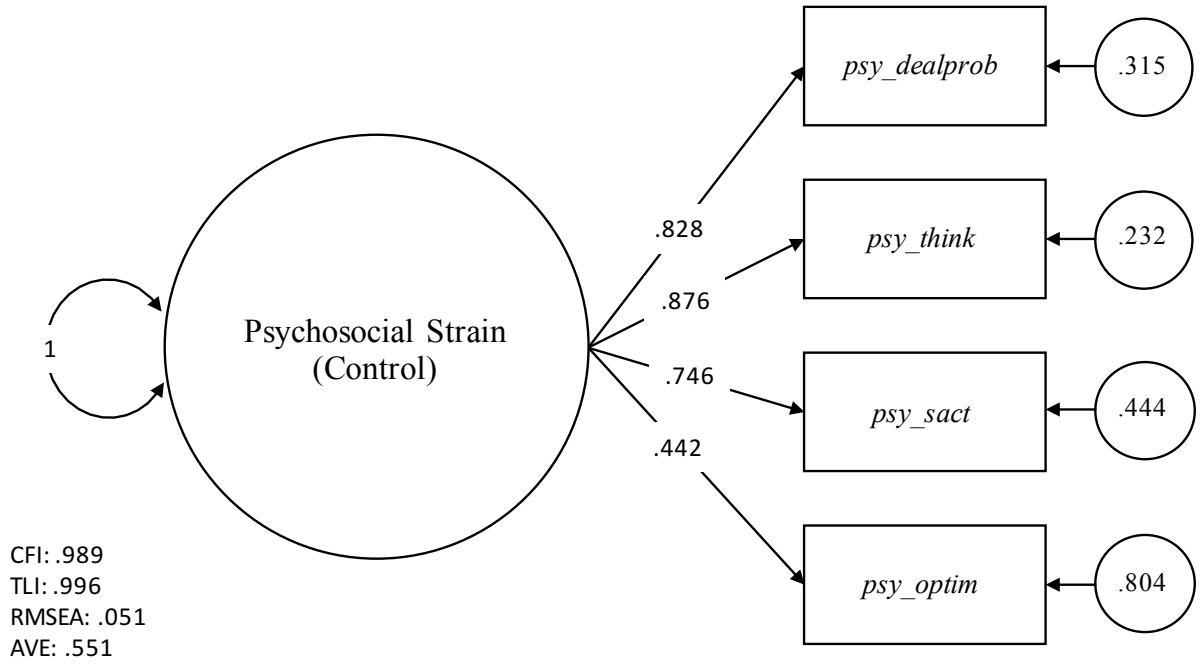


Figure 25: *Psychosocial Stain (Control) CFA*

Finally, a construct for an underlying psychosocial strain factor representing control was created using items from the WEMWBS. This included questions about whether the participant had been dealing with problems well, thinking clearly, was able to make up their own mind about things, and whether they were optimistic about the future. This latent variable had very good model fit statistics (CFI = .989, TLI = .996, RMSEA = .051), full measurement invariance over six waves, and approximately equal factor loadings across all imputations. The AVE was also moderate-to-strong.

The item with the lowest loading on this factor was in reference to participants' optimism about the future. Although in general those who are feeling confident about their ability to think clearly, deal with problems well, and make decisions, are not necessarily also feeling positive about future prospects. This may be because the other questions are strictly introspective, with the expectation being that the respondent reflects solely on themselves, whereas the question about future optimism is worded in a more general way. It may therefore be answered with a person's point of reference being more focused on external aspects in life which may be relational or societal. It has been retained as it is still a salient item, although given a choice of a greater variety of questions could potentially be removed in future research.

5.3.12 Psychosocial Strain (Second Order Factor)

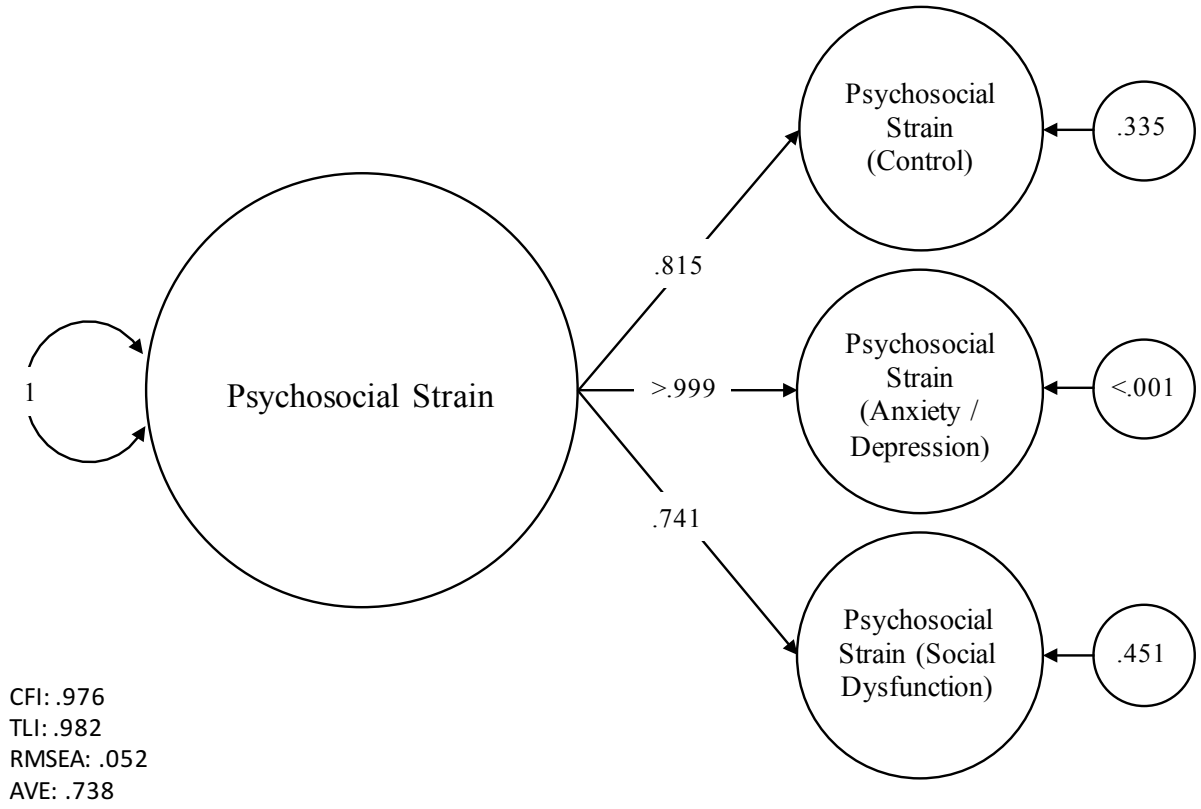


Figure 26: *Psychosocial Strain (Second order) CFA*

Research by Gnambs & Staufenbiel (2018) on the factorial structure of GHQ-12 questions suggests that a single psychological distress factor provides an equally good and more parsimonious factor solution than the three separate dimensions used by Shevlin & Adamson (2005). This research adopts the same strategy of Tomlinson, et al., (2008) by constructing a second-order factor for general psychosocial strain using each of the first-order psychosocial strain dimensions. Support for treating these factors that do not sufficiently capture distinguishable dimensions of psychosocial strain as opposed to just general psychosocial strain is raised in the following discussion of Fornell-Larcker criterion tests. Combining these dimensions within a general psychosocial strain factor results in a latent variable model with very good fit statistics (CFI = .976, TLI = .982, RMSEA = .052) that shows full measurement invariance across all six waves of *Understanding Society* (appendix 4.11.1). Factor loadings and fit statistics were also near-identical for all five missing data imputations. Finally, the AVE value is strong, suggesting that this factor is likely to be discriminantly valid from similar scales if they do not have a correlation of over around 0.85.

The factor loadings seem to suggest that the underlying construct is most strongly associated with the anxiety and depression factor, followed by the control factor, and lastly by the social dysfunction factor. An interpretation of this may be that psychological strain, for whatever reason, is almost always associated with some onset of feelings related to anxiety and depression, but that these do not necessarily predispose a person to a perceived loss of control around their life – in other words, their ability to manage is often, but not always, impeded. Furthermore, their capacity to maintain normal or healthy social functioning is often, but far from always, likely to be affected by these underlying feelings of anxiety and depression. The strong association between these types of mental health difficulties illustrates that research should be cautious about treating them as independent from one another in analyses. In general though, this seems to represent what would be expected from comorbidity research in mental wellbeing (Hirshfield, 2001).

5.3.13 Environment – Local Services

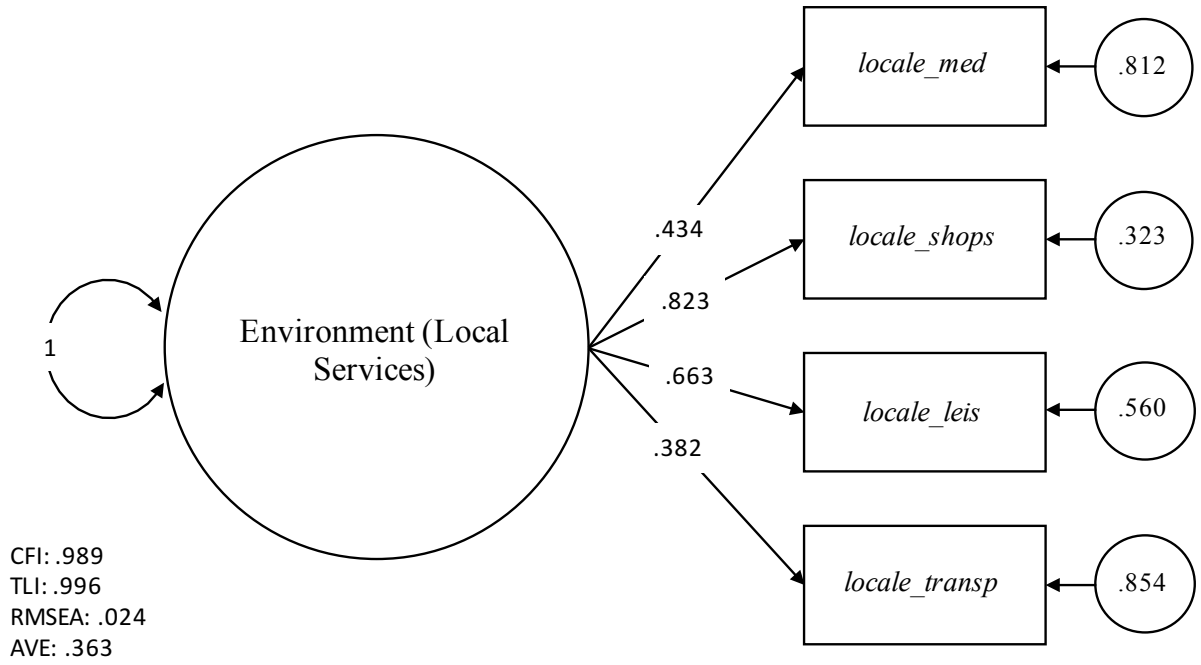


Figure 27: *Environment (Local Services) CFA*

Lastly, a factor was created that reflected respondents' feelings towards the quality of their local services. This was restricted to medical, shopping, leisure, and transport services due to their universal users. This created a variable with very good model fit statistics (CFI = .989, TLI = .996, RMSEA = .024) but a weak AVE value, due to poor loadings of two of the items used in the confirmatory factor analysis. The construct showed measurement invariance over the six waves of *Understanding Society*, and loadings did not show more than .006 deviation

between imputations. Worse shopping facilities and leisure facilities were likely to be strongly associated with an underlying local services construct, but worse medical services and transport services were only moderately associated. Using the exploratory subset of data an additional item for perceived crime levels, as a proxy for quality of police services, was included but did not have a salient loading (0.013, $p > .05$), so was subsequently removed. As previously mentioned in the exploratory factor analysis, this variable did not appear to have many cross-interactions with other clusters of items. This was confirmed using the exploratory subset of data in a higher-order multidimensional poverty model where it was shown that the variable had a non-salient loading and resulted in the overall model having worse model fit.

5.4. Fornell-Larcker Validity of Each Factor

Before presenting the third-order factor of multidimensional poverty an omnibus of Fornell-Larcker criterion tests (Fornell & Larcker, 1981) was completed using a model that correlated all of the first-order dimensions that were included in the final model, once it had been validated on a confirmatory subset of data after initial construction on an exploratory subset. The mean correlations across all imputations of the data between each of the latent variables are presented alongside each variable's factor AVE, the correlation squared, and the consequences for this when applied to the Fornell-Larcker criterion. To remind the reader, the Fornell-Larcker criteria for discriminant validity between factors is that a factor's Average Variance Extracted should exceed the squared correlation of other similar factors, otherwise the similar factors cannot be said to have discriminant validity. The purpose of the introduction of this type of testing into measures of poverty is to broaden the discussion and tools available beyond face and content validity, which focus only on subjective expert and layperson assessments of how well a measurement reflects its real-world counterpart, to justify dimensions as unique (Carmines & Zeller, 1979; DeVon, et al., 2007).

Many of the factors developed here do have discriminant validity from a majority of other factors, but there are a few dimensions that are evidently capturing significantly similar concepts. Or, in other words, the instruments used to measure them do not sufficiently isolate them from other, similar constructs. There are two responses to such cases. The first is to conclude that the instruments used to model a factor scale to measure the underlying construct do not have sufficiently high commonalities and the measure is therefore too vague or poorly defined by the manifest variables. This is a problem with the research design and survey items. The second conclusion one might make regarding the lack of discriminant validity is that the

theoretical basis for two distinct factors instead of one more general factor is unfounded. The results of the Fornell-Larcker criteria comparison do not tell us which of or what combination of these two conclusions should be made.

Many of the factors reported here could be defined as 'unintentionally emergent factors'. These are factors that have been constructed based on observations of clusters of individual variables available in survey data that have been grouped together firstly dependent on whether they correspond with theoretically defined dimensions, e.g. in conceptual illustrations of multidimensional poverty by Lister (2004), Spicker (2009), Baulch (1996), or in the theories of authors like Townsend (1979) or Sen (1999). After being filtered through this qualitative lens, the variables were then explored using the hierarchically clustered correlation plot to identify whether there is a strong basis to support these conceptual groupings. Many of these items were not designed with the explicit purpose to be used in factor analysis, in the same way that psychometric scales such as the GHQ-12 or the WEMWBS have been. These include the items found in the financial strain, material deprivation, and life satisfaction dimensions. Any lack of discriminant validity in these factors is therefore probably more likely to be derived from an inadequate measurement design that prevents a precise, highly valid, factor from being constructed. These tests are rarely conducted in sociological studies, where factors are not typically checked for measurment invariance, reliability, or discriminant validity, but these tools are essential in establishing a consensus around measures. As has been shown, many of these emergent factors do actually have reasonable high AVE values.

In contrast, factors derived from commonly used psychometric measurements are more likely to have discriminant validity because this technique is more frequently used in psychological studies. It is not within the scope of this thesis to explore the validity of competing psychological measurements, but existing reviews have highlighted that even within this pioneering field there is a large amount of underreporting of construct validity (DeVon, et al., 2007).

Table 6: Fornell-Larcker Criterion Validity Testing for Lower-Order Dimensions of Poverty

Model fit statistics	**Mean Fit**
CFI:	0.968
TLI:	0.971
RMSEA:	0.032
WRMR:	26.671

Correlation estimates		**Mean r**	**R-squared**	**Factor**	**Factor AVE**	**Valid?**
FINSTR	MATDEP1 (Cons)	0.954	0.910	FINSTR	0.597	No
FINSTR	MATDEP2 (Com)	0.672	0.452	FINSTR	0.597	Yes
FINSTR	PSYCN	0.335	0.112	FINSTR	0.597	Yes
FINSTR	PSYSD	0.252	0.064	FINSTR	0.597	Yes
FINSTR	PSYAD	0.379	0.144	FINSTR	0.597	Yes
FINSTR	LSAT	0.617	0.381	FINSTR	0.597	Yes
MATDEP1 (Cons)	FINSTR	0.954	0.910	MATDEP1	0.619	No
MATDEP1 (Cons)	MATDEP2 (Com)	0.431	0.186	MATDEP1	0.619	Yes
MATDEP1 (Cons)	PSYCN	0.258	0.067	MATDEP1	0.619	Yes
MATDEP1 (Cons)	PSYSD	0.190	0.036	MATDEP1	0.619	Yes
MATDEP1 (Cons)	PSYAD	0.297	0.088	MATDEP1	0.619	Yes
MATDEP1 (Cons)	LSAT	0.453	0.205	MATDEP1	0.619	Yes
MATDEP2 (Com)	FINSTR	0.672	0.452	MATDEP2	0.413	No
MATDEP2 (Com)	MATDEP1 (Cons)	0.431	0.186	MATDEP2	0.413	Yes
MATDEP2 (Com)	PSYCN	0.097	0.009	MATDEP2	0.413	Yes
MATDEP2 (Com)	PSYSD	0.095	0.009	MATDEP2	0.413	Yes
MATDEP2 (Com)	PSYAD	0.084	0.007	MATDEP2	0.413	Yes
MATDEP2 (Com)	LSAT	0.145	0.021	MATDEP2	0.413	Yes
PSYCN	FINSTR	0.335	0.112	PSYCN	0.551	Yes
PSYCN	MATDEP1 (Cons)	0.258	0.067	PSYCN	0.551	Yes
PSYCN	MATDEP2 (Com)	0.097	0.009	PSYCN	0.551	Yes
PSYCN	PSYSD	0.583	0.339	PSYCN	0.551	Yes
PSYCN	PSYAD	0.824	0.680	PSYCN	0.551	No
PSYCN	LSAT	0.539	0.291	PSYCN	0.551	Yes
PSYSD	FINSTR	0.252	0.064	PSYSD	0.552	Yes
PSYSD	MATDEP1 (Cons)	0.190	0.036	PSYSD	0.552	Yes
PSYSD	MATDEP2 (Com)	0.095	0.009	PSYSD	0.552	Yes
PSYSD	PSYCN	0.583	0.339	PSYSD	0.552	Yes
PSYSD	PSYAD	0.753	0.567	PSYSD	0.552	No
PSYSD	LSAT	0.459	0.211	PSYSD	0.552	Yes

PSYAD	FINSTR	0.379	0.144	PSYAD	0.572	Yes
PSYAD	MATDEP1 (Cons)	0.297	0.088	PSYAD	0.572	Yes
PSYAD	MATDEP2 (Com)	0.084	0.007	PSYAD	0.572	Yes
PSYAD	PSYCN	0.824	0.680	PSYAD	0.572	No
PSYAD	PSYSD	0.753	0.567	PSYAD	0.572	Yes
PSYAD	LSAT	0.613	0.376	PSYAD	0.572	Yes
LSAT	FINSTR	0.617	0.381	LSAT	0.522	Yes
LSAT	MATDEP1 (Cons)	0.453	0.205	LSAT	0.522	Yes
LSAT	MATDEP2 (Com)	0.145	0.021	LSAT	0.522	Yes
LSAT	PSYCN	0.539	0.291	LSAT	0.522	Yes
LSAT	PSYSD	0.459	0.211	LSAT	0.522	Yes
LSAT	PSYAD	0.613	0.376	LSAT	0.522	Yes

Under the Fornell-Larcker test criteria, the financial strain dimension shows discriminant validity from latent variables for material deprivation of commodities, all psychosocial strain dimensions, and life satisfaction. Financial strain's AVE did not however exceed the squared correlation between it and the material deprivation of consumable activities, goods, and services (r = .954). This implies that financial strain and relative material deprivation are, when measured as constructs, so closely related that we cannot claim they reflect different things in reality. In substantive theory, this may mean that the ability to keep up with regular bills for housing, council tax, and energy can be assumed within the wider ability to participate in the activities and attain the kinds of diets, comforts, and securities, expected within wider society, and vice versa. In other words, material deprivation and financial strain are so fundamentally linked that there is not much justification for treating them as distinct constructs: where one happens, the other follows. The reason for such a high correlation might be related to Townsend's (1979) theorised 'breakpoint' in participation. Townsend (1979; 1985) theorised that below a certain point in the income distribution levels of participation in society fell at a greater rate, and evidence for a 'participation' floor has been reported elsewhere (Ferragina, et al., 2013). As Townsend (1985: 662) explains:

"It is as if people strive to conform with what is expected of them when income shrinks (they economise in what they do but still undertake the same activities) but once it shrinks below a particular level they withdraw (or withdraw their children) from fulfilling certain social obligations or well-established customs or activities. They no longer meet friends, children are occasionally absent from school, heating is turned off, conventional diets are no longer regularly observed, visitors are not longer invited into the home, ill-health and disability become more common."

These findings may reflect this. This 'certain level' may be when financial strain becomes a significant factor in life, which, despite being technically more difficult to develop a continuous measure of, is likely to be more successful in predicting participation than using only income, given the vast number of caveats and adjustments that need to be taken for housing and living costs and household size.

This supports the idea of an irreducible material and income based core of poverty. Furthermore, we have established that this alternate measure of financial strain, as a core of poverty more complex than income, can be tested for its reliability, invariance, and validity, making it a methodologically more desirable metric than income. When a person's situation is at a point in which their finances are strained to the extent that they are unable to pay, or be confident in their ability to pay, their bills regularly, an increase in their relative deprivation – including their ability to participate in society – is almost inevitable.

Even when people define poverty as using income and payments for essentials, they must undoubtedly be capturing a great many people who also more broadly are lacking a material quality of life because of this strong relationship (Pantazis, et al., 2006). With this information, there should not be much difficulty making a technical argument for moving away from a blunt income and housing costs based measure of poverty, such as the HBAI, and towards a material deprivation based measure. This evidence helps overcome the political argument that can be made against such a move. Methodologically, this means that we cannot be confident that our financial strain dimension is truly a different dimension to relative material deprivation and, in theory, these two may be better conceived of as a single dimension.

The next result of interest concerns the test for discriminant validity between the two material deprivation dimensions. The earlier decision based on goodness of fit statistics to separate material deprivation items on a theoretical basis into two groups, one concerning consumable goods that need regular replenishment and another concerning household commodities that require far less frequent replacement, and are more socially transient and culturally specific, appears to be supported by the finding that these two dimensions met the Fornell-Larcker criteria. In fact, these two dimensions only showed a moderate correlation. This supports the

idea that these different measures of material deprivation are distinct concepts and warrant different measurements (Bradshaw & Finch, 2003). The correlation does however show that they are related. A truly holistic multidimensional measure of poverty should recognise that there is more than a single way to be materially deprived.

Psychosocial dimensions show discriminant validity with all material dimensions and with life satisfaction, however, are not always distinct from one another, supporting the decision to combine these into one higher-order factor, following the findings of Gnambs and Staufenbiel (2018). There is some evidence to challenge this conclusion. The psychosocial dimension designed to measure a sense of control is valid in that it captures a substantially different set of variance from social dysfunction, but not necessarily from anxiety and depression. One reason for this may be the lack of volition that often accompanies anxiety and depression, where the former might be characterised by paralytic fear of the consequences of any action and the latter by a loss of motivation, self-esteem, and willpower (Peveler, et al., 2002; House & Stark, 2002) or further comorbidities with limiting physical health (Niles, et al., 2015). Mixed evidence for discriminant validity exists between social dysfunction and anxiety and depression, where the Fornell-Larcker criterion is not met when using the Social Dysfunction factor's AVE but is met when using the Anxiety and Depression factor's. Measurement instruments relating to social dysfunction may therefore benefit from being re-evaluated and made more precise, if possible, potentially with the inclusion of additional indicators of social dysfunction to further distinguish it from more general anxiety and depression. Developing valid measures of these psychosocial dimensions that are sufficiently distinct from general anxiety and depression is likely to be difficult considering the high rates of comorbidity between anxiety, depression, and other psychological differences (Hirschfeld, 2001).

Life satisfaction was positioned with similar relationships to both psychosocial and material deprivation factors, with strong correlations with financial strain and anxiety and depression, but was also a valid factor that met the Fornell-Larcker criterion for discriminant validity when tested against all other factors. Although it could be argued that life satisfaction as a concept would be largely capturing the variance of either general psychosocial distress or levels of consumption/material attainment, in reality, it sits at the intersection of these two things, related to each but sufficiently distinct.

It is clear from this analysis that in order to claim that we have a strong and valid multidimensional measure of poverty more work needs to be done to develop valid scales of each financial and material dimension that emphasise the subtle differences in their ontological basis. It is not enough to be able to rely on these emergent factors that already exist throughout secondary surveys, although it has been established here that they can be reliable if needed, and are good intermediates, they cannot replace purpose designed instruments. More work needs to be done in both qualitative and quantitative research exploring whether, for instance, financial strain and material deprivation can really be said to be distinct concepts and, if so, additional efforts need to be made to explore how these are best operationalised to develop measures that can contribute to reliable and valid latent constructs.

There are clear applications for the use of this method beyond this thesis. Future research, for instance, might contribute to the discussion around social quality indicators (Van Der Maesen & Walker, 2012) and poverty by testing whether reliable latent variable constructs of the two concepts show discriminant validity. They are likely to be inversely related: as poverty increases, social quality decreases. However, if their strength of association is so strong that their indicators show no discriminant validity it would suggest that they are inverted measurements of the same underlying concept. If they did show discriminant validity, however, they could be treated as complimentary indicators that show us legitimately different dimensions of human experience and outcomes.

5.5 Multidimensional Poverty

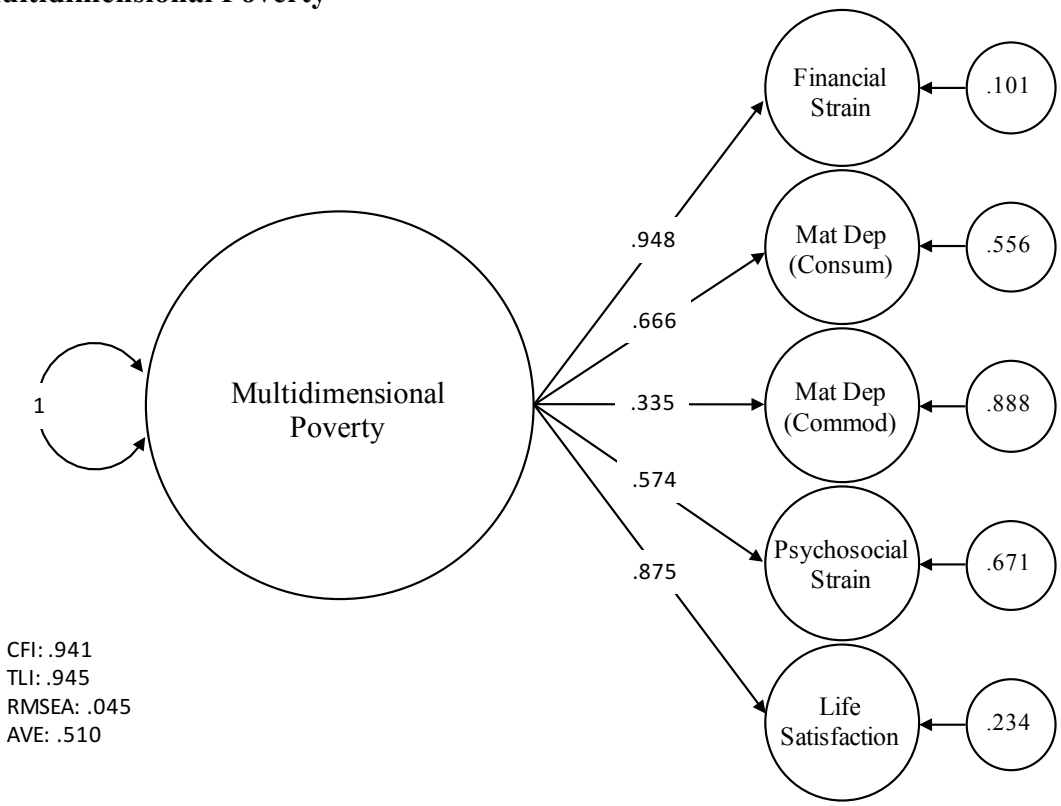


Figure 28: *Multidimensional poverty CFA*

Figure 28 shows the final third-order factor model for multidimensional poverty using four first-order factors and one second order factor. During exploratory factor analysis the measure included the civil participation and social isolation from personal relationships dimensions, as used by Tomlinson, et al. (2008). This resulted in a model that had marginally better fit statistics (CFI = .943, TLI = .947, RMSEA = 0.044) but civic participation and social isolation did not have salient factor loadings, both being smaller than ±0.15. This was not surprising considering the earlier correlation matrix visual inspection hinted at this. This indicates that, using this data, the two factors are not substantive dimensions in a broader construct that also captures life satisfaction, material deprivation, and psychosocial strain. This is contradictory to the findings of Tomlinson, et al. (2008), although these two factors had relatively low loadings in their study.

There may be two reasons for this. The first may be that the manifest variables available in *Understanding Society* that were chosen actually relate to a different underlying concept because of subtle differences in wording and coding. This is certainly the case for the social isolation variable, which focuses very closely on personal relationships rather than general societal isolation. As outlined in the exploratory data analysis section there is little substantive

evidence within this data to suggest that people in poverty are always likely to have much lower social capital. Alternatively, if the social isolation and civic participation variables were capturing the same underlying concept that was captured by Tomlinson, et al. (2008), it may be that the configural structure of poverty – the dimensions that make up multidimensional poverty - has shifted since the 1991-2003 time period covered by Tomlinson, et al., (2008). For example, a shift in political climate following the 2008 financial crises may have redefined the disengagement from civic life of those living in poverty. The continuation of adoption and expansion in ownership of new technologies and social media networks may also have democratised the maintenance of strong interpersonal relationships further. More in-depth research is required to understand exactly why these differ from similar analyses. As previously identified, more precise definitions of what is meant by social isolation and social exclusion, and different types of social isolation and exclusion, may be needed to improve survey methodology in the measurement of these concepts.

The second-order financial pressure factor was disaggregated back to three first-order factors. The inclusion of two second-order factors (psychosocial strain and financial pressure) resulted in instability in the model estimation which resulted in models not converging. It was not possible to isolate why this was the case. As an applied user engaging in model structures that push some of the boundaries of the existing software, namely using third-order factors based on first and second order factors that use a mixture of categorical, ordinal, and continuous variables, and complex longitudinal survey data, this was always a risk inherent in the research and acknowledged in the methodology. The final model using four first-order factors and one second-order factor to construct a third-order multidimensional poverty factor represents the limits of my own applied statistical ability at the time of data analysis. This solution converged for four of the five data imputations, and it was not possible to establish why it would not converge for the fifth. Nonetheless, I believe this to be a robust example of the feasibility and quality of a multidimensional poverty measure using latent variable analysis and as software, computing power, and diagnostic tools improve in the future it should become more and more viable.

The final model fit for the multidimensional poverty index was satisfactory to good, but fell short of criteria for very good, by commonly accepted standards (CFI = .941, TLI = .945, RMSEA = .045; Kenny, 2014; Bollen & Long, 1993). Furthermore, factor loadings were approximately equal regardless of missing data imputation (appendix 4.13.1). Model fit

improved after constraints for wave invariance were added, suggesting the parsimony adjustment after a reduction in estimated parameters far outweighs the additional precision of wave-specific estimates. In other words, treating each year as having its own specific construct of multidimensional poverty is not necessary: the underlying items, the scale, and the mean structure of the multidimensional poverty factor do not change substantially between years.

The consequences of undertaking this invariance testing appear to be that, when constructs are very complicated, the addition of invariance constraints can substantively change the conclusions made by researchers. When constraints for measurement invariance are not added the risk of something akin to a type I error - rejecting an accurate hypothesised structure in favour of a less accurate structure, or rejecting the construct entirely - increases. This is in addition to the more general problems with using latent variables in models without first adjusting for invariance, in that they are non-comparable (Millsap, 2011). If sociologists wish to seriously develop reliable and valid multidimensional measures, this kind of testing must become more commonplace.

Substantively this factor analysis shows that it is possible to create a measure that, when compared with theory and existing qualitative literature, reflects an underlying latent construct of poverty that fits within the conceptualisation of poverty as multidimensional, combining material and psychosocial/relational factors (Lister, 2004; Walker, 2014), and defined by levels of participation (Townsend, 1979). Specifically, several underlying factors that are found within experiential and qualitative accounts of poverty – financial strain, material deprivation, life satisfaction, and psychosocial strain - can be predicted well using a single latent variable. This validates on a quantitative, macro scale, the interconnectedness of financial/material deprivation and psychosocial wellbeing commonly reported in lived experiences of poverty (Lister, 2004; Walker, 2014).

Financial strain is the highest loading factor for the multidimensional poverty construct. The questions these relate to include a variation of objective and subjective assessments of household financial circumstances, rather than strict income criteria such as those found from the HBAI or $2-a-day criteria (Anand, et al., 2010). This reflects the material deprivation or unacceptable hardship hub or core dimension of poverty which features in theoretical frameworks of poverty (Spicker, 2007; Lister, 2004). Lack of security characterises this dimension, defined by Ferge (1992, in Spicker, et al., 2007: 177) as a combination of security

of employment, security of income, and security of accommodation. However, the construct is more weighted towards security of income and security of accommodation as basic rights than income. While security of income and security of accommodation are universal basic needs, security of employment ignores valuable labour outside of employment contracts.

While security of employment is a legitimate need for those in employment (Standing, 2014) it lacks universal applicability to those outside of employment, people who have retired from work or have longstanding illnesses, or, indeed, children and those fulfilling essential caring roles. Although formal employment in the public or private sector is seen, especially in political rhetoric, as a universal condition or, at least as a universal aspiration, and largely neglects the emotional, caring, and domestic labour undertaken by many who are deemed 'economically inactive': 21 per cent of the working age population in 2018 (Office for National Statistics, 2019: online). If employment is emphasized as central to defining poverty a real problem is created which reinforces the false dichotomy of the deserving and undeserving poor (Shildrick, 2018). Any quantitative measure should be cautious to not treat unemployment as a prerequisite that needs to be met to qualify for being classified as experiencing poverty, and such a view is supported by recent work by the Joseph Rowntree Foundation (2018) that a large number of working households can be defined as living in poverty. Security of income and security of accommodation can be justified as a dimension of universal basic needs, security of employment cannot.

Life satisfaction also takes a central role in the factor analysis. Indeed, life satisfaction could be considered a cross-cutting measure of poverty that is applicable to any society, if statistically invariant. The dimension itself may be reasonably contentious, as life satisfaction is conceptualised as a higher-order, rather than basic need based on Maslow's (1943) original ranking of physiological and psychological human needs. However, more recent research into the applicability and reliability of Maslow's conceptualisation of needs has found no support for the proposed ranking of each need (Tay & Diener, 2011). In other words, the meeting of certain 'higher-level' needs is not contingent on meeting more basic needs.

This contradicts some of the assumptions found in narratives about poverty, and supports literature around the role of income and happiness. As early as the 1840s there was a popular narrative that what the poor lost in terms of wealth was made up for in terms of physical health and 'vigor' (Brisbane, 2012 [1840]). More recently, the stigmatising narrative around the poor

is focused on ‘laziness’, the idea of ‘skivers’ versus ‘strivers’, with the skivers often painted as people who have minimal income from state benefits but are flush with free time to occupy with leisure (Shildrick, 2018; Shildrick & MacDonald, 2013). Although this may be a comforting delusion for people frustrated with what they perceive as injustices, it is demonstrably nonsense. As a general rule, if someone is less satisfied with their income, they are less satisfied with their leisure time, and less satisfied with their health, and all of these factors are associated with financial and psychosocial strain. The poor are not living ‘the easy life’; there is no ‘happiness upside’ to being in poverty. This factor captures a sense of underlying life-satisfaction, which could be considered analogously to Maslow’s (1943) higher order dimension of need titled ‘self-actualisation’. A sense of satisfaction and contentment with life may reflect the meeting of an innate need to be realising ones’ potential, however, is still only an approximation.

This raises a question. Why should satisfaction with ones life and dignity not be seen as a valid measure of a fundamental, universal, basic need? Why is contentment with life seen as a privelege to be earned by a few and not a right? If we argue that how satisfied someone is with their life has no place in our assessment of what constitutes poverty then, even in a multidimensional framework, we are reducing the needs of human life to subsistence, consumption, economic productivity, and freedom from ill-health. Although all of these are clearly important, if a person is not happy or at least content in their existence, can it really be said that they are liberated from deprivation? Surely they are still deprived of something? At the very least, the hierarchy of needs and Baulch’s pyramid need inverting – some acknowlegement of dignity and self-regard should be central to measuring poverty and need, not a ‘wooly’ extension to hard financial and material measures. It comes up again and again in peoples’ recollections of lived experiences, while material and financial aspects can often be downplayed (Walker, 2014; Shildrick & MacDonald, 2013). Things that are very complex to operationalise are not necessarily very difficult to experience, describe and feel tacitly. This model suggests that poverty is equally as dehumanising as it is fiscally straining, and the need to experience humanity in all its complexity is at the core of these dimensions (Veit-Wilson, 1999; Lister, 2004; Shildrick & MacDonald, 2013).

Relative deprivation of goods and services has been highlighted in this research as an approximation of consumption based participation in society. This factor is strongly associated with the latent multidimensional poverty construct, which suggests that participation and the

ability to play social roles often falls as financial strain increases and life satisfaction decreases. The fact that this dimension is not as strongly associated with the underlying construct is probably accounted for by the efforts that people make to maintain such activities through economising them, where possible, as financial strain increases (Townsend, 1985; Ferragina, et al., 2013). Such a dimension can be seen as a manifestation of the point at which unacceptable hardship begins to interrupt the attainment of the conditions of life required for material and social-relational participation, and economisation of participation starts to increase.

The other side of these manifestations of financial and life satisfaction strain can be manifestations of poor psychosocial health. This distinction, supported by discriminant validity, proposes that there are material manifestations of need which can be measured by participation and psychological dimensions of need, which can be measured using psychometric tools. Where these two types of needs meet, the social-relational impact, might be approximated externally by participation in activities that foster relationships, and internally by the internalisation of the impact of such activities, and how they form an inner view of the self in relation to others. External participation can be seen in the observable experiences of participation; internal participation might be seen in our feelings of self-worth relative to others in society. Both are related, but distinct. The factor loadings suggest that a person is likely to experience psychological distress due to a loss of participation in society, or vice versa.

More broadly this psychological distress acts as an operationalised indicator of the relationship between poverty and mental health (World Health Organisation, 2014), as well as a potential indicator of repeated assaults on dignity, which may take the form of stigmatised narratives and dehumanisation. The strong factor loading supports the earlier hypothesis that psychological wellbeing is a significant dimension of poverty in itself, but only when the conceptualisation of poverty is broad enough to include it (Lister, 2004: 36; Baulch, 1996). Using a higher order factor model these specific subdimensions can be included with clarity on how they relate to one another in a way that avoids '[losing] sight of what is unique to the phenomenon': material deprivation and financial strain are included in the model so only psychosocial strain associated with these unique parts of the phenomenon is included in the overall measure (Lister 2004: 36).

Finally, the lowest loading factor was the factor constructed from items relating to the ownership of certain goods and appliances. This reflected the extent to which someone was able to attain items associated with convenience designed to improve material quality of life. Their changing affordability and lack of cultural nuance in their desirability has been brought into question in earlier discussion and it is not clear whether their low factor loading reflects the inadequate construction of the questionnaire items or the fact that the ownership of such items legitimately has little association with poverty as imagined using a combination of these other factors. Nevertheless, the construct has enough salience to be included in the third-order factor.

5.6 Contributions to the development of theory

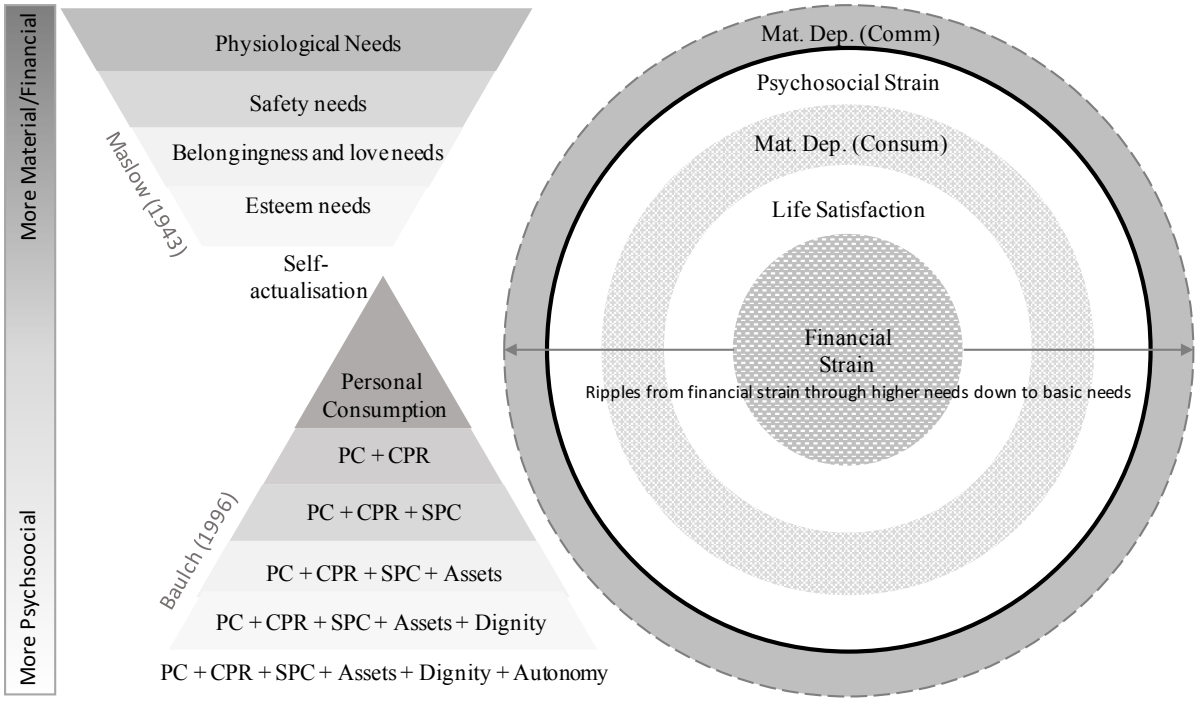


Figure 29: *A theory of multidimensional poverty that illustrates the dynamics between psychosocial needs, participation, and its material core. The right hand side shows the arrangement of items in the multidimensional poverty higher-order factor with their proximity to the centre indicated by how strongly weighted their factor loadings are. Dimensions that are more closely associated with material/financial deprivation are shaded in darker grey, dimensions that are more closely associated with psychosocial dimensions are shaded in lighter grey through to white. These dimensions show clear parallels to the taxonomies of Baulch (1996) and Maslow (1943). Financial strain operates as the central 'core' of poverty – unacceptable hardship. The effects of financial strain then ripple outwards in alternating patterns of psychosocial and material dimensions. Firstly affecting the universal need of life satisfaction, then the universal material need of consumption of activities, events, and goods required to participate in society. In more compound poverty, it is psychosocial strain that is affected next as a recognisable loss of self-actualisation, dignity, and esteem, which may later manifest in poor mental health. Lastly, deprivation of societally*

common material goods of convenience may become apparent. As a person's condition of financial strain persists over prolonged spells, this may bring them closer towards states of destitution these dimensions may be more likely to surface. However, these states at the edge of the ripples of poverty are more changeable and dependant on what cultural differences.

The findings suggest that at the centre of poverty you have the financial/material core that is present in the conceptualisations of Lister (2004) and Spicker (2009). However, although the detail and evidence for what encircles this material core is rich and detailed, the nature of the processes underlying how unacceptable hardship translates to symbolic-relational, psychosocial, and participation dimensions is less clear. This model develops the relationship using the findings from the factor analyses as a basis. The propositions themselves are theoretical extensions as possible explanations for the findings here, and will remain as such until they can be validated with in-depth qualitative research. Central to poverty is a material component, and a succinct way to capture this regardless of time is by observing whether a person is facing financial strain. Financial strain may not always be the highest loading factor in all cases, but it can be argued that without an aspect of financial strain the phenomenon that is observed is less likely to fit within the definition of poverty (Lister, 2004: 36), largely that there must be some material dimension associated with the measurement of the concept; it cannot be *purely* psychological or symbolic. Financial strain is central to the experience of poverty.

Financial strain may then cause a ripple outwards that first affects 'higher order' psychosocial and material needs. Maslow's (1943) hierarchy of needs is used as an example of this taxonomy of needs that can be distinguished as being higher or lower than others, and Baulch's (1996) pyramid is used to represent a hierarchy of material needs. The interpretation of Baulch's pyramid in this context is slightly harder, however, but equally we can imagine that higher-order material needs are associated more with the ideas surrounding relative poverty – the ability to attain the 'conditions of life… which allow [people] to play the roles, participate in the relationships, and follow the customary behaviour which is expected of them by virtue of their membership of society' (Townsend, 1979: 31). Lower order material needs are more similar to those found in definitions of absolute poverty. They are more concerned with subsistance, shelter, and the facilities and goods required to attain these. These base physiological needs share common ground with those found in Maslow's hierarchy.

What is illustrated here is the following: if poverty is agreed to represent unmet universal needs of the individual, the meeting of which is dictated by societally accepted 'satiaters' – things that sate needs (Doyal & Gough, 1984), and if these needs are agreed to be multidimensional (Baulch 1996, Lister 2004, Spicker 2009), and if those dimensions are agreed to be economic, relational, and psychosocial, it therefore follows that any quantitative operationalisation (Lister, 2004: 6) of poverty should represent a combination of conceptualisations of relative economic, social, and psychological needs. More importantly, poverty should therefore not only be imagined as the sum of these unmet needs, but also as the state in which they interact. The dimensions do not just exist independently, they contribute to the whole and occur together. Together they form an indicator of the extent that people are able to attain the multitude of different resources required to participate in the 'full experience of humanity' (Veit-Wilson, 1999: 85).

The implication that each dimension has a ripple effect is supported by the factor loadings and commonalities that suggest some ripples 'go out further' than others when the conditions of financial strain and low life satisfaction are met, and possibly as the span of time a person living in poverty increases, resulting in externalisation of indicators of poverty, from more complex universal sociorelational needs to more physiological and material needs needs, from participation and dignity to subsistance. This pattern is also mimicked in some of the first-order factors that make up the dimensions of poverty, for example, in the financial strain factor where council tax bills are more likely to go unpaid before energy bills. This implies some kind of causal relationship that cannot be tested with the existing data in *Understanding Society* because not all manifest variable items are asked in all years. Because the order of events is so important for establishing causation, it would not be wise to rely on imputations for waves when questions were not asked to test a causal model which would likely underestimate any causal effect. To do this, there needs to be an agreement to measure all multidimensional poverty variables in a longitudinal study for an extended length of time, with a sample size of people in persistent poverty large enough for such a complex model to have enough explanatory power to obtain reliable estimates. This is clearly well-beyond the limitations of this thesis and existing data, but can be done, and the above theory can be tested more robustly.

Of course, this theoretical contribution is not without its problems either. The proposition may be specific to certain types of communities, particularly Western, individualised communities, where there is a higher value placed on individual needs and consumption with less attention

paid to societal or collective needs (Hofstede, 1984). Although this construction and theory of multidimensional poverty may be appropriate for the UK, it might not be applicable to other cultures, and would need to be independently validated. As previously stated, concepts like those found in Maslow's (1943) hierarchy have been validated over 123 countries, although there is limited support for the idea in the theory that there is a requirement for lower needs such as subsistence and safety to be met before higher needs can be (Tay & Diener, 2011). Indeed, Doyal and Gough's (1984) conceptualisation of needs argues that all needs are related to societal expectations in some way, so a lack of a strict universal hierarchy of needs is to be expected. If this theory is to be believed then it is important to state that higher-order needs are not always necessarily affected first, and it is not necessarily the case that worse scores on higher-level needs is a precondition for lower-order needs being affected. The ripple from financial strain can skip higher-order needs and affect lower order needs directly and its likelihood of doing so might be mediated by factors such as individual resilience (Luthar, 1999), or social and cultural forms of capital (Bourdieu, 1986). However, this model seems to show that, for the majority, a decrease in life satisfaction almost always coincides with an increase in financial strain.

To fully understand this intricacy within multidimensional poverty it is important to understand the process of *becoming poor*, and the order that manifestations of poverty take. Quantitative research into the experience of becoming poor is limited, and does not contain the same complexity as the multidimensional poverty measure presented here, so cannot truly test the idea of a ripple effect through dimensions, with financial strain at the heart. However, there is some literature that provides a moderate level of support for this chain of events when looking at specific dimensions. Wickham, et al., (2017: 146) studied the effect on maternal mental health when households without any sign of baseline maternal distress transitioned into poverty, defined using the HBAI60% standard, and found that the odds of maternal psychological distress after having falling below the HBAI poverty line were 1.4 times higher than if they had remained out of poverty. Similarly, some evidence from studies tracing transitions out of poverty in India suggests a process that reflects that proposed above. Krishna (2004) queried villages in India about the 'stages-of-process' of exiting poverty, and what kinds of spending and priorities these were associated with. Krishna (2004) found complete similarities across the thirty-five villages in the study, with all of them reporting that the four stages of process were, in order, 'buying food to eat, sending children to school, possessing clothes to wear outside of the house, and retiring debt in regular installments' (123). Although clearly a very

culturally different setting the same general framework can be applied to see how the effect is reversed to that which is proposed in entering poverty; the priority of dimensions went from fulfilling physiological and subsistance needs to meeting the need to participate and avoid stigma.

The strongest support for the multidimensional ripple-effect theory should come from longitudinal or individual-history experiential qualitative accounts of 'becoming poor' that document explicitly how, when, and in what order, dimensions of poverty are experienced. These kinds of ethnographic or longitudinal studies of transitions into and out of poverty are very limited, although feasible narrative accounts can be constructed from general experiences of poverty, where the focus is often on the inability to afford food and energy but lack of social participation is frequently recounted (Hill, et al., 2016). The research sample of interest for these kinds of reports is often people whose level of poverty has reached the point in which they worry frequently about subsistence needs and research about the economisation of participation and gradual decline of life satisfaction in the early stages of financial strain and unacceptable hardship appear to be far more underexplored. This may be because of the longstanding narrative dominance and epistemological authority of absolute poverty, but it is more likely because it would be very difficult, costly, and possibly unethical, to do research on a population of people who the researcher had sampled on the suspicion that they would soon be transitioning into poverty. If this theory is to be utilised beyond this thesis it would be helpful to hear reflections about the applicability of such a framework to the observations of qualitative researchers and people with lived experience, to explore whether it makes sense to them.

This chapter has demonstrated that it is possible to create a time-invariance measure of poverty using the proposed methodology, and argues that this provides some modest insights into the nature of poverty. Poverty is dehumanising and materially draining, and the two can be considered inseparable but distinct dimensions. However, Chapter 2 discussed the need for any measurement of poverty to be sensitive to differences between different groups – as reflected in differences between groups in existing measures and the work of Pantazis, et al. (2004), as well as others. A truly reliable multidimensional measure should broadly be able to equally represent the universal experiences of poverty regardless of a persons membership of certain social groups. The following chapter tests whether the measure developed here meets this criteria.

Chapter 6: Multigroup Invariance

So far the establishment of this multidimensional measure of poverty has focused on its stability over time and underlying structure, which has been shown through the fit statistics of models with various constraints. However, as a dynamic concept, poverty may differ in its construction across social divisions. This can be reflected in loss of good fit when the same dimensions are used for different groups (configural variance), or when the loadings on each dimension are fixed rather than freely estimates (metric variance). A truly universal measure should be able to be applied as identically as possible for all groups. Such a consideration is not made in typical poverty metrics, with the exception of the Minimum Income Standard, and specific levels of low income or certain baskets of goods are assumed to apply identically to different groups with different characteristics. This does not mean that differences between groups in estimates such as factor loadings are unimportant, just that they are not so substantively different that the latent concept for quantitative operationalisation should be changed substantially, meaning that quantitative comparisons can be made. For this reason, this section reports whether the multidimensional measure fits well for different types of household, for men and women, and for people from different cohorts. Visualisations and tables also identify where differences exist, and the extent to which they differ from fixed parameters. It is important to note that any differences do not mean that certain groups are more or less at risk of certain types of poverty, but instead that the underlying construct has different weightings for different groups. In practice, this may be that certain dimensions are overstated or suppressed in their importance when generalised across all groups, so it is useful to know where these differences lie.

The multidimensional poverty measure was assessed for group invariance across three different groupings. This included: seventeen different household types found in the *Understanding Society* derived variables; gender, limited by the data in this analysis to only people who identified as either male or female; and cohort, grouped by decade with the exception of those born between 1910 and 1940 in order to create a group large enough to derive reliable estimates (Wolf, et al., 2013). The main defining criteria for household types was the number of working age or pensionable age adults and dependent children living in a household, and whether adults were couples living in the same household, married, cohabiting, or otherwise or single, classified in the data as either not being in a relationship or not living with their partner/s.

The reason for focusing on these aspects, particularly household type, is to mirror both the approach used to validate the Minimum Income Standard definition of poverty (Hirsch, 2015) and the chapter structure of social divisions outlined in Pantazis, et al.'s analysis of poverty and social exclusion in Britain (2006). The basket of goods for this income-based definition of poverty measure is determined through the consensual method resembling that outlined by Robert Walker (1987). Typically, the needs of certain types of households, such as lone parents, are very different to the needs of other types of household, such as retired couples living without children (Hirsch, 2015). Similarly, dimensions may differ significantly for men and women, particularly when considering the hidden poverty caused by unequal distribution of resources within households, self-sacrifice made by women particularly in relation to caring for children and relatives, and its associated time poverty (Daly, 1992; Glendinning & Millar, 1987; Pantazis & Ruspini, 2006; Vogler, 1994). Likewise, older people 'do not exhibit the same consumption and expenditure patterns as their younger counterparts' (Patsios, 2006: 436), and it is reasonable to suggest that tastes and preferences may differ substantially between generations that may result in the proposed material dimensions being variant between different age groups.

If this measure truly captures some universal material and relational-symbolic dimensions of poverty, we should expect to see a construct that works well regardless of household characteristics, gender, and age/cohort and avoids some of the pitfalls that an overreliance on income can introduce. However, it is important to reiterate that not all potential dimensions of poverty found in theory and qualitative evidence could be reconstructed here using existing secondary data. These missing dimensions may be significantly different between groups. Therefore, any claim that the configuration and weighting of dimensions of multidimensional poverty is invariant between different groups as a result of these findings should be made with caution, as this can only be claimed for the dimensions that have been measured by *Understanding Society*.

These three areas do not represent a complete list of all social divisions that should be assessed for measurement invariance. For one, the main *Understanding Society* individual response questionnaire only covers adults. The same measure of multidimensional poverty cannot be applied directly to children. The dimensions of poverty that exist for people in childhood may look very different to those that have been found to fit well for adults. The metric that has been developed and much of the literature and theory it derives from is adult-centric. Rather than

being treated as a stage in the life course as deserving as any other, children and young people's concerns and experiences can often be trivialised or treated very differently, usually for logistical and ethical reasons (Corsaro, 2017; McAuley & Rose, 2010). However, the breakpoint between adult and child is quite arbitrarily imposed by law and the kind of data collected shows very little resemblance in its design to the equivalent items for adults. Although there are some good reasons for this, including children's ability to fully understand certain questions, and their willingness to complete a lengthy and sometimes complicated questionnaire being at the forefront, there may be legitimate concerns to them being treated so differently, to the extent that exploring transitions of universal experiences between childhood, adolescence, and adulthood become nearly impossible. There are equally good reasons why quantitative social inquiry should seek to be designed in a way that is inclusive for both children and adults, however difficult this may be. By virtue of their membership of society, and in pursuit of truly universal theories of human need, it would be excellent to see fewer dichotomies where one theory is developed for adults and another for children using completely different data. Unfortunately, the reliance on secondary data makes this impossible in this thesis.

A further limitation is that invariance has not been tested for different geographies or ethnicities, nor for intersections of gender, ethnicity, and place. Invariance testing for geographies across the UK beyond basic groupings of the four nations is not possible without negotiating special access to geographical identifiers in the *Understanding Society* survey, which was not possible in the time span of the project. In addition, ethnicity is a complex factor that includes the interplay between self-identification, heritage, and the self-defined strength of association with such heritage, which is frequently oversimplified and diluted in quantitative research (Aspinall, 2000). Meaningful categorisations of ethnic groups involves a commitment to sufficiently detailed categorisation to avoid conflating the lived experiences of different ethnic groups or identities by what may be arbitrary criteria that often ignore self-identification. Although *Understanding Society* has variables that facilitate these more nuanced groupings (Burton, et al., 2008), there are considerable problems with trying to meaningfully analyse differences between these groups in more complex methods. Even considering the oversampling of certain groups undertaken in the survey, the problem is that structural equation modelling and latent variable estimation require a large sample size. Using Monte-Carlo simulation, Wolf, et al. (2013: 21) show that even for a 3-factor model with factors with 3 or 4 manifest variables and where item loadings are between 0.5 and 0.8, sample sizes for reliable estimation of between

150, for loadings of 0.8, and around 420, for loadings of 0.5, are needed. Our multidimensional poverty construct is a higher order factor of four first-order and one second-order factors, many of which contain only 4 indicator variables, so the sample size requirement is likely to be a magnitude higher and not able to be obtained using a meaningfully refined categorisation of ethnic groups.

This leaves a choice of whether to use a large combined ethnic grouping variable, or to forgo invariance testing for ethnicity because of an insufficient sample size. When ethnic groups are combined to 'make up' sample size this often results in suppression effects, for example, where one ethnic group scores lower than a reference group on a specific variable and another scores higher - when they are combined the new estimate has less error but becomes closer to the reference group estimate and does not really reflect the reality of either group (Connelly, et al., 2016). Is it better to risk making sweeping generalisations about the adequacy of a certain metric for a group in society that are at greater risk of poverty (Alcock, 2012), but to make an attempt at assessing it, no matter how inadequately, or is it better to admit that it cannot be done responsibly with the existing data? Although being able to claim that this measure of poverty is invariant across ethnic groups would be valuable, it is not ethically justifiable to risk misrepresenting its adequacy for certain groups of people to do so. Perhaps with enough pooled data with ethnic group boosted samples it will be possible in future.

Lastly, there are similar difficulties doing meaningful invariance testing between groups of people that live with or have lived with disabilities. Using a social model of disability definition, people living with impairments which lead to them being disabled by a societal structure in which inadequate provisions are made to enable their full participation, are frequently overrepresented in poverty statistics, suggesting they are at greater risk of poverty (Heslop & Gordon, 2014; Beresford, 1996). Survey methods do not usually capture the amount of nuance within types of impairment and disabilities and typically rely on a broad categorisation of whether someone reports having a long-standing illness or not. This ignores the diverse types of disabling experiences that exist, the subtlety between impairment and disability in the social model, and the important intersections that exist between, say, disability and age (Goethals, et al., 2015; Beresford, 1996). For example, people with impairments in older age may face less disability in society due to the growing societal adaptation and expectation of health-related impairments in older age; the experience of equivalent levels of impairment may lead to far

greater disabling experiences for younger and middle-aged people. This may fundamentally change the dimensions of poverty that apply most strongly to them.

This results in the same dilemma that is faced in deciding whether to conduct tests of measurement invariance between different ethnic groups. Invariance could be established between two groups as defined by the very blunt 'long-standing illness' variable, however the extent to which this would represent a fair assessment of whether the measure can be applied equivalently to quantitatively understand the experience of poverty for people with differing levels of impairment and disability is questionable. The assumption is that all of those living with long-standing illnesses face disability, but according to the social model this may not necessarily be true. If those categorised as having long-standing illnesses are largely not precluded from full participation this may mask the real difference in the experience of poverty of those for whom society imposes disability. Under the very large presupposition that adequate measures exist to evaluate impairment versus disability, and that these are then partitioned into groups to allow for intersectionality at least in relation to age, there are likely to be very few cases within the resultant groups in a longitudinal survey, leaving the same problem as trying to conduct invariance testing on meaningful categorisations of ethnic groups (Wolf, et al., 2013). For this reason it was again decided that no analysis was ethically better than a poor analysis that risks misrepresenting the applicability of the measure to people living with disability.

6.1 Measurement Invariance between Household Types

Results from multi-group invariance testing for household types shows that the multidimensional measure of poverty still has satisfactory fit statistics when assessed across the 17 different types of household identified in *Understanding Society* with fixed loadings, residual variances, intercepts and thresholds (Table 8, full invariance fit statistics: CFI = .917, TLI = .925, RMSEA = .047). This means that, within reason, the multidimensional poverty construct can be applied across different types of household with the same interpretation, allowing for valid comparisons. There are differences in the salience of different dimensions of poverty for certain household types, but overall these differences are not so extreme that a different construct is required to capture poverty for certain households. These differences are visualised in figure 30 and are transcribed in table 7.

As an example to demonstrate how to interpret one of the radar plots shown in figure 30, consider the top-left plot showing the factor loadings for households that meet the criteria for being 'retired couples living with no dependent children'. This household type makes up 15 per cent of the households in the *Understanding Society* survey. The thick black line indicates the position of the radar points if the household factor loadings were identical to the weighted average factor loadings (deviations of 0). Radar plot points inside the black line indicate that dimensions are less salient, meaning they fit less well with the underlying construct than the average. Radar plot points outside the line indicate higher than average salience, implying they are a more consistent feature of the larger construct. A group-invariant construct should load approximately the same as average for all groups. PSYST refers to Psychosocial Strain, FINSTR refers to Financial Strain, LSAT refers to Life Satisfaction, MATDEP1 refers to Material Deprivation of Consumables, and MATDEP2 refers to Material Deprivation of Commodities. In this first example, then, we can see that psychosocial strain and material deprivation of commodities load slightly higher onto the multidimensional poverty factor than average, and life satisfaction loads about the same. By contrast, material deprivation of consumables and financial strain load lower than average.

Retired couples and single female pensioner household types show the lowest levels of salience across multiple proposed dimensions of poverty, specifically in relation to financial strain and material deprivation of activities and services for couples, and financial strain and material deprivation of household goods for lone female pensioners. On one hand, this may reflect that the importance of psychosocial and life satisfaction dimensions begin to diverge from material dimensions after retirement; on the other hand, this may be an artefact of the problems associated with the wording of material deprivation interview questions for older people (McKay, 2008). As is shown later in the comparison of different cohorts this difference is even more profound for those born between 1910 and 1939, with financial strain and both material deprivation dimensions showing between -0.10 and -0.16 lower loadings than the weighted average across households. Although material dimensions are still central in the multidimensional construct of poverty for these retired households, with loadings above 0.7, the psychosocial health dimension and life satisfaction dimension begin to become more important.

Other household types show notable differences in the salience of certain dimensions. For example, material deprivation of household goods appears to have a much lower loading for

households with three single adults and no children and for households with two people of pensionable age sharing a home that are not in a relationship. Any proposed reasons for this would be speculative at this point, but it may reflect the associated age demographics or particular living preferences of people living in house shares. Similarly, couples with children have lower salience in the estimates for the psychosocial strain factor in their multidimensional poverty construct. This may be a consequence of the positive association that relationship status has on self-assessed psychological wellbeing (Kamp Dush & Amato, 2005), where the negative psychosocial impact of financial strain and material deprivation may be negated slightly by a supportive personal relationship. Life satisfaction, in contrast, appears identical between all household types.

Figure 30: *Differences in loadings for dimensions of poverty for different family types. Higher deviations indicate the dimensions are more salient, lower loadings indicate they are less salient.*

Table 7: Loadings, fit statistics, and difference from weighted average loadings from a configural model for 17 types of households (2016) in the United Kingdom

Model Fit Indices	**All Imputation Mean**
CFI:	0.901
TLI:	0.911
RMSEA:	0.052
WRMR:	15.629

Weighted Average Loadings (Weighted by % HHs)	
Financial Strain	0.812
Material Deprivation (Consumables)	0.578
Material Deprivation (Commodities)	0.261
Psychosocial Strain	0.614
Life Satisfaction	0.912

Male Pensioner No Children (2.1% of households)	**All Imputation Mean**	**Difference**
Financial Strain	0.819	0.007
Material Deprivation (Consumables)	0.638	0.061
Material Deprivation (Commodities)	0.224	-0.038
Psychosocial Strain	0.569	-0.045
Life Satisfaction	0.906	-0.006
	Sum of Squared Differences	0.0072

Female Pensioner No Children (5.2% of households)	**All Imputation Mean**	**Difference**
Financial Strain	0.750	-0.062
Material Deprivation (Consumables)	0.542	-0.036
Material Deprivation (Commodities)	0.176	-0.086
Psychosocial Strain	0.668	0.053
Life Satisfaction	0.915	0.003
	Sum of Squared Differences	0.0152

Single WA Adult No Children (6.9% of households)	All Imputation Mean	Difference
Financial Strain	0.874	0.062
Material Deprivation (Consumables)	0.712	0.134
Material Deprivation (Commodities)	0.310	0.049
Psychosocial Strain	0.692	0.078
Life Satisfaction	0.915	0.003
	Sum of Squared Differences	0.0304

Single WA Adult 1 Child (2% of households)	All Imputation Mean	Difference
Financial Strain	0.817	0.005
Material Deprivation (Consumables)	0.614	0.036
Material Deprivation (Commodities)	0.220	-0.041
Psychosocial Strain	0.640	0.026
Life Satisfaction	0.927	0.015
	Sum of Squared Differences	0.0039

Single WA Adult 2+ Children (2.3% of households)	All Imputation Mean	Difference
Financial Strain	0.780	-0.032
Material Deprivation (Consumables)	0.539	-0.039
Material Deprivation (Commodities)	0.187	-0.074
Psychosocial Strain	0.648	0.034
Life Satisfaction	0.921	0.009
	Sum of Squared Differences	0.0092

Couple WA No Children (12.3% of households)	All Imputation Mean	Difference
Financial Strain	0.854	0.042
Material Deprivation (Consumables)	0.649	0.071
Material Deprivation (Commodities)	0.267	0.006
Psychosocial Strain	0.627	0.013
Life Satisfaction	0.909	-0.003
	Sum of Squared Differences	0.0071

Couple Retired No Children (15% of households)	All Imputation Mean	Difference
Financial Strain	0.745	-0.067
Material Deprivation (Consumables)	0.467	-0.110
Material Deprivation (Commodities)	0.286	0.025
Psychosocial Strain	0.628	0.013
Life Satisfaction	0.918	0.006
	Sum of Squared Differences	0.0174

Couple WA 1 Child (7.5% of households)	All Imputation Mean	Difference
Financial Strain	0.837	0.025
Material Deprivation (Consumables)	0.602	0.024
Material Deprivation (Commodities)	0.287	0.026
Psychosocial Strain	0.563	-0.051
Life Satisfaction	0.903	-0.009
	Sum of Squared Differences	0.0045

Couple WA 2 Children (10.2% of households)	All Imputation Mean	Difference
Financial Strain	0.836	0.024
Material Deprivation (Consumables)	0.596	0.018
Material Deprivation (Commodities)	0.274	0.013
Psychosocial Strain	0.556	-0.058
Life Satisfaction	0.898	-0.014
	Sum of Squared Differences	0.0046

Couple WA 3+ Children (5.2% of households)	All Imputation Mean	Difference
Financial Strain	0.844	0.032
Material Deprivation (Consumables)	0.601	0.023
Material Deprivation (Commodities)	0.253	-0.008
Psychosocial Strain	0.584	-0.031
Life Satisfaction	0.909	-0.003
	Sum of Squared Differences	0.0026

Two WA Adults, Non-couple, No Children (2.6%)	All Imputation Mean	Difference
Financial Strain	0.832	0.020
Material Deprivation (Consumables)	0.614	0.037
Material Deprivation (Commodities)	0.224	-0.037
Psychosocial Strain	0.694	0.080
Life Satisfaction	0.915	0.003
	Sum of Squared Differences	0.0096

Two Retired Adults, Non-couple, No Children (1.8%)	All Imputation Mean	Difference
Financial Strain	0.782	-0.029
Material Deprivation (Consumables)	0.592	0.015
Material Deprivation (Commodities)	0.182	-0.079
Psychosocial Strain	0.662	0.048
Life Satisfaction	0.917	0.005
	Sum of Squared Differences	0.0096

Two WA Adults, Non-couple, living with Children (1.8%)	All Imputation Mean	Difference
Financial Strain	0.832	0.020
Material Deprivation (Consumables)	0.560	-0.018
Material Deprivation (Commodities)	0.291	0.030
Psychosocial Strain	0.620	0.005
Life Satisfaction	0.921	0.009
	Sum of Squared Differences	0.0017

3+ Adults, no couples, 1 child (12.3% of households)	All Imputation Mean	Difference
Financial Strain	0.807	-0.005
Material Deprivation (Consumables)	0.549	-0.029
Material Deprivation (Commodities)	0.253	-0.008
Psychosocial Strain	0.614	-0.001
Life Satisfaction	0.914	0.002
	Sum of Squared Differences	0.0009

3+ Adults, one couple, 1-2 children (8.2% of households)	All Imputation Mean	Difference
Financial Strain	0.802	-0.010
Material Deprivation (Consumables)	0.543	-0.034
Material Deprivation (Commodities)	0.280	0.018
Psychosocial Strain	0.576	-0.038
Life Satisfaction	0.914	0.002
	Sum of Squared Differences	0.0031

3+ Adults, no couples, no children (2.3% of households)	All Imputation Mean	Difference
Financial Strain	0.795	-0.017
Material Deprivation (Consumables)	0.584	0.006
Material Deprivation (Commodities)	0.147	-0.114
Psychosocial Strain	0.634	0.020
Life Satisfaction	0.918	0.006
	Sum of Squared Differences	0.0138

3+ Adults, no couples, 1+ children (2.3% of households)	All Imputation Mean	Difference
Financial Strain	0.793	-0.018
Material Deprivation (Consumables)	0.567	-0.010
Material Deprivation (Commodities)	0.309	0.048
Psychosocial Strain	0.565	-0.049
Life Satisfaction	0.913	0.001
	Sum of Squared Differences	0.0051

There are other household types for which the dimensions were more salient, suggesting that the proposed multidimensional poverty construct works better for these households than average. Most noticable are the better than average factor loadings for households comprised of single working age adults with no children. For single working age adults all dimensions have much higher factor loadings than average, reflecting that the chosen configurations of items into factors explains a small but noticeable proportion more variance (around 3 per cent). This may reflect that our theoretical perspectives are often more centred around the individual, particularly the working age individual, and their view of themselves than they are around the household or community. These individual dimensions may become less applicable later in the life course, or in different stages throughout, by more collective dimensions of poverty. To a lesser extent this is also found in couple households with no children.

It should be reiterated that despite these differences the tests show that, as a universal constrct approximating poverty along these proposed dimensions, the measure shows that it can be employed in a meaningful and reliable way regardless of household type.

Table 8: Model Fit and Factor Loadings for Full Invariance between Household Types

Model Fit Statistics	**All imputation mean**
CFI:	0.917
TLI:	0.925
RMSEA:	0.047
WRMR:	16.000
Factor Loadings	**All imputation mean**
Financial Strain	0.822
Material Deprivation (Consumables)	0.588
Material Deprivation (Commodities)	0.261
Psychosocial Strain	0.614
Life Satisfaction	0.916

6.2 Measurement Invariance between Men and Women

Measurement invariance testing for gender showed that the multidimensional poverty measure had factor loadings and intercepts that were virtually identical for men and women. There were miniscule differences in that psychosocial strain and material deprivation of consumable activities, goods, and services were slightly higher weighted for women than for men. On a standardised scale, this difference was only around -0.015 for both. Although this may represent a difference between men and women's experiences of poverty that may surface in qualitative accounts, the model shows that in terms of a general universal measure of poverty the dimensions can be weighted approximately equally without losing model fit (full invariance fit statistics: CFI = .939, TLI = .941, RMSEA = .045). The construct can therefore be applied to measuring poverty for both men and women.

This corroborates with previous findings from the poverty and social exclusion survey that the mean response for men and the mean response for women across all material deprivation items were in 92 per cent agreement, with women very slightly less likely than men to report things were necessities across the board (Pantazis, et al., 2006: 99). The authors go so far as to say:

> "[this] demonstrate[s] that with regard to most items and activities the views of men and women are indistinguishable, with effectively identical percentages of men and women in agreement (or disagreement) about what constitutes the necessities of life."
>
> Pantazis, et al., *The necessities of life*, 2006: 100

The proposed measure can be equally informative for both men and women, and allows comparison on an equivalent scale. Income based metrics, and their associated equivalisation, have long been considered problematic for being 'gender-blind' (Pantazis & Ruspini, 2006: 379). This multidimensional measure therefore represents an alternative that can be assessed for validity explicitly, as above, to ensure that the questionnaire items and their arrangement into a factor is meaningful for both men and women. There are still problems with this approach, certain items, for instance, may not be measured due to the gendered dimensions of survey research, however it is arguably better than the 'black-box' approach used in most income-based measures.

This leaves a significant question of whether this measure can contribute to quantitative understandings of *intra*-household poverty, the highly gendered and hidden 'dark-side' of poverty (Pantazis & Ruspini, 2006; Burchardt & Karagiannaki, 2018). The answer is yes, but only for some dimensions. Because poverty is typically conceptualised at the household level there is the potential for poverty of individuals to be underestimated when unequal distribution of resources within households exists. A household may therefore appear to not be at risk of poverty but members of that household could already be living in poverty. Certain variables that make up the multidimensional poverty measure are measured at the individual level, where each individual in a household provides their own answer, and some are measured at the household level, where the question is answered as a household and assumed to be the same for all individuals (see table 4 in the methodology chapter for a list of which variables are at the household level). Individual dimensions, which include psychosocial strain and life satisfaction, can be assessed for differences within households. Unfortunately, both material deprivation and the financial strain factors contain items that are almost exclusively measured at the household level. This measure would allow researchers to investigate whether there are intra-household differences in psychosocial strain and life satisfaction dimensions of poverty, which may represent some inequalities in financial distribution, but it falls short of being able to provide a way to measure intra-household material deprivation. This may be a consideration that survey methodologists wish to take forwards to provide better data on which to explore intra-household inequality.

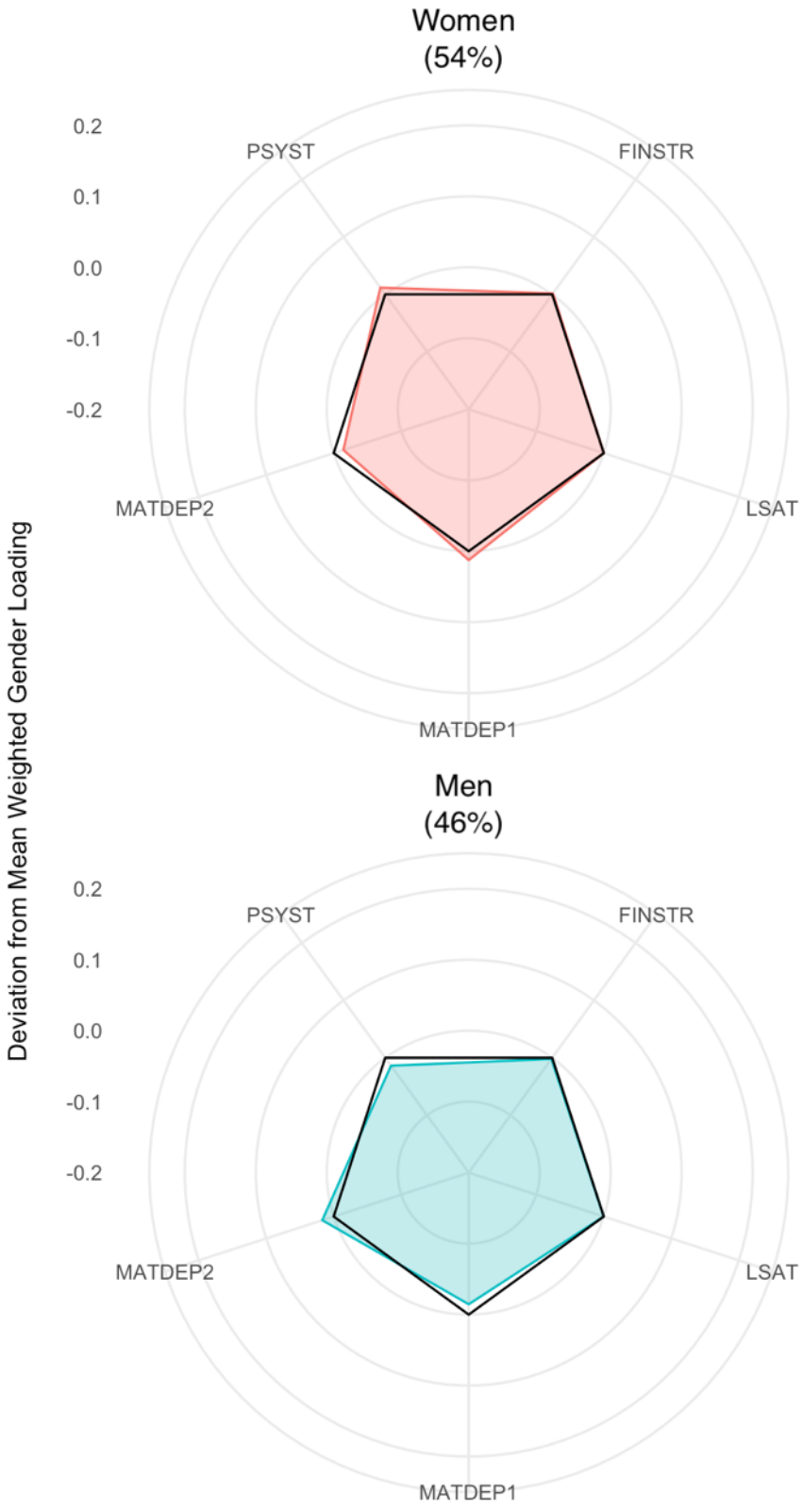


Figure 31: *Differences in factor loadings for men and women. The five dimensions appear to be as salient for men as for women, with some very slightly increased emphasis on psychosocial strain (PSYST) and participation (MATDEP1) for women that may be reflected in experiential accounts.*

Table 9: Model fit statistics and loadings for configural invariance models for multidimensional poverty and gender.

Model fit statistics	**All imputation mean**	
CFI:	0.925	
TLI:	0.926	
RMSEA:	0.050	
WRMR:	12.299	
Weighted Average Loadings		
Financial Strain	0.878	
Material Deprivation (Consumables)	0.641	
Material Deprivation (Commodities)	0.285	
Psychosocial Strain	0.593	
Life Satisfaction	0.904	
Male	**All imputation mean**	**Difference**
Financial Strain	0.876	-0.002
Material Deprivation (Consumables)	0.627	-0.015
Material Deprivation (Commodities)	0.302	0.017
Psychosocial Strain	0.579	-0.014
Life Satisfaction	0.904	0.000
	Sum of Squared Differences	0.0006
Female		
Financial Strain	0.880	0.002
Material Deprivation (Consumables)	0.654	0.012
Material Deprivation (Commodities)	0.271	-0.014
Psychosocial Strain	0.605	0.012
Life Satisfaction	0.905	0.000
	Sum of Squared Differences	0.0005

Table 10: Model Fit and Factor Loadings for Full Invariance between Genders.

Model Fit Statistics	**All imputation mean**
CFI:	0.939
TLI:	0.941
RMSEA:	0.045
WRMR:	12.329

Factor Loadings	**All imputation mean**
Financial Strain	0.879
Material Deprivation (Consumables)	0.643
Material Deprivation (Commodities)	0.285
Psychosocial Strain	0.592
Life Satisfaction	0.904

6.3 Measurement Invariance between Cohorts

Before outlining any differences and invariance between cohorts it is necessary to remind the reader of the impossibility of separating the effects of age, periods of time (years of measurement/events), and cohorts because of their inherent exact collinearity, this is known as APC-analysis and the APC-identification problem (Bell & Jones, 2013; 2014). It is not possible *quantitatively* to evaluate whether differences between people of different ages are a consequence of their cohort membership, conditions unique to their upbrining in a specific historical context, or of the ageing process and specific changes associated with growing older, or of the specific context of the time period at which they are being measured. This is because on identical scales 'age', 'period', and 'cohort' are perfectly linearly related. If we know any two of the three we know the value of the third. Regression models will only converge when the three are on different scales. This means that the decomposition of variance of any dependent variable between the three factors is completely arbitrary, as there are multiple equally good solutions (*ibid*). This matters because although the following analysis is referred to in terms of cohort membership, it may also reflect differences between age groups related to ageing rather than related to the context their cohort's upbrining. Both are equally valid explanations for any differences, from a quantitative standpoint.

Multigroup invariance tests showed that full invariance conditions of fixed factor loadings, intercepts/thresholds, and residual variances across all cohort groups retained satisfactory-to-good model fit statistics (table 12: CFI = .929, TLI = .935, RMSEA = .047). This implies that

the poverty construct can be applied to all age groups without a loss in its substantive meaning and application. There is sometimes a tendency in popular discourse to imagine the experience of poverty as so qualitatively different for older people that it requires an explicit redirection of focus onto the dimensions of health and social exclusion, or exclusive sets of material deprivation items. This is quite patronising and debilitating view of older people as detatched from the rest of society, living with the wants of the past, not engaging in new societal desires or advancements, preoccupied with health concerns or plagued by chronic illness (Walker, 1980; Walker & Naegele, 1999). There is no strong evidence that the meaning of poverty changes drastically by the virtue of getting older or being born at a certain time, these dimensions appear to be universal.

Figure 32: *Differences in factor loadings for different cohorts. The multidimensional measures seems to have the most salience with people born between the 1950 and 1980. It fits data from those born between 1910 and 1950, and those born after the 1980s, worse (but not significantly worse).*

Table 11: Model fit statistics and loadings for configural invariance models for multidimensional poverty and cohort.

Model fit	**Average**	
CFI:	0.916	
TLI:	0.923	
RMSEA:	0.052	
WRMR:	14.096	
Global Weighted Loadings		
Financial Strain	0.854	
Material Deprivation (Consumables)	0.624	
Material Deprivation (Commodities)	0.337	
Psychosocial Strain	0.604	
Life Satisfaction	0.898	
Cohort		
1910-1939	**All imputation mean**	**Difference**
Financial Strain	0.736	-0.117
Material Deprivation (Consumables)	0.470	-0.154
Material Deprivation (Commodities)	0.199	-0.138
Psychosocial Strain	0.594	-0.010
Life Satisfaction	0.903	0.006
	Sum of Squared Differences	0.057
1940s	**All imputation mean**	**Difference**
Financial Strain	0.813	-0.041
Material Deprivation (Consumables)	0.532	-0.092
Material Deprivation (Commodities)	0.340	0.003
Psychosocial Strain	0.617	0.013
Life Satisfaction	0.906	0.008
	Sum of Squared Differences	0.010
1950s	**All imputation mean**	**Difference**
Financial Strain	0.883	0.029
Material Deprivation (Consumables)	0.684	0.060
Material Deprivation (Commodities)	0.420	0.083
Psychosocial Strain	0.662	0.058
Life Satisfaction	0.905	0.007
	Sum of Squared Differences	0.015

1960s	**All imputation mean**	**Difference**
Financial Strain	0.896	0.042
Material Deprivation (Consumables)	0.705	0.081
Material Deprivation (Commodities)	0.393	0.056
Psychosocial Strain	0.638	0.034
Life Satisfaction	0.897	-0.001
	Sum of Squared Differences	0.013

1970s	**All imputation mean**	**Difference**
Financial Strain	0.889	0.035
Material Deprivation (Consumables)	0.683	0.059
Material Deprivation (Commodities)	0.356	0.019
Psychosocial Strain	0.592	-0.012
Life Satisfaction	0.887	-0.011
	Sum of Squared Differences	0.005

1980s	**All imputation mean**	**Difference**
Financial Strain	0.869	0.016
Material Deprivation (Consumables)	0.631	0.007
Material Deprivation (Commodities)	0.258	-0.079
Psychosocial Strain	0.573	-0.031
Life Satisfaction	0.891	-0.007
	Sum of Squared Differences	0.008

1990s	**All imputation mean**	**Difference**
Financial Strain	0.812	-0.042
Material Deprivation (Consumables)	0.546	-0.078
Material Deprivation (Commodities)	0.310	-0.027
Psychosocial Strain	0.539	-0.065
Life Satisfaction	0.902	0.004
	Sum of Squared Differences	0.013

Because the differences between people inadvertently become the focus of any analysis this can often detract from the similarities they share. Pantazis, et al., (2006: 101) similarly found that although there were some differences and that, overall, younger people were less likely to agree that certain material deprivation factors constituted necessities than older people, there

was a high rate of agreement between the two groups in the general correspondence of agreement ($R^2 = 0.81$). Although these differences are important, and should be highlighted, attempts to measure poverty using one multidimensional measure needs to focus on universal needs. Focus on social exclusion and health concerns that are more pronounced in older populations (Hoff, 2008; Dominy & Kempson, 2006) should not subvert attempts to develop an indicator that can be applied to the entire population.

Table 12: Model Fit and Factor Loadings for Full Invariance between Cohorts.

Model Fit	**All imputation mean**
CFI:	0.929
TLI:	0.935
RMSEA:	0.047
WRMR:	14.679

Factor Loadings	**All imputation mean**
Financial Strain	0.870
Material Deprivation (Consumables)	0.642
Material Deprivation (Commodities)	0.339
Psychosocial Strain	0.604
Life Satisfaction	0.895

Where differences exist between cohorts they appear to follow a trend where the multidimensional poverty index fits better for those in middle-age at the time of the survey administration compared to those who are either older than around sixty-years old or younger than around thirty-years old. The most pronounced differences for older people are in the salience of the material deprivation and financial strain dimensions, while life satisfaction and psychosocial strain retain almost identical factor loadings to the weighted average.

To some extent, this may be an artefact of the finding that some of the wording of material deprivation items in surveys is interpreted differently by older people (McKay, 2008) however, in and of itself, the model fit statistics here suggest that this difference is not large enough to suggest that the survey items are unreliable for measuring deprivation in the older population. This was also found by McKay (2008: 14) using Cronbach's alpha as a measure of question item reliability. The specific dimensions that are less salient for older people contradict the findings of McKay (2008) where regular savings for funeral costs and unexpected expenses,

and regular holidays were considered more important to those over 60, although this may be because the classification of over-60 is too broad to capture the diversity of opinions from older people, especially as other research finds notable differences between younger retired couples and older retired couples (Patsios, 2006). Ultimately, the lack of salience may be due to lifestyle or health changes in older age resulting in them receeding from certain activities or withdrawing from or having certain responsibilities withdrawn from them. Examples of this might be discontinuing driving due to health conditions, or no longer having responsibility for maintaining living conditions, furnishings, or major electrical goods in instances where people have moved into supported or otherwise assisted living conditions either with family or in social care arrangements.

One reason for financial strain being less salient for older people, and especially those aged over around 70-years old, may be that financial strain captures more of the experience of insecure or unstable income, consumption, and one-off unexpected expenses. Although pensioners have historically been more likely to experience low income (Price, 2006; Patsios, 2006; Dominy & Kempson, 2006; JRF, 2018), a number of policy decisions, especially the introduction of the 'triple-lock' pension, the council tax pension credit, in addition to winter fuel allowances and the relatively stable delivery of pensions compared to the delivery of frozen and increasingly means-tested working age benefits are likely to afford some additional security, or at least, some security specific to housing and council tax payments on which the financial strain dimension is based. This is probably compounded by the lessened degree of unexpected costs as a result of responsibilities towards dependents as the children of people in their seventies and above become more financially secure in their middle-age. However, older people may still experience different kinds of financial strain that are not captured by the dimension.

For younger people, particularly those in their twenties or younger at the time of the survey, financial strain, psychosocial strain, and material deprivation fit less well than for middle-aged people. This may be for some of the same reasons that financial strain and material deprivation fit less well for older people. Younger people may not yet be in a stage of the life course where they are responsible for significant regular payments that can lead to financial strain, they may be more likely to be receiving support from broader family networks, and they may not yet have accrued the material responsibilities that are common in later life. These include responsibilities for the state of repair of property and furnishings, the provision of meals,

holidays, or major electrical goods, especially white goods, which may still variably be the responsibility (or joint responsibility) of private landlords, housing associations, parents or guardians. These responsibilities for many will invariably transition further onto the individual as they continue through the life course, accruing economic capital.

The slight negative deviation in the salience of the psychosocial strain factor for younger people may reflect the correlation between general physical health and general mental health. While the GHQ and WEMWBS questions are construed as relating directly to mental health they are likely to capture a large amount of covariance from physical health, as shown in research exploring the relationship between the GHQ-12 (general psychological) and EQ-5D (general physical) questionnaire items (Knott, 2012). However, there is no strong evidence for the opposite occuring at the other end of the age distribution, considering the psychosocial strain factor does not show the inverse relationship in older age groups. The full reasons for this non-significant difference in salience may only be uncoverable with a more focussed analysis on measures of psychosocial strain and their interpretation amongst younger age groups, as well as a better understanding of the relationship between general psychosocial health and poverty. However, it may be reasonable to suggest that any analysis of poverty using the proposed multidimensional measure include controls for general physical wellbeing, for which the SF-12 questionnaire items in *Understanding Society* can be used (Mukuria, 2016: 13). Regardless, the invariance testing has shown that the scale of these differences is not so large as to suggest that the construct cannot be applied equally between all cohorts.

6.4 Conclusions of multigroup invariance testing

This analysis suggests that the multidimensional poverty measure constructed using items from the *Understanding Society* survey is reliable regardless of the group it is applied to. Despite there being clear differences between groups along the lines of social divisions, for the purpose of a universal metric the measure does not differ substantially in its construction between household groups, cohort membership, or between men and women. This does not mean that differences in the nature of poverty between different groups are not important, merely that there is not strong evidence that unique measures are required for specific groups to be meaningful – a general measure can be applied. This, in turn, makes a more complex multidimensional measure of poverty derived from this method more desirable as one way of quantitatively operationalising the concept of poverty (Lister, 2004).

One point of additional interest which has not been raised in the discussions of each set of groups itself is the complete consistency of the factor loading for the life satisfaction domain. For every group: all cohorts, men and women, different households, life satisfaction is equally and very strongly weighted as a dimension of poverty. This again raises the important question of why life satisfaction, at the intersection of subjective feelings of success, strain, and stability, which inherently capture both absolute and relative positionalities of people relating simultaneously to the material, relational, physical, psychological, and financial dimensions of life is not considered more seriously as a valuable indicator of poverty, especially when used in tandem with material measures. In one sense, it suggests a lack of trust in peoples' own interpretations of the quality of their lives which is endemic in quantitative research in its search for positivist objectivity (Jenkins, 2002). A search that we considered in the methodology to be broadly impossible. Poverty is experienced at the intersection of all of these factors, and life satisfaction may represent a universal 'hub' for the sociorelational ones, in the same way that financial hardship sits as the hub for material deprivation factors. It is important to represent in some way some dimension that may manifest in peoples' assessments of themselves and their position in society in response to voicelessness and injustice. The way the measure is constructed seems to show that happiness is important to everyone, regardless of time, cohort, gender, or household, and that it is grossly impeded by exposure to the material and psychological dimensions of poverty: material deprivation, financial strain, and poor psychosocial health.

It has already been mentioned that the measure itself probably does not go far enough in its complexity. A number of dimensions were not able to be measured, and therefore could not be tested for invariance. This means that the invariance testing undertaken has been limited in its scope due in part to sample size, and in part due to the variables available. This leaves the quality of the measure in an unusual position. Factor analysis is not often used in sociology; invariance testing of factors is an even rarer occurrence, usually only seen in interdisciplinary work with psychologists. In this regard, the work is an advancement in the methodological sophistication of measurement construction in the field, but still falls short of testing all of the dimensions that emerge from the literature, both with regards to social divisions and with regards to poverty. Most importantly, we do not know if the measure that was constructed can be applied equivalently to those living with disabilities and those from different ethnic groups, but also at the intersections of social divisions. If a holistic multidimensional measure of

poverty is to be adopted in social research it is essential that these challenges, in addition to the technical and survey methodology challenges, are addressed.

Chapter 7: Conclusions

The conclusions of the empirical findings of the thesis are constructed around responses to four questions:

- Is there support from exploratory data analysis for the construction of a multidimensional measure of poverty using the proposed domains identified theories and qualitative accounts of the lived experience of poverty?
- What does a quantitative operationalisation of multidimensional poverty consist of, how well does it work, and what can it tell us about the accuracy of existing theories of poverty?
- Is the proposed multidimensional measure of poverty and its individual dimensions statistically reliable and valid over time and between different groups?
- What are the technical, methodological, and political limitations of the measure?

In addition to this, it is hoped that the logic and substantive meanings behind each stage of the process has been accessible and interpretable even for a reader without specialism in quantitative research methods. Some of the technical aspects of this research are difficult to communicate, but the hope is that the meanings have not been obscured by too much statistical jargon. Research should lead to dialogue, and dialogue cannot happen if one party is made to feel shut out of participation. Further to this, there has been an attempt to describe challenges as reflexively and openly as possible, in an effort to present methodological objectivity without the assumption of value objectivity (Jenkins, 2002).

7.1 Exploratory Analysis

By constructing a visual representation of the clusters of relationships shared between multiple indicators in the *Understanding Society* survey that were chosen because of their relevance to literature examining the nature of poverty, the research has shown that it is possible to begin to construct proposed factor solutions for Confirmatory Factor Analysis without the reliance on *p*-values (Gelman & Loken, 2013). The meanings of these clusters of items has been derived through a more qualitative examination of the patterns, and largely supports the dimensions of poverty that are expressed very clearly in theories of multidimensional poverty that have been developed by Townsend (1979), Baulch (1996), Lister (2004), Tomlinson, et al., (2008) and Spicker (2009). These dimensions can be labelled as: social isolation (with regards to personal

friendships); financial strain; material deprivation of consumable goods, activities, and services; material deprivation of household commodities; civic participation and community cohesion; life satisfaction; general psychosocial wellbeing; and local service quality.

Of these, financial strain, material deprivation of both commodities and consumables, life satisfaction, and general psychosocial health seemed to share a very strong relationship to an underlying factor that could be seen in the intersections of clusters. This raised some early questions about the definition of social isolation versus social exclusion, and the dynamics this involves, especially in that social exclusion in poverty does not mean that people suffer worse interpersonal relationships. Researchers should be cautious about inadvertently conflating these two factors, and more focus may be needed on how to adequately measure social exclusion (or to what extent it should be thought of as political or societal, rather than social, exclusion). This was similar to what was seen with civic participation and exclusion, where the types of measurements needed to construct an appropriate measure for civic exclusion are not really available in the chosen dataset. This analysis provided a foundation with which to conduct confirmatory factor analysis and assess the reliability and validity for any factors and a higher-order multidimensional poverty measure, while safeguarding against the temptation of pursuing on statistically significant results that may have been a product of random noise in the data.

7.2 Confirmatory Factor Analysis and Reliability/Validity Testing

With the data separated into exploratory and validation subsets across five iterations, and the use of the Fornell-Larcker criterion (Fornell & Larcker, 1981), it was possible to establish whether the factors that had been identified through the exploratory analysis and informed by the literature review represented reliable and valid underlying constructs and suggest improvements for measurement. All of the first-order constructs identified had excellent model fit, although civil exclusion had other methodological problems. This reflected the effectiveness of the exploratory data analysis method using hierarchical clustering and non-parametric correlation heatmapping, as well as suggesting that the concepts evoked by the literature and theory have substantive meaning. It also demonstrates that these concepts that are difficult to operationalise can be measured more appropriately using factor analysis as opposed to relying on single-indicator approximations or uncritical aggregations which may attribute too much weight to survey items that are capturing variance from some other unrelated concepts.

The real contribution to the theoretical understanding of multidimensional poverty came from the construction of the higher-order multidimensional poverty factor. The primary iteration of this model on an exploratory subset of data showed that some factors that might have been expected to be dimension of poverty, in particular the idea of social isolation (Eckhard, 2018), when measured as related to the quality of a person's interpersonal relationships, actually had no association with the underlying construct. In this case it is important to make the distinction between social isolation and social exclusion, especially in quantitative measures. The final measure was restricted to five dimensions which incorporated five factors: financial strain, life satisfaction, material deprivation of consumables, psychosocial strain, and material deprivation of commodities. Financial strain and life satisfaction had the highest factor loadings, representing central material and psychosocial dimensions of the underlying construct. These tap into subjective realities of peoples' interconnected fiscal, material, relational, and psychological hardship. Material deprivation of consumables, which broadly represented the households' ability to participate in society as developed by Pantazis, et al., (2006) and earlier by Townsend (1979, 1985; Walker, et al., 2010), represents the strongest external indicator of multidimensional poverty. Many of the items are objective indicators, but reflect relative and subjective, observations of what constitutes participation in UK society. Psychosocial strain can be imagined to be the manifestations of low life satisfaction and material deprivation in a person's general psychosocial health, which increases at a considerable rate alongside dissatisfaction with life, financial strain, and material deprivation. Material deprivation of household commodities is more weakly associated with the overall factor, but still represents an important dimension. Although this may reflect a change in the structure of multidimensional poverty since the work of Tomlinson, et al., (2008), further revision of questions related to household goods may be needed to assess more fully its relevance. Ultimately, the method demonstrates that it is possible to construct a measure of poverty that can incorporate appropriately weighted psychosocial dimensions without sacrificing a definitive material core (Lister, 2004).

Embedded into the construction of these factors have been tests to show their levels of invariance over time. Townsend's (1979) conceptualisation of poverty identifies it as both a dynamic lived experience, a situation that people move in and out of rather than a static feature of many peoples' lives, and as a dynamic concept. The idea of a dynamic constructs has applicability in sociology more broadly. Briefly, the empirical observable indicators of such

concepts are expected to change over time as they reflect societies' changing manifestations of their underlying ontological basis. For example, material deprivation of commodities may have not included access to a personal computer and the internet in the mid-1990s, but due to the wide roll-out of vital services in the 2010s on a digital platform, including services for welfare benefit and job applications, access to these facilities has become absolutely essential.

However, instantly problematising these indicators ignores the fact that they can be relatively stable for set periods of time. In some ways, certain metrics like the Indices of Multiple Deprivation (DCLG, 2015) have reacted to this dynamic nature drastically, by employing a methodology that, by design, cannot be used to assess changes in poverty over time, under the assumption that relative deprivation is so context specific that its levels cannot be compared. However, this decision has no strong empirical basis due to the lack of invariance testing, and greatly limits how we can study and understand poverty and its effects on individuals and places over time.

Invariance testing is not designed explicitly with dynamic concepts in mind, and therefore a strategy of omnibus tests needed to be developed and employed to explore the extent to which the derived latent variable remained stable over time. This use of invariance testing showed that, at least for this multidimensional poverty measure, there was a strong case that poverty could be measured consistently well using a singular measurement developed and tested on a representative sample of the UK population between the years of 2009 and 2016. This is an important finding that helps to answer the question of whether a multidimensional metric that is contextually relative can be meaningfully compared between years. In the same way that income-based measures are able to be compared across years by adjusting for inflation and family size, the multidimensional poverty measure can be compared across years by imposing constraints on the various statistical parameters using these techniques. Since change over time and life course dynamics are critical for understanding poverty and its impact on peoples' lives the importance of employing a method that shows highly contextual and relative measures but can be used over shifting circumstances as society changes, even if this is for short periods of time before the measure needs to be 'reimagined', should not be understated. The same methodology might also be employed to prevent unnecessary and potentially wasteful 'reimaginings' of poverty that, as Piachaud (1987) warns, may do nothing to help the poor.

7.3 Group Invariance and Social Divisions

Similarly, any proposed measure of poverty needs to be applicable across a range of people based on social divisions and living arrangements in society. This is done to some extent in existing income-based poverty metrics, such as the HBAI measure. However, because of this need to ensure equivalence between different households by adjusting for economies of scale, income-based metrics are extremely sensitive to the choice of equivalence scale methodology used (Spicker, et al., 2007; Deaton, 2018). This method uses equivalence adjusted household income as only one of four indicators of financial strain, reducing its overall influence and error on the final measure of poverty. The use of model fit statistics to assess invariance between different groups, including a diverse range of household types, allowed for the assumption that the indicators chosen for each dimension and each dimension itself as a part of poverty are capturing approximately the same underlying construct could be explicitly tested. If the indicators used when fixed to be equal across different types of household, or across gender or cohort, resulted in substantially worse model fit statistics we could then argue that different measurements of poverty need to be derived for different groups, and that a universal measure may not be the best option. However, it was found that model fit actually improved under the assumption that all groups assign the same meaning to each indicator and lower-order factor.

This means that it was possible to show that a multidimensional poverty measure derived in this way could be reliably applied to any type of household equally. This is often implicitly assumed when using measures of poverty, but the ability to conclude this is important to justifying the integrity of a multidimensional measure. As a consequence it can be argued that there is good evidence that it is possible to derive a meaningful, theoretically-informed, reliable and valid multidimensional measure of poverty. However, it is still worthwhile addressing the fact that there are important divisions in society, specific social groups at greater risk of poverty, for whom it was not possible to assess the reliability of the measure. These included social groupings of people by geography, ethnic group, and disability, as well as the intersections between these and other social divisions. More data with a far higher sample size of these marginal groups is needed to be confident in the relevance of this measure for them.

7.4 Introducing a multidimensional measure to the policy sphere: some challenges

An astute observation may be that thus far in the thesis there has been no effort made to apply this measure, either in terms of exploring changes in poverty over time, looking at the relationship between mean levels of poverty on the index and different social groups, or looking at the relationship between multidimensional poverty and other factors. This has been

a deliberate choice. Earlier in this section the idea that different poverty metrics can tell different stories was raised, and this was visualised in figure 14. There are concerns that the validity of a particular metric is judged tacitly based on the reviewer's subjective political viewpoint and experience and whether the measure shows what they believe a measure of poverty should be showing (Atkinson, 1985). For example, someone who believes poverty has been increasing at a dramatic rate since 2008 may be more predisposed to accept the JRF's MIS as the most valid measure of poverty. Someone who feels like poverty has been largely stable over time but still remains quite high will probably be predisposed to putting faith in the HBAI measure after housing costs. Someone who has a preexisting feeling that poverty has been decreasing continuously over time might implicitly prefer the absolute HBAI before housing costs measure. By avoiding applying trend analysis to this measure, or trying to find a cut off to estimate a level of poverty, the measure can be judged firstly on the strength of the argument for the dimensions used and, secondly, on the statsitical foundations that provide merit for its relevance, reliability, and validity independent of the story it tells.

To move on from reliability and validity and address the measure's substantive theoretical basis, it is clearly a departure from competing absolute and relative approaches to the measurement of poverty using income alone. In using income, a metric becomes relative or absolute based on whether it is calculated using a function of inequality, in the former, or not, in the latter. This is regardless of how this function is derived: whether it be through a consensual method to establish a minimum threshold as a percentage of income (Walker, 1987; Padley, et al., 2017) in the former, or through the use of an 'objective' basket of goods in the latter (Anand, et al., 2010). By contrast, the relative and absolute natures of indicators in this multidimensional measure are dictated by the design of the questionnaire items, not by the dividing up of the distribution of responses to create a 'line'.

The need to feel satisfaction in one's life or to be able to pay one's bills can be described as absolute and universal, but using this method does not mean that the conclusions drawn have to be based on the distribution of how much more satisfied or able to pay bills others are. It is the respondents choice as to whether they interpret the criteria as absolute or relative. In other words, specific poverty lines can be drawn using either approach from the same measure, but the interpretation of the indicators is based on respondents' understandings which are likely to be relative. An absolute moral minimum can be taken from indicative responses to specific questions that indicate unacceptable hardship and then inverted as an expected score on the

measure, or a line can be drawn at a certain relative point in the distribution for a certain year, as in Tomlinson, et al. (2008). These approaches could even be combined in creative ways to result in a measure that assesses whether someone is either in 'absolute' poverty by some specific criteria, such as income, or whether they are at the equivalent level of that absolute criteria in other dimensions in relative terms, even when they may not necessarily meet the absolute criteria within the specific dimension itself. This brings the potential for greater understanding of the economisation of poverty (Townsend, 1985). However, scholars must be critical of imposing any such binary classification in the first place.

We have a good fitting, time-invariant, theoretically informed way of measuring poverty: but that does not mean that it is an acceptable way of measuring poverty. The approach creates a measurement of poverty that shows a continuum (Townsend, 1985), that combines the multidimensional aspects developed by Baulch (1996), Lister (2004; 2015), and Spicker (2007). This creates an opportunity for more theoretically meaningful representations of poverty that might be more appropriate for assessing research and policy questions that further our understanding. This might, for example, be used for testing theories of Townsend's (1985) participation breakpoint (Ferragina, et al., 2013). However, this method is exceedingly complex and specialist. It can be difficult to interpret and produce, and the quality of the data and measurement tools used is not quite high enough to be unproblematic. One of the strengths of existing poverty measures is in their simplicity: they give a clear figure, with clear criteria. These kinds of measures are attractive to both policy makers and activists. But what about the unknown risks of this kind of simplification?

In contrast, this continuum does not diminish the multiplicity of experiences that people in poverty face as much as conventional income based measures. It does not sidestep the issues they cause for methodological convenience. Yes, it is more technically difficult to model. It is also harder to interpret directly. But are those good enough reasons to abandon adding complexity to our metrics? The trade off for this complexity is that we can have greater confidence in our knowledge of poverty, because it is possible to demonstrate the measure's validity and reliability. The arguments for the impact and extent of poverty can be made far more strongly due to this. People who have lived or are living in poverty are currently fighting to have their voices heard and their complex experiences acknowledged by those with the power to do something about the problem. If quantitative social scientists cannot support this

effort with equal vigor we may end up detracting from, rather than enchancing, their argument (Piachaud, 1987).

7.5 Methodological reflections

We find similar points of complexity when considering the strengths and weaknesses of the methodological approach. As the model grew in complexity, with the introduction of higher order factors, the number of technical difficulties increased and the available, accessible resources for the applied user to draw upon to resolve them became more and more scarce. Simply put: as we build up theoretical complexity, things begin to fall apart technically. This reflects an underlying tension between ideas and measures, between ontology and methodology, and the difficult process of operationalising concepts. We can only measure complex ideas if our tools are sufficiently complex. Any compromise on the complexity of our measure is also a compromise on our representation of the concept. They are two sides of the same coin. Currently, the decision to compromise is based largely on considerations of practicality and transferral of ideas: what is achievable and what is easily interpretable. However, is this right? Should the only measures that are used be ones that anyone can assess easily? Working with and communicating the results of advanced quantitative methods is hard and can be frustrating. These factors almost invariably drive measures towards simplicity, and limit their capacity to reflect the ontological reality of what they aim to measure. Should policy be based on measures that are restricted to one, or a small number of, narrowly defined dimensions of a concept?

So, what is the solution? It's hard to argue that compromises in measurement complexity can be justified by time or resources when such decisions may result in misrepresenting what they claim to be showing. The complexity and nuance that this research has shown is possible is sacrificed readily for convenience and ease of use and interpretation, or for the sake of continuity of existing processes. Following the arguments of Tomlinson, et al., (2008) and Lister (2015), I instead propose that a social justice perspective should be applied to the construction of metrics. Instead of metrics being assessed on their policy impact, interpretability, or convenience, they should primarily be assessed on the extent to which they adequately reflect the lived experience of what is being measured. Do they allow for the necessary nuance? Can they be applied fairly to all, especially the most vulnerable and unheard in society? Any compromises made should be forced to withstand this kind of ethical scrutiny – have the compromises themselves pushed a measure so far from the lived reality of people concerned that they now 'fail to do justice' (Tomlinson, et al., 2008: 600) to those whose

experiences it is supposed to be representing? Can any simplification of the measurement be morally justified? This, not efficiency, could be the driving force behind any measure, and we should always be striving to ensure that future iterations of our operationalisations better describe the conditions they purport to describe. If a measure no longer works well, and this thesis demonstrates a way of assessing how well it works, it should be reassessed. Currently there are few statistical checks and balances on how well poverty measures work, leading to a cacophony of competing approaches with little hard evidence to inform decisions between them.

There is an additional methodological concern that the integration of logistic and probit regression of categorical and ordinal data that are specifically needed for developing such methods within a SEM framework result in artificially high fit indices, or fit indices that do not perform as expected, on complex and large sample models. The fit statistics for many dimensions were exceptionally high, and it would be valuable to ensure that this is really the case and does not lead to a false sense of confidence around the metric. Methodological developments to deal with these problems have been germinating in quantitative psychology (for example, see the special issue of Frontiers in Psychology edited by Rens Van De Schoot, et al., 2015), but these are yet to be implemented widely in statistical software. Nonetheless, the tools that are available now have demonstrated a way to assess the quality of poverty metrics beyond face validity checking. For this to become common and develop further in sociology, the effort needs to begin somewhere.

The development of these metrics can be arduous and such invariance findings might seem trite on first reading. After all, most measures already implicitly operate under the assumption they can be applied equally regardless of time or group circumstances without undergoing any testing. Nor do the findings suggest an exciting or radical reformulation of how to measure poverty based on social divisions. The measure largely fits with what we would have expected from theoretical contributions over the years. However, this ignores the fact that what should be a procedure applied by rote in the development of all measures, and which is done so in fields such as psychology, is almost completely passed over and practically non-existent in the field of sociological research, even though much of what we are interested in can be approximated by latent constructs. In fact, it is so unusual that its use can be described as novel. Without using these methods, we cannot be sure that measures developed using a similar methodology – let alone measures that do not use factor analysis and therefore cannot even be

tested for invariance – have the same meanings for different groups of people, or indeed as in time-invariance, at different points in time.

The consequence of these methodological checks is that, when applied, there is no ambiguity as to whether differences in levels of the rate and depth of poverty in any dimension - experienced by women, or older people, or younger people, or retired households, or lone parents, or in different years - actually represents differences in taste, preference, or interpretation of questions. Different levels of poverty mean different levels of poverty, nothing else. We can be confident in that. If it is our ethical responsibility as researchers of poverty to represent the lived experiences of people living in poverty with the strongest evidence, in a way that does justice to its multidimensional nature, we need this level of robustness and confidence in our measures. With this confidence, we can make stronger arguments on behalf of those living in poverty and avoid dragging them into 'semantic and statistical squabbles' (Piachaud, 1987). Quantitative measures will always be a feature of the wider policy and public narrative of poverty, no matter how much the voices of the poor are amplified. We need – and people deserve – measures that are both conceptually strong in their theoretical foundations, grounded in lived experiences, and technically strong in their statistical rigour, with assumptions tested explicitly and never assumed. The measure proposed here has its own problems, but I believe it is a step forward in this direction.

7.6 Summary: Implications for policy and future research

This thesis has provided evidence to suggest that it is possible to construct a reliable and valid measure of multidimensional poverty using sufficiently complex quantitative techniques and providing high quality, large sample survey data is available. As it stands, there is a heavy reliance on imputation of missing data, as the quality of survey data is somewhat lacking. However, the need for such a measure becomes apparent when considering, side-by-side, the disconnects between theories of poverty and measures of poverty. In many cases, discussions of poverty become co-opted by discussions about measures (Lister, 2004), for which there is no strong methodological or theoretical justification. There is an emerging impasse in the measurement of poverty, where several proposed measures have emerged which sometimes tell contradictory stories about the extent and depth of poverty. None of these measures in their current form can be assessed for their robustness, reliability, and validity using statistical means,

and all overemphasize the material dimensions of poverty due to their focus on income and consumption.

This is flawed in two major ways beyond their ignorance of nonmaterial dimensions. Firstly, under a definition of poverty as being a dynamic concept, one that changes relative to society and between different groups, there are no clear-cut ways to determine when measurement should be changed to reflect a shifting definition. Either changes in the definition and measurement have to be made arbitrarily through fixed periods of reassessment, such as in the MIS, or changes are simply not made and an assumption is made that the measurement still fits the definition equally well, which does not fit with theory. Secondly, because such competing income-based measures cannot be assessed for their statistical reliability and validity, they are largely judged based on their face validity: how well they are imagined to represent what they propose to represent. Because poverty itself is such a political concept, this undoubtedly results in certain people and policymakers favouring some measurements over others because they show results that fit within their worldview (Atkinson, 1985).

There are some good reasons for these simplified measures; they allow for easy quantification and interpretation of poverty. However, when these oversimplified measures begin to co-opt the definition of poverty and become used as meaningful measures of the concept in academic research, there is a risk that they are not truly representing what it means to be poor to different extents, and therefore not truly doing justice to nor representing the experience of those who have lived with poverty. What I mean to say is, even if we agree that oversimplified measures are acceptable for policy purposes, does that mean we should also agree that they are suitable for quantitative research into, say, the relationship between poverty and health, especially when alternative multidimensional poverty scales can be constructed?

The question of measurement based on a theoretical conceptualisation of poverty is one which naturally leads to a factor analysis solution, it being the standard way that concepts that are not possible to measure directly can be constructed as ‘latent observations’ based on multiple indicators. Such approaches are commonplace in other fields, such as psychology and political science, but are less frequently used in sociology, despite the fact that many sociological concepts are impossible to measure using a single indicator. Quantitatively measured sociological constructs are therefore absent of a range of statistical assessments for their quality, including tests for their stability over time or comparative use between groups.

Introducing such methods neatly solves the two major problems in poverty measurement that were identified from the review of existing approaches. By employing these, the previous section has shown that a multidimensional measure can be constructed and offers a validation of existing theoretical conceptualisations. Furthermore, the measure can be tested to explore whether its suitability is a byproduct of existing randomness in the data, or whether it reflects a true underlying concept. The use of exploratory data analysis using hierarchically clustered correlation plots, in addition to the use of construction and validation subsets and multiple imputations of missing data, shows that there is evidence for the existence of an underlying construct of poverty that contains material, social, and psychological dimensions. The measure has then been tested to explore how dynamic or stable the construct is over time, and it was found to be stable for all six time points that it was measured across. Furthermore, it was directly possible to address the question of whether the underlying structure of multidimensional poverty differed significantly depending on household type, gender, and cohort membership. This particular construct was found to be stable across all of these groups.

Not only does this demonstrate that a measure can be developed which can be explicitly tested for its arbitrariness or irrationality (Townsend, 2002), but also allows the development of new methodological approaches to explore exactly when such a measure *does* become arbitrary or irrational. This approach would break the divide between one method, that assumes poverty is so dynamic that it cannot be compared at different time points or between different groups or needs to be reassessed with arbitrary frequency, and the other, that the same measure can be for an indeterminate amount of time and mean the same thing throughout. Poverty *can* be measured in a multidimensional way and a measure *can* be compared across certain points in time and between certain groups within a framework that is sensitive to points at which the underlying dynamic concept changes so substantially that it needs to be reassessed. This method makes such changes known.

That is not to say that there are not substantial problems with this approach. There is a high degree of complexity involved in constructing the measure in the first place, which requires complex survey data with a great number of variables and very large samples. Often this data also has a high degree of missingness which can introduce further difficulties. Certain tests were not possible to conduct either. Certain dimensions from theory were not possible to observe or model. The measure could not be tested to be reliable for different ethnic groups

and people living with disabilities. The final model seemed to show some instability, the cause of which could not be identified. The aim of extending the factor analysis to a multilevel factor analysis was not yet possible due to the computational complexity involved. However, the outcome of this research may demonstrate that pursuing such an approach to measurement is possible, valuable and, ideally, might lead to the adoption of a standardised set of multidimensional poverty indicator items that can be used to construct a valid and reliable scale in much the same way that new psychometric measures become adopted.

There is some potential for such an adoption of this scale to contribute to our understanding of poverty. While existing research has mapped the dimensions and thematic elements of poverty, there is still much that is not known about how these dimensions relate to one another or broader societal changes. This research acts as a preliminary stage at which to unpick some of these relationships. Most importantly, I believe what is observed reflects that multidimensional poverty can be imagined to have two 'hubs' or 'cores' – one for material dimensions and one for social and emotional dimensions. These twin cores refer to financial strain and life satisfaction respectively. They both reflect the most universal dimensions of multidimensional poverty, reflected in equal measure across the entire continuum that forms the construct. Financial strain consists of being unable to reliably make payments for housing, council tax, and energy, in a given year, and also captures low income. Life satisfaction, on the other hand, consists of a holistic measure reflecting a person's happiness across a number of factors: leisure, income, health, and overall.

From this approach emerges the possibility that we can explore how poverty in its initial stages manifests as financial strain and a loss of self-esteem and self-actualisation. Such a proposition is in line with theories of poverty and a joined conceptualisation of Baulch's (1996) pyramid and Maslow's (1943) hierarchy of needs, and does not prioritise material needs over psychosocial emotional needs, or vice-versa, but considers them together. This demonstrates a way to incorporate what Walker (2014) and colleagues have identified as a universal dimension of poverty – shame - into a quantitative metric. With better data it would be possible to demonstrate how experiencing these twin cores of poverty might later manifest in a reduction in participation in society as well as worse psychosocial health, which offers a range of opportunities for social policy. It also allows a quantitative tether on reimaginings – or at least, remeasurement - of poverty, by providing a methodological framework with which to assess the salience of different dimensions of poverty as lives and societies change. It can, in theory,

keep in check the tendencies to create new measures for the sake of creating new measures, something which may well undermine the very concept of relative poverty in the developed world. The approach also creates a more symbiotic companion to qualitative research on poverty. Quantitative methods have thusfar been inadequate for incorporating findings from qualitative research; quantitative findings have been too bulwarked behind a narrow definition of poverty to provide much insight into theory development beyond describing very general patterns. Since qualitative research draws so heavily on lived experience, it is a social responsibility for quantitative researchers to make efforts to close this gap.

A manifesto for quantitative researchers wanting to do so may begin with a commitment to developing, jointly with people with lived experience and expertise, a standardised survey instrument for gauging multidimensional poverty that can be validated using the validity and reliability testing methods used throughout this thesis. Moreover, any such set of questions must assess the hitherto unmeasured dimensions of poverty, outlined in section 5.1.12. This includes capturing more directly experiences of dehumanisation, the denial of rights, time poverty, and social disabling of people living with impairments. Thirdly, such a design must move beyong a fixation on the household and be focused foremost on the individual. While the experience of poverty at the household level is important, to aggregate up to this level by default, and subsequently lose much information on the differences between people within the household, is irresponsible as it masks intra-household inequalities. Lastly, a commitment must be made by larger longitudinal surveys to adopt such a set of items for measuring multidimensional poverty, and to administer them regularly.

Ultimately, this thesis presents few surprises with regards to changing theory. The quantitative evidence supports the various conceptualisations of poverty made by scholars over the years. Importantly, it also appears to represent more of the dimensions that feature in accounts from those with lived experience than existing quantitative operationalisations do. I believe this demonstrates an approach to measurement that does more justice to the underlying concept of poverty, and therefore better respects the reality faced by those living in its grasp. This measure has been deliberately presented without its implementation: the methodological approach to its justification should be the primary means by which its quality is assessed, not what it seems to be suggesting about trends in the extent or depth of poverty, where it can be easily dismissed out of hand (Atkinson, 1985). The difficulties in implementation and interpretation are the cost of using such an approach, the value of which needs to be seriously assessed. It raises a question

of whether such complicated approaches can play a meaningful role in policy, or whether they should be restricted to purely sociological inquiry. If so, would this divergence truly be such a bad thing and, if not, should we be so skeptical that a complex measure explained in understandable terms would not be preferable? As Atkinson (1985: 16-17) argues, such an approach can lead to 'less definite answers, but, one hopes, commands a wider degree of support'. My personal hope is that such evidence may lead to a step away from what David Piachaud (1987) identifies as a situation of semantic and statistical arguments that do little for those in poverty, and towards a measure that does justice to the experiences of those living in, or who have lived in, poverty.

Bibliography

Abel-Smith, B. & Townsend, P., 2005. The poor and the poorest. In: D. Banks & M. Purdy, eds. *The Sociology and Politics of Health: A Reader.* London: Routledge, pp. 66-70.

Acock, A. C., 2005. Working with Missing Values. *Journal of Marriage and Family,* 67(4), pp. 1012-1028.

Adams, N. et al., 2012. *Households below average income: an analysis of the income distribution 1994/95 - 2010/11,* London: Department for Work & Pensions.

Afonso, P., Fonseca, M. & Pires, J. F., 2017. Impact of working hours on sleep and mental health. *Occupational Medicine,* 67(5), pp. 377-382.

Alcock, P., 2006. *Understanding Poverty.* Basingstoke: Palgrave Macmillan.

Alcock, P., May, M. & Wright, S., 2012. *The Student's Companion to Social Policy (4th ed.).* London: Wiley-Blackwell.

Aldridge, H. & Hughes, C., 2016. *Informal Carers & Poverty in the UK: An analysis of the Family Resources Survey,* York: Joseph Rowntree Foundation.

Allison, P. D., 2001. *Missing Data.* Thousand Oaks, CA: Sage.

Anand, S., Segal, P. & Stiglitz, J., 2010. *Debates on the Measurement of Global Poverty.* Oxford: Oxford University Press.

Anyaegbu, G., 2010. Using the OECD equivalence scale in taxes and benefits analysis. *Economic & Labour Market Review,* 4(1), pp. 49-54.

Aspinall, P., 2000. The challenges of measuring the ethno-cultural diversity of Britain in the new millennium. *Policy & Politics,* 28(1), pp. 109-118.

ATD Fourth World, N.D.. *Understanding Poverty in All its Forms.* [Online] Available at: https://atd-uk.org/projects-campaigns/understanding-poverty/ [Accessed 1 Jan 2019].

Atkinson, A. B., 1985. *How Should We Measure Poverty? Some Conceptual Issues.* ESRC Programme on Taxation, Incentives and the Distribution of Income, Discussion Paper No. 82.

Azur, M. J., Stuart, E. A., Frangakis, C. & Leaf, P. J., 2011. Multiple Imputation by Chained Equations: What is it and how does it work?. *International Journal of Methods in Psychiatric Research,* 20(1), pp. 40-49.

Bailey, N., Bramley, G. & Hastings, A., 2015. Symposium Introduction: Local Responses to 'Austerity'. *Local Government Studies,* 41(4), pp. 571-581.

Bar-Tal, D., Raviv, A. & Raviv, A., 1991. The Concept of Epistemic Authority and the Process of Political Knowledge Acquisition: The effect of similarity. *Representative Research in Social Psychology,* 19(2), pp. 119-138.

Barnardo's, 2012. *Paying to Work: Childcare and Child Poverty,* Ilford: Barnardo's.

Barnes, M., Cullinane, C., Scott, S. & Silvester, H., 2013. *People living in bad housing – numbers and health impacts,* London: NatCen Social Research.

Bartlett, A., Frew, C. & Gilroy, J., 2013. *Understanding Material Deprivation Among Older People,* London: Department for Work and Pensions.

Baulch, B., 1996. The New Poverty Agenda: A Disputed Consensus. *IDS Bulletin,* 27(1), pp. 1-10.

Baumberg, B., 2016. The stigma of claiming benefits: a quantitative study. *Journal of Social Policy,* 45(2), pp. 181-189.

BBC, 2018. *Poverty causing 'misery' in UK, and ministers are in denial, says UN official.* [Online] Available at: https://www.bbc.co.uk/news/uk-46236642 [Accessed 26 Dec 2018].

Beaujean, A. A., 2014. *Latent Variable Modeling Using R.* Hove: Routledge.

Becker, G., 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education (3rd Ed.).* London: University of Chicago Press.

Becker, S., Bryman, A. & Ferguson, H., 2012. *Understanding Research for Social Policy and Social Work: Themes, Methods and Approaches.* Bristol: Policy Press.

Béland, D., 2005. Ideas and Social Policy: An institutionalist perspective. *Social Policy & Administration,* 39(1), pp. 1-18.

Bell, A. & Jones, K., 2014. Current practice in the modelling of age, period and cohort effects with panel data: a commentary on Tawfik et al. (2012), Clarke et al. (2009), and McCulloch (2012). *Quality & Quantity,* 48(4), p. 2089–2095.

Bell, A. J. & Jones, K., 2013. The impossibility of separating age, period and cohort effects. *Social Sciences and Medicine,* Volume 93, pp. 163-165.

Bell, A., Johnston, R. & Jones, K., 2015. Stylised fact or situated messiness? The diverse effects of increasing debt on national economic growth. *Journal of Economic Geography,* 15(2), pp. 449-472.

Bell, A., Jones, K. & Fairbrother, M., 2017. Understanding and misunderstanding group mean centering: a commentary on Kelley et al.'s dangerous practice. *Quality & Quantity.*

Bennett, F. & Daly, M., 2014. *Poverty through a Gender Lens: Evidence and Policy Review on Gender and Poverty,* Oxford: Department of Social Policy and Intervention, University of Oxford.

Bennett, F., 2008. Distribution within the household. In: *Measuring Poverty: Seven key issues, Briefing Paper.* Colchester: Institute of Social and Economic Research, University of Essex.

Benzeval, M. et al., 2014. *How does money influence health?,* York: Joseph Rowntree Foundation.

Beresford, P., 1996. Poverty and Disabled People: Challenging dominant debates and policies. *Disability and Society,* 11(4), pp. 553-568.

Beresford, P., Green, D., Lister, R. & Woodward, K., 1999. *Poverty First Hand: Poor People Speak for Themselves.* London: Child Poverty Action Group.

Best, J., 2012. *Damned Lies and Statistics: Untangling numbers from the Media, Politicians and Activists.* London: University of California Press.

Bishop, J. A. & Rodríguez, J. G., 2018. *Inequality, Taxation, and Intergenerational Transmission.* Bingley: Emerald Publishing.

Bollen, K. A. & Long, J. S., 1993. *Testing Structural Equation Models.* London: Sage.

Bourdieu, P., 1986. The Forms of Capital. In: J. Richardson & R. Nice, eds. *Handbook of Theory and Research for the Sociology of Education.* New York: Greenwood, pp. 46-58.

Bradshaw, J. & Finch, N., 2003. Overlaps in Dimensions of Poverty. *Journal of Social Policy,* 32(4), pp. 513-525.

Bradshaw, J. et al., 2003. *Gender and Poverty in Britain,* Manchester: Equal Opportunities Commission Working Paper Series 6.

Bradshaw, J. et al., 2008. *A minimum income standard for Britain: What people think,* York: Joseph Rowntree Foundation.

Bradshaw, J., 2016. *The Well-Being of Children in the UK.* Bristol: Policy Press.

Brady, H. E. & Collier, D., 2010. *Rethinking Social Inquiry.* Lanham, MD: Rowman and Littlefield.

Brisbane, A., 2012 [1840]. *Social Destiny of Man, Or, Association and Reorganization of Industry.* London: Forgotten Books.

Brown, T. A., 2015. *Confirmatory Factor Analysis for Applied Research.* London: Guilford.

Bryman, A., 2008. *Social Research Methods.* Oxford: Oxford University Press.

Burchardt, T. & Karagiannaki, E., 2018. *Intra-household inequality and material deprivation and poverty in Europe.* Brussels: Social Situation Monitor Seminar 'Multidimensional Poverty in the EU'.

Burchardt, T., 2008. *Time and income poverty.* York: Joseph Rowntree Foundation.

Burton, J., Nandi, A. & Platt, L., 2008. *Who are the UK's minority ethnic groups? Issues of identification and measurement in a longitudinal study.* Essex: Insitute for Social and Economic Research.

Burton, P., Phipps, S. & Woolley, F., 2007. Inequality within the household reconsidered. In: *Inequality and Poverty Re-examined.* Oxford: Oxford University Press, pp. 103-127.

Byrne, B., 2009. *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming.* London: Routledge.

Cade, J., Upmeier, H., Calvert, C. & Greenwood, D., 1999. Costs of a healthy diet: analysis from the UK women's cohort study. *Public Health Nutrition,* 2(4), pp. 505-512.

Calandrino, M., 2003. *Low-Income and Deprivation in British Families,* London: Department for Work and Pensions Working Paper 10.

Cambell, D. T. & Fiske, D. W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* Volume 56, pp. 81-105.

Carmines, E. & Zeller, R., 1979. *Reliability and Validity Assessment.* London: Sage.

Carter, J., 1998. *Postmodernity and the fragmentation of welfare.* London: Routledge.

Carter, M. R. & Barrett, C. B., 2006. The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies,* 42(2), pp. 178-199.

Castell, S. & Thompson, J., 2007. *Understanding attitudes to poverty in the UK: Getting the public's attention,* York: Joseph Rowntree Foundation.

Centre for Research in Social Policy, N.D.. *What is MIS?.* [Online] Available at: http://www.lboro.ac.uk/research/crsp/mis/whatismis/ [Accessed 22 May 2016].

Cernat, A., Couper, M. P. & Ofstedal, M. B., 2016. Estimation of Mode Effects in the Health and Retirement Study Using Measurement Models. *Journal of Survey Statistics and Methodology,* 4(4), pp. 501-524.

Chatzitheochari, S. & Arber, S., 2012. Class, gender and time poverty: a time-use analysis of British workers' free time resources. *British Journal of Sociology,* 63(3), pp. 451-471.

Cheng, H. G. & Phillips, M. R., 2014. Secondary analysis of existing data: opportunities and implementation. *Shanghai Archives of Psychiatry,* 26(6), pp. 371-375.

Cohen, J., 1960. A coefficient for agreement for nominal scales. *Education and Psychological Measurement,* Volume 20, pp. 37-46.

Collins, L. M., Schafer, J. L. & Kam, C. M., 2001. A comparison of inclusive and restrictive strategies in modern missing-data procedures. *Psychological Methods,* Volume 6, pp. 330-351.

Connelly, R., Gayle, V. & Lambert, P., 2016. Ethnicity and ethnic group measures in social survey research. *Methodological Innovations,* 9, pp. 1-10.

Cooper, N., Purcell, S. & Jackson, R., 2014. *Below the breadline: the relentless rise in food poverty in Britain,* Oxford: Church Action on Poverty/Oxfam/Trussell Trust.

Corsaro, W. A., 2017. *The Sociology of Childhood.* London: Sage.

Cramer, D. & Howitt, D., 2004. *The SAGE Dictionary of Statistics.* London: SAGE.

Crisp, R., Eadson, E. & While, A., 2016. *Tackling poverty through housing and planning policy in city regions,* York: Joseph Rowntree Foundation.

Cronbach, L. J., 1951. Coefficient alpha and the internal structure of tests. *Psychometrika,* 16(3), pp. 297-334.

Crouch, C., 2016. *Society and Social Change in 21st Century Europe.* London: Palgrave Macmillan.

Daly, M., 1992. Europe's poor women? Gender in research on poverty. *European Sociological Review,* 8(1), pp. 1-12.

Davies, K., 2011. Knocking on doors: recruitment and enrichment in a qualitative interview-based study. *International Journal of Social Research Methodology,* 14(4), pp. 289-300.

Davies, S., Finney, A. & Hartfree, Y., 2016. *Paying to be Poor: Uncovering the scale and nature of the poverty premium,* Bristol: University of Bristol Personal Finance Research Centre.

Davis, A., Hirsch, D. & Padley, M., 2014. *A Minimum Income Standard for the UK in 2014,* York: Joseph Rowntree Foundation.

de Souza, A. C., Alexandre, N. M. C. & de Brito Guirardello, E., 2017. Psychometric properties in instruments evaluation of reliability and validity. *Applications in Epidemiology,* 26(3).

Deaton, A., 2018. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Washington, D. C.: The World Bank.

DEFRA, 2014. *Family Food 2013,* London: National Statistics.

Department for Communities and Local Government, 2015. *English indices of deprivation 2015.* [Online] Available at: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 [Accessed 17 Mar 2016].

Department for Work & Pensions, 2016. *Households Below Average Income: An analysis of the UK income distribution: 1994/95-2014/15,* London: DWP.

Department for Work & Pensions, 2016. *Income-Related Benefits: Estimates of Take-up.* [Online] Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535362/ir-benefits-take-up-main-report-2014-15.pdf [Accessed 1 August 2016].

DeVon, H. et al., 2007. A Psychometric Toolbox for Testing Validity and Reliability. *Journal of Nursing Scholarship,* 39(2), pp. 155-164.

Dijkstra, A., Smith, J. & White, M., 2006. *Measuring care dependency with the Care Dependency Scale (CDS): A manual.* Groningen: Eurecare.

Dominy, N. & Kempson, E., 2006. *Understanding older people's experiences of poverty and material deprivation (Research Report No 363).* Norwich: Department for Work and Pensions.

Dong, Y. & Peng, C. J., 2013. Principled missing data methods for researchers. *SpringerPlus,* 2(222).

Dowler, E. & Lambie-Mumford, H., 2015. How can households eat in austerity? Challenges for Social Policy in the UK. *Social Policy & Society,* 14(3), pp. 417-428.

Doyal, L. & Gough, I., 1984. A Theory of Human Needs. *Critical Social Policy,* 4(10), pp. 6-38.

Eckhard, J., 2018. Does Poverty Increase the Risk of Social Isolation? Insights Based on Panel Data from Germany. *The Sociological Quarterly,* 59(2), pp. 338-359.

Eidlin, B., 2014. Class Formation and Class Identity: Birth, Death, and Possibilities for Renewal. *Sociology Compass,* 8(8), pp. 1045-1062.

Enders, C. & Bandalos, D. L., 2001. The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling,* 8(3), pp. 430-457.

Evans, G. W. & English, K., 2003. The Environment of Poverty: Multiple Stressor Exposure, Psychophysiological Stress, and Socioemotional Adjustment. *Child Development,* 73(4), pp. 1238-1248.

Everitt, B., 1974. *Cluster Analysis.* London: Heinemann.

Ferdman, R. A., 2014. *The slow death of the microwave.* [Online] Available at: https://qz.com/187743/the-slow-death-of-the-microwave/ [Accessed 18 Jan 2018].

Ferragina, E., Tomlinson, M. & Walker, R., 2013. *Poverty, Participation & Choice: The legacy of Peter Townsend,* York: Joseph Rowntree Foundation.

Ferrer-i-Carbonell, A. & Ramos, X., 2012. *Inequality and happiness: a survey (Discussion Paper 38).* Amsterdam: GINI Project (European Commission).

Field, A., 2013. *Discovering Statistics Using IBM SPSS Statistics.* London: Sage.

Field, F., 2010. *The Foundation Years: preventing poor children becoming poor adults,* London: HM Government.

Finch, N. & Groves, D., 1983. *A Labour of Love: Women, Work and Caring.* London: Routledge.

Fitzpatrick, S. et al., 2016. *Destitution in the UK,* York: Joseph Rowntree Foundation.

Flaherty, J., 2008. *"I mean we're not the richest but we're not poor": discourses of 'poverty' and 'social exclusion',* Loughborough: Loughborough University.

Fornell, C. & Larcker, D. F., 1981. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research,* 18(1), pp. 39-50.

Foster, L., Tomlinson, M. & Walker, A., 2018. Older people and Social Quality – what difference does income make? *Ageing and Society,* pp. 1-26.

Foster, R. & Lunn, J., 2007. Briefing Paper: 40th Anniversary Briefing Paper: Food availability and our changing diet. *British Nutrition Foundation Nutrition Bulletin,* Volume 32, pp. 187-249.

Foucault, M., 2001 [1970, 1966]. *The Order of Things: Archaeology of the Human Sciences.* Oxford: Routledge.

Fraser, N., 1997. *Justice Interruptus.* London: Routledge.

Friedli, L. & Stearn, R., 2015. Positive affect as coercive strategy: conditionality, activation and the role of psychology in UK government workfare programmes. *Medical Humanities,* Volume 41, pp. 40-47.

Gaille, D., Paugam, S. & Jacobs, S., 2003. Unemployment, Poverty and Social Isolation: Is there a vicious circle of social exclusion?. *European Societies,* 5(1), pp. 1-32.

Gamble, A., 2016. *Can the Welfare State Survive?.* Oxford: Wiley.

Garthwaite, K. & Bambra, C., 2015. Food poverty, welfare reform and health inequalities. In: *In Defence of Welfare 2.* London: Policy Press, pp. 121-123.

Gelman, A. & Loken, E., 2013. *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time.* New York: Department of Statistics, Columbia University.

Gelman, A. & Loken, E., 2014. The statistical crisis in science. *American Scientist,* 102(6), p. 460.

Gelman, A., Shor, B., Bafumi, J. & Park, D., 2007. Rich state, poor state, red state, blue state: what's the matter with Connecticut?. *Quarterly Journal of Political Science,* Volume 2, pp. 345-367.

Glendinning, C. & Millar, J., 1987. *Women and Poverty in Britain.* Brighton: Wheatsheaf.

Glennerster, H., Hills, J., Piachaud, D. & Webb, J., 2004. *One hundred years of poverty and policy,* York: Joseph Rowntree Foundation.

Gnambs, T. & Staufenbiel, T., 2018. The structure of the General Health Questionnaire (GHQ-12): two meta-analytic factor analyses. *Health Psychology Review,* 12(2), pp. 179-194.

Goethals, T., De Schauwer, E. & Van Hove, G., 2015. Weaving Intersectionality into Disability Studies Research: Inclusion, Reflexivity and Anti-Essentialism. *Journal of Diversity and Gender Studies,* 2(1-2), pp. 75-94.

Goldberg, D. & Williams, P., 1988. *A User's Guide to the General Health questionnaire.* Windsor: NFER-Nelson.

Goode, J., Callender, C. & Lister, R., 1998. *Purse or wallet? Gender inequalities and income distribution within families on benefits,* London: Policy Studies Institute.

Gordon, D., 2006. The concept and measurement of poverty. In: *Poverty and Social Exclusion in Britain.* Bristol: Policy Press, pp. 29-69.

Gottschall, A. C., West, S. G. & Enders, C. K., 2012. A Comparison of Item-Level and Scale-Level Multiple Imputation for Questionnaire Batteries. *Multivariate Behavioural Research,* 47(1), pp. 1-25.

Graafland, J. & Lous, B., 2018. *Income Inequality, Life Satisfaction Inequality and Trust: A Cross Country Panel Analysis, Journal of Happiness Studies.* [Online] Available at: https://doi.org/10.1007/s10902-018-0021-0 [Accessed: 12 Dec 2018]

Graham, H., 1986. *Caring for the Family, Research Report No.1,* London: Health Education Council.

Graham, H., 1992. Budgeting for Health: Mothers in Low Income Households. In: *Women and Poverty in Britain: the 1990s.* London: Harvester Wheatsheaf.

Grant, N., Wardle, J. & Steptoe, A., 2009. The Relationship Between Life Satisfaction and Health Behavior: A Cross-cultural Analysis of Young Adults. *International Journal of Behavioral Medicine,* 16(3), pp. 259-268.

Greenland, S. et al., 2016. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology,* Volume 31, pp. 337-350.

Grimm, K. J., Ram, N. & Estabrook, R., 2017. *Growth Modeling: Structural equation and multilevel modeling approaches.* London: Guildford Press.

Grolemund, G. & Wickham, H., 2016. *R for Data Science.* Farnham: O'Reilly.

Hallquist, M. & Wiley, J., 2016. *MplusAutomation: Automating Mplus Model Estimation and Interpretation. R package version 0.6-4. https://CRAN.R-project.org/package=MplusAutomation.*

Harper, C., Marcus, R. & Moore, K., 2003. Enduring Poverty and the Conditions of Childhood: Lifecourse and Intergenerational Poverty Transmissions. *World Development,* 31(3), pp. 535-554.

Hastings, A., 2009. Poor Neighbourhoods and Poor Services: Evidence on the 'Rationing' of Environmental Service Provision to Deprived Neighbourhoods. *Urban Studies,* 46(13), pp. 2907-2927.

Heck, R. H., 2015. *An Introduction to Multilevel Modeling Techniques: MLM and SEM Approaches Using Mplus, 3rd Ed..* London: Routledge.

Herrnstein, R. J. & Murray, C., 1994. *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

Heslop, P. & Gordon, D., 2014. Trends in poverty and disadvantage among households with disabled people from 1999–2012: from exclusion to inclusion?. *Journal of Poverty and Social Justice,* 22(3), pp. 209-226.

Hewison, K., 2016. Precarious Work. In: S. Edgell, H. Gottfried & E. Granter, eds. *The SAGE Handbook of the Sociology of Work and Employment.* London: SAGE, pp. 428-443.

Hick, R., 2016. Material poverty and multiple deprivation in Britain: the distinctiveness of multidimensional assessment. *Journal of Public Policy,* 36(2), p. 277–308.

Hill, K., Davis, A., Hirsch, D. & Marshall, L., 2016. *Falling short: the experiences of families living below the Minimum Income Standard.* York: Joseph Rowntree Foundation.

Hillquit, M., 1910. *History of socialism in the United States.* New York: Funk & Wagnallis.

Hills, J., 2015. *Good times, bad times: the welfare myth of them and us.* Bristol: Polity Press.

Hills, J., Le Grand, J. & Piachaud, D., 2002. *Understanding Social Exclusion.* Oxford: Oxford University Press.

Hirsch, D., 2015. *A minimum income standard for the UK in 2015,* York: Joseph Rowntree Foundation.

Hirschfeld, R. M. A., 2001. The Comorbidity of Major Depression and Anxiety Disorders: Recognition and Management in Primary Care. *The Primary Care Companion to the Journal of Clinical Psychiatry,* 3(6), p. 244–254.

Hoaglin, D. C., Mosteller, F. & Tukey, J. W., 1983. *Understanding robust and exploratory data analysis.* New York: Wiley.

Hoff, A., 2008. *Tackling Poverty and Social Exclusion of Older People – Lessons from Europe.* Oxford: Oxford Institute of Ageing (Working Paper 308).

Hofstede, G., 1984. The cultural relativity of the quality of life concept. *Academy of Management Review,* 9(3), pp. 389-398.

Honneth, A., 1995. *The Struggle for Recognition.* Cambridge: Cambridge University Press.

Hornikx, J., 2011. Epistemic authority of professors and researchers: differential perceptions by students from two cultural-educational systems. *Social Psychology of Education,* Volume 14, pp. 169-183.

House, A. & Stark, D., 2002. Anxiety in medical patients. *British Medical Journal,* 325(7357), pp. 207-209.

Howard, M., 2001. *Paying the price: Carers, poverty, and social exclusion,* London: Child Poverty Action Group.

Humphries, M., Sanchez, R. & Windt, P., 2013. Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration. *Political Analysis,* Volume 21, pp. 1-20.

Jenkins, R., 2002. *Foundations of Sociology: Towards a better understanding of the human world.* Basingstoke: Palgrave Macmillan.

John, O. P. & Benet-Martinez, V., 2000. Measurement: Reliability, construct validation, and scale construction. In: H. T. Reis & C. M. Judd, eds. *Handbook of research methods in social psychology.* Cambridge: Cambridge University Press, pp. 339-369.

Jones, O., 2012. *Chavs: The Demonization of the Working Class.* London: Verso Press.

Jöreskog, K., 1993. Testing Structural Equation Models. In: *Testing Structural Equation Models.* London: Sage, pp. 294-317.

Joyce, P., 1995. *The Oxford Reader on Class.* Oxford: Oxford University Press.

Joyce, R., 2014. *Child poverty in Britain: recent trends and future prospects,* London: IFS Working Paper W15/07.

JRF, 2018. *UK Poverty 2018,* York: Joseph Rowntree Foundation.

Kamp Dush, C. M. & Amato, P., 2005. Consequences of relationship status and quality for subjective well-being. *Journal of Social and Personal Relationships,* 22(5), p. 607–627.

Kasper, J. D. et al., 2008. Effects of Poverty and Family Stress Over Three Decades on the Functional Status of Older African American Women. *The Journals of Gerontology: Series B,* 63(4), S201–S210.

Kempson, E., Bryson, A. & Rowlingson, K., 1994. *Hard times: how poor families make ends meet.* London: Policy Studies Institute.

Kenny, D. A., 2014. *Measuring Model Fit.* [Online] Available at: http://davidakenny.net/cm/fit.htm [Accessed 15 July 2015].

Kenny, D. A., Kaniskan, B. & McCoach, D. B., 2015. The Performance of RMSEA in Models With Small Degrees of Freedom. *Sociological Methods & Research,* 44(3), pp. 486-507.

Keynes, J. M., 1963. Economic Possibilities for our Grandchildren [1930]. In: *Essays in Persuasion.* London: W. W. Norton, pp. 358-375.

Klein, R. & Mayblin, J., 2012. *Thinking about rationing.* London: The Kings Fund.

Klein, R., Day, P. & Redmayne, S., 1996. *Managing Scarcity: Priority Setting and Rationing in the National Health Service.* London: Open University Press.

Kline, R. B., 2010. *Principles and Practice of Structural Equation Modeling.* New York, NY: Guilford Press.

Knies, G., 2017. *Understanding Society: The UK longitudinal study wave 1-7 user guide.* Essex: Institute for Social and Economic Research.

Knott, C., 2012. General mental and physical health. *Health Survey England 2012,* 1(4).

Krishna, A., 2004. Escaping Poverty and Becoming Poor: Who Gains, Who Loses, and Why?. *World Development,* 32(1), pp. 121-136.

Kruglanski, A. W., 1980. Lay epistemologic process and contents: another look at attribution theory. *Psychological Review,* Volume 87, pp. 70-87.

Kuhn, T., 1970 [1962]. *The Structure of Scientific Revolutions, 2nd ed.* Chicago: University of Chicago Press.

Kuroda, S. & Yamamoto, I., 2018. Why Do People Overwork at the Risk of Impairing Mental Health?. *Journal of Happiness Studies,* https://doi.org/10.1007/s10902-018-0008-x.

Lambert, M. & Crossley, S., 2016. 'Getting with the (troubled families) programme': a review. *Social Policy & Society*. pp. 87-97.

Land, H. & Rose, H., 1985. Compulsory Altruism for All, or an Altruistic Society for Some? In: *In Defence of Welfare.* London: Routledge.

Lavrakas, P. J., 2008. *Encyclopedia of Survey Research Methods.* London: Sage.

Lee, E. S. & Forthofer, R. N., 2006. *Analyzing complex survey data: second edition.* London: Sage.

Lister, R., 2004. *Poverty.* Cambridge: Polity Press.

Lister, R., 2006. Recognition and Voice, Gender and Poverty: The Challenge for Social Justice. In: U. Degener & R. B., eds. *Die Neuverhandlung sozialer Gerechtigkeit.* Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 53-63.

Lister, R., 2015. 'To count for nothing': Poverty beyond the statistics. *Journal of the British Academy,* Volume 3, pp. 139-165.

Little, T. D., 2013. *Longitudinal structural equation modeling.* London: Guilford.

Little, T. D., Jorgensen, T. D., Lang, K. M. & Moore, E. G. M., 2014. On the Joys of Missing Data. *Pediatric Psychology,* 39(2), pp. 151-162.

Liu, M. & Lin, T. I., 2014. Skew-normal factor analysis models with incomplete data. *Journal of Applied Statistics,* 42(4), pp. 789-805.

Liu, X., Zhu, X., Qiu, P. & Chen, W., 2012. A Correlation-Matrix-Based Hierarchical Clustering Method for Functional Connectivity Analysis. *Journal of Neuroscience Methods,* 211(1), pp. 94-102.

Loh, W., Eltinge, J., Cho, M. & Li, Y., 2016. Classification and regression tree methods for incomplete data from sample surveys. *arXiv: 1603.01631.*

LSE (London School of Economics), 2016. *The Welfare Trait: how state benefits affect personality.* [Online] Available at: http://www.lse.ac.uk/publicEvents/events/2016/06/20160629t1830vLSE.aspx [Accessed 18 October 2016].

Luthar, S. S., 1999. *Poverty and children's adjustment.* London: Sage.

Lynn, P. & Knies, G., 2016. *Understanding Society: The UK Household Longitudinal Study Waves 1-5 Quality Profile,* Colchester: Understanding Society.

Macdonald, R., Shildrick, T. & Furlong, A., 2013. In search of 'intergenerational cultures of worklessness': Hunting the Yeti and shooting zombies. *Critical Social Policy,* 34(2), pp. 199-220.

MacInnes, J., 2017. *An Introduction to Secondary Data Analysis with IBM SPSS Statistics.* London: Sage.

MacInnes, T. et al., 2013. *Monitoring Poverty and Social Exclusion 2013,* York: Joseph Rowntree Foundation.

Macmillan, L., 2010. The Intergenerational Transmission of Worklessness in the UK. *CMPO Working Paper No. 10/231.*

Maplethorpe, N., Chanfreau, J., Philo, D. & Tait, C., 2010. *Families with Children in Britain: Findings from the 2008 Families and Children Study, Research Report 656,* London: Department for Work and Pensions.

Maslow, A. H., 1943. A theory of human motivation. *Psychological Review,* Volume 50, pp. 370-396.

Mason, J., 2002. *Qualitative Researching.* London: Sage.

Matthews, C. K. & Harrington, N. G., 2000. Invisible disability. In: D. O. Braithwaite & T. L. Thompson, eds. *LEA's communication series. Handbook of communication and people with disabilities: Research and application.* Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 405-421.

McAuley, C. & Rose, W., 2010. *Child Well-being: Understanding children's lives.* London: Jessica Kingsley.

McKay, S., 2008. *Measuring material deprivation among older people: Methodological study to revise the Family Resources Survey questions,* Norwich: Department for Work and Pensions.

Millar, J. & Glendinning, C., 1989. Gender and Poverty. *Journal of Social Policy,* 18(3), pp. 363-381.

Millsap, R. E., 2011. *Statistical Approaches to Measurement Invariance.* London: Routledge.

Milton, S. et al., 2015. Understanding welfare conditionality in the context of a generational habitus: A qualitative study of older citizens in England. *Journal of Aging Studies,* Volume 34, pp. 113-122.

Ministry of Housing, Communities & Local Government, 2018. *English Housing Survey.* [Online] Available at: https://www.gov.uk/government/collections/english-housing-survey [Accessed 5 Jan 2019].

Minton, J. et al., 2017. Visualising and quantifying 'excess deaths' in Scotland compared with the rest of the UK and the rest of Western Europe. *Journal of Epidemiology and Community Health,* Volume 71, pp. 461-467.

Mkandawire, T., 2005. *Targeting and Universalism in Poverty Reduction,* Geneva: United Nations Research Institute for Social Development.

Morris, L., 1995. *Social Divisions: Economic Declineand Social Structural Change.* London: University College London Press.

Mukuria, C., 2016. *An empirical comparison of well-being measures used in the UK.* Sheffield: Policy Research Unit in Economic Evaluation of Health and Social Care Interventions.

Murphy, M. J., 1996. The dynamic household as a logical concept and its use in demography. *European Journal of Population,* Volume 12, pp. 363-381.

Muthén, L. K. & Muthén, B. O., 2017. *Mplus User's Guide.* Eighth Edition ed. Los Angeles, CA: Muthén & Muthén.

Newsome, K. et al., N.D.. *Living on the Edge: Experiencing workplace insecurity in the UK,* London: Trade Union Congress.

Niles, A. N. et al., 2015. Anxiety and Depressive Symptoms and Medical Illness Among Adults with Anxiety Disorders. *Journal of Psychosomatic Research,* 78(2), pp. 109-115.

Nosek, B. A., Spies, J. R. & Motyl, M., 2012. Scientific Utopia: II. Restructuring Incentives and Practicies to Promote Truth Over Publishability. *Perspectives on Psychological Science,* 7(6), pp. 615-631.

O'Brien, R. M., 2011. The age–period–cohort conundrum as two fundamental problems. *Quality & Quantity,* 45(6), pp. 1429-1444.

Oakley, A., 1976. *Housewife.* Middlesex: Penguin Books.

Office for National Statistics, 2013. *Poverty and Social Exclusion in the UK and EU, 2005-2011,* Newport: Office for National Statistics.

Office for National Statistics, 2017. *Families and Households: 2017.* [Online] Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2017 [Accessed 5 January 2019].

Office for National Statistics, 2019. *UK labour market: January 2019.* [Online] Available at: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/uklabourmarket/january2019#economic-inactivity [Accessed 6 Feb 2019].

Oliver, M., 1983. *Social Work with Disabled People.* Basingstoke: Macmillan.

Oliver, M., 2013. The social model of disability: thirty years on. *Disability & Society,* 28(7), p. 1024–1026.

Padley, M. & Hirsch, D., 2016. *Households below a Minimum Income Standard: 2008/09 to 2013/14,* York: Joseph Rowntree Foundation.

Padley, M., Hirsch, D. & Valadez, L., 2017. *Households Below a Minimum Income Standard: 2008/9 - 2014/15,* York: Joseph Rowntree Foundation.

Pahl, J., 1989. Money and Marriage. In: *Women and Poverty: Exploring the Research and Policy Agenda.* London: Thomas Coram Research Unit, p. 215.

Pahl, J., 1989. Money and Marriage. In: *Women and Poverty: Exploring the Research and Policy Agenda.* London: Thomas Coram Research Unit.

Palmer, M., 2011. Disability and Poverty: A Conceptual Review. *Journal of Disability Policy Studies,* 21(4), p. 210–218.

Pantazis, C. & Ruspini, E., 2006. Gender, poverty and social exclusion. In: C. Pantazis, D. Gordon & R. Levitas, eds. *Poverty and Social Exclusion in Britain.* Bristol: Policy Press.

Pantazis, C., Gordon, D. & Levitas, R., 2006. *Poverty and Social Exclusion in Britain: the Millenium Survey.* Bristol: The Polity Press.

Pantazis, C., Gordon, D. & Townsend, P., 2006. The necessities of life. In: C. Pantazis, D. Gordon & R. Levitas, eds. *Poverty and social exclusion in Britain.* Bristol: Policy Press, pp. 89-123.

Pantazis, C., Gordon, D. & Townsend, P., 2006. The necessities of life. In: *Poverty and Social Exclusion in Britain.* Bristol: Policy Press, pp. 89-123.

Patrick, R., 2011. Disabling or Enabling: The Extension of Work-Related Conditionality to Disabled People. *Social Policy & Society,* 10(3), pp. 309-320.

Patsios, D., 2006. Pensioners, poverty and social exclusion. In: C. Pantazis, D. Gordon & R. Levitas, eds. *Poverty and Social Exclusion in Britain.* Bristol: Policy Press.

Payne, G., 2013. *Social Divisions.* London: Palgrave Macmillan.

Peveler, R., Carson, A. & Rodin, G., 2002. Depression in medical patients. *British Medical Journal,* 325(7356), p. 149–152.

Piachaud, D., 1987. Problems in the Definition and Measurement of Poverty. *Journal of Social Policy,* 16(2), pp. 146-164.

Piketty, T., 2014. *Capital in the Twenty-First Century.* London: Belknap Press.

Piredda, M. et al., 2016. Psychometric testing of the modified Care Dependency Scale (Neuro-CDS). *NeuroRehabilitation,* 38(2), pp. 211-219.

Platt, L., 2005. *Discovering Child Poverty: The creation of a policy agenda from 1800 to the present.* Bristol: Policy Press.

Platt, L., 2007. *Poverty and Ethnicity in the UK.* Bristol: Policy Press / JRF.

Ponthieux, S., 2013. *Income pooling and equal sharing within the household: what can we learn from the 2010 EU-SILC module?.* Luxembourg: European Commission: Eurostat Population and Social Conditions: Methodologies and Working Papers.

Price, D., 2006. The Poverty of Older People in the UK. *Journal of Social Work Practice,* 20(3), pp. 251-266.

R Core Team, 2016. *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Raykov, T., 2011. Intraclass Correlation Coefficients in Hierarchical Designs: Evaluation Using Latent Variable Modeling. *Structural Equation Modeling: A Multidisciplinary Journal,* Volume 18, pp. 73-90.

Reeve, K., 2017. Welfare conditionality, benefit sanctions and homelessness in the UK : ending the 'something for nothing culture' or punishing the poor?. *Journal of Poverty and Social Justice,* 25(1), pp. 65-78.

Remnant, J. & Adams, J., 2015. The nutritional content and cost of supermarket ready-meals. Cross-sectional analysis. *Appetite,* 1(92), pp. 36-42.

Roberts, K., 2001. *Class in Modern Britain.* Basingstoke: Palgrave Macmillan.

Robson, K. & Pevalin, D., 2015. *Multilevel Modeling in Plain Language.* London: Sage.

Rolstad, S., Adler, J. & Rydén, A., 2011. Response Burden and Questionnaire Length: Is Shorter Better? A Review and Meta-analysis. *Value in Health,* 14(8), pp. 1101-1108.

Rossi, P. H., Wright, J. D. & Anderson, A. B., 2013. *Handbook of Survey Research.* London: Academic Press.

Rowntree, S., 1901. *Poverty: A Study of Town Life.* London: Macmillan.

RStudio Team, 2018. *RStudio: Integrated Development for R.* Boston, MA: RStudio, Inc.

Rubin, D. B., 1976. Inference and Missing Data. *Biometrika,* Volume 63, pp. 581-590.

Sampson, R., 1988. Local Friendship Ties and Community Attachment in Mass Society: A Multilevel Systemic Model. *American Sociological Review,* 53(5), pp. 766-779.

Savage, M. et al., 2013. A New Model of Social Class? Findings from the BBC's Great British Class Survey Experiment. *Sociology,* 47(2), pp. 219-250.

Savage, M., 2015. *Social Class in the 21st Century.* London: Pelican Books.

Schafer, J. L. & Graham, J. W., 2002. Missing Data: Our View of the State of the Art. *Psychological Methods,* 7(2), pp. 147-177.

Schroeder, L. D., Sjoquist, D. L. & Stephan, P. E., 1986. *Understanding Regression Analysis: an introductory guide.* London: Sage.

Sen, A., 1983. Poor, relatively speaking. *Oxford Economic Papers,* Volume 35, pp. 135-169.

Sen, A., 1989. Development as capability expansion. *Journal of Development Planning,* 19(1), pp. 41-58.

Sen, A., 1995. *Commodities and Capabilities.* Oxford: Oxford University Press.

Shale, J. et al., 2015. *Households Below Average Income: An analysis of the income distribution 1994/95–2013/14,* London: Department for Work & Pensions.

Shevlin, M. & Adamson, G., 2005. Alternative factor models and factorial invariance of the GHQ-12: a large sample analysis using confirmatory factor analysis. *Psychological Assessment,* 17(2), pp. 231 - 236.

Shildrick, T. & MacDonald, R., 2013. Poverty Talk: How People Experiencing Poverty Deny Their Poverty and Why They Blame 'The Poor'. *The Sociological Review,* 61(2).

Shildrick, T., 2018. *Poverty propaganda: Exploring the myths.* Bristol: Policy Press.

Shildrick, T., MacDonald, R., Webster, C. S. & Garthwaite, K., 2012. *Poverty and Insecurity: Life in low-pay, No-pay Britain.* Bristol: Policy Press.

Shrout, P. & Fleiss, J. L., 1979. Intraclass correlation: uses in assessing rater reliability. *Psychological Bulletin,* 86(2), pp. 420-428.

Simmons, D., 2011. *Escalating conditionality.* [Online] Available at: http://www.cpag.org.uk/content/escalating-conditionality [Accessed 1 August 2016].

Smart, C., 2007. *Personal Life: New directions in sociological thinking.* Cambridge: Polity.

Snowdon, C., 2014. *The Minimum Income Standard: the wisdom of crowds?.* [Online] Available at: https://iea.org.uk/blog/the-minimum-income-standard-the-wisdom-of-crowds [Accessed 12 Jan 2018].

Social Metrics Commission, 2018. *A new measure of poverty for the UK,* London: Legatum Institute.

Spicker, P., 2004. Developing Indicators: Issues in the Use of Quantitative Data about Poverty. *Policy and Politics,* 32(4), pp. 431-440.

Spicker, P., 2007. *The Idea of Poverty.* Britsol: Policy Press.

Spicker, P., Leguizamón, S. A. & Gordon, D., 2007. *Poverty: An international glossary 2nd ed..* London: Zed (CROP).

Standing, G., 2014. *The Precariat: The new dangerous class.* London: Bloomsbury.

Steele, F., Clarke, P. & Washbrook, E., 2013. Modeling household decisions using longitudinal data from household panel surveys, with applications to residential mobility. *Sociological Methodology,* 43(1), pp. 220-271.

Stewart-Brown, S. L. et al., 2011. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a valid and reliable tool for measuring mental well-being in diverse

populations and projects. *Journal of Epidemiology & Community Health,* Volume 65, pp. A38-A39.

Strine, T. W. et al., 2008. The associations between life satisfaction and health-related quality of life, chronic illness, and health behaviors among U.S. community-dwelling adults. *Journal of Community Health,* 33(1), pp. 40-50.

Stubbs, S., 2008. *In place of drums and samosas, The Guardian.* [Online] Available at: https://www.theguardian.com/society/2008/may/14/equality [Accessed 2 Jan 2019].

Sutherland, H. et al., 2008. *The impact of benefit and tax uprating on incomes and poverty,* York: Joseph Rowntree Foundation.

Sweney, M., 2018. *Netflix to overtake Sky's satellite TV subscriptions by end of year.* [Online] Available at: https://www.theguardian.com/media/2018/dec/23/netflix-to-overtake-sky-satelite-tv-subscriptions-by-end-of-year [Accessed 18 Jan 2019].

Tabachnick, B. G. & Fidell, L. S., 2013. Cleaning up your act. In: B. G. Tabachnick & L. S. Fidell, eds. *Using Multivariate Statistics.* London: Pearson, pp. 93-151.

Tabachnick, B. G. & Fidell, L. S., 2013. Principal Components and Factor Analysis. In: B. G. Tabachnick & L. S. Fidell, eds. *Using Multivariate Statistics.* London: Pearson, pp. 659-730.

Tabachnick, B. G. & Fidell, L. S., 2014. *Using Multivariate Statistics.* Harlow: Pearson.

Tay, L. & Diener, E., 2011. Needs and Subjective Well-Being Around the World. *Journal of Personality and Social Psychology,* 101(2), pp. 354-365.

Tinson, A., Aldridge, H., Born, T. B. & Hughes, C., 2016. *Disability and poverty: Why disability must be at the centre of poverty reduction,* London: New Policy Institute.

Tinson, A., Kenway, P., Bushe, S. & MacInnes, T., 2014. *Poverty and the Cost of Living: An evidence review,* York: Joseph Rowntree Foundation.

Tomlinson, M. & Walker, R., 2009. *Coping with complexity: Child and adult poverty.* London: Child Poverty Action Group.

Tomlinson, M. & Walker, R., 2010. *Recurrent poverty: the impact of family and labour market changes,* York: Joseph Rowntree Foundation.

Tomlinson, M., Walker, R. & Williams, G., 2008. Measuring poverty in Britain as a multi-dimensional concept, 1991 to 2003. *Journal of Social Policy,* pp. 37(4): 597-620.

Tomlinson, M., Walker, R. & Williams, G., 2010. The Problem with Poverty: Definition, Measurement & Participation. In: *The SAGE Handbook of Measurement.* London: Sage, pp. 353-394.

Tonkin, R., 2016. *Persistent Poverty in the UK and EU: 2014.* [Online] Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/articles/persistentpovertyintheukandeu/2014 [Accessed 17 May 2016].

Townsend, P. & Gordon, D., 1989. *What is enough? New evidence on poverty in Greater London allowing the definition of a minimum benefit: Memorandum of evidence to the House of Commons Social Services Select Committee on Minimum Income 579,* London: HMSO.

Townsend, P., 1962. The meaning of poverty. *British Journal of Sociology,* 13(3), pp. 210-227.

Townsend, P., 1979. *Poverty in the United Kingdom.* Harmondsworth: Allen Lane.

Townsend, P., 1985. A Sociological Approach to the Measurement of Poverty: A Rejoinder to Professor Amartya Sen. *Oxford Economic Papers,* Volume 37, pp. 659-668.

Townsend, P., 2010. The international analysis of Poverty. In: *The Peter Townsend Reader.* Bristol: Policy Press, pp. 57-82.

Townsend, P., 2010. The need to construct an international welfare state. In: *The Peter Townsend Reader.* Bristol: Policy Press, pp. 82-83.

Tukey, J. W., 1977. *Exploratory data analysis.* Reading, PA: Addison-Wesley.

UK Data Service, N.D.. *Help - Frequently Asked Questions (FAQs).* [Online] Available at: https://www.ukdataservice.ac.uk/help/faq [Accessed 31 May 2018].

Ullman, J. B., 2014. Structural Equation Modeling. In: *Using Multivariate Statistics.* Harlow: Pearson Education Ltd, pp. 731-836.

Understanding Society, 2017. *Annual findings report.* [Online] Available at: https://www.understandingsociety.ac.uk/research/findings [Accessed 31 May 2018].

University of Essex. Institute for Social and Economic Research, NatCen Social Research, Kantar Public, 2017. *Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 9th Edition.* UK Data Service. SN: 6614.

van Buuren, S. & Groothuis-Outshoorn, K., 2011. mice: Multivariate imputation by chained equations. *Journal of Statistical Software,* 45(3).

van de Schoot, R. et al., 2015. Editorial: Measurement Invariance. *Frontiers in Psychology,* Volume 6, pp. 1-4.

van de Schoot, R., Lugtig, B. & Hox, J., 2012. A checklist for testing measurement invariance. *European Journal of Developmental Psychology,* Volume 9, pp. 486-492.

van der Maesen, L. & Walker, A., 2012. *Social Quality: From Theory to Indicators,* London: Palgrave Macmillan.

Van Der Maesen, L. & Walker, A., 2012. *Social Quality: from Theory to Indicators.* London: Palgrave Macmillan.

Vecchio, G. M. & Caprara, G. V., 2007. Multi-faceted self-efficacy beliefs as predictors of life-satisfaction in late adolescence. *Personality and Individual Differences,* 43(7), pp. 1807-1818.

Veit-Wilson, J., 1999. Poverty and the adequacy of social security. In: *Introduction to Social Security.* London: Routledge.

Velotti, P., Garofalo, C., Bottazzi, F. & Caretti, V., 2017. Faces of Shame: Implications for Self-Esteem, Emotion Regulation, Aggression, and Well-Being. *The Journal of Psychology,* Volume 151, pp. 171-184.

Verhaeghe, P., 2014. *What about Me? the struggle for identity in a market-based society.* London: Scribe.

Vickery, C., 1977. The Time Poor: A new look at Poverty. *The Journal of Human Resources,* 12(1), pp. 27-48.

Vogler, C., 1994. Money in the household. In: M. Anderson, F. Bechhofer & J. Gershuny, eds. *The Social and Political Economy of the Household.* Oxford: Oxford University Press.

Walker, A. & Mollenkopf, H., 2007. International and Multidisciplinary Perspectives on Quality of Life in Old Age. In: *Quality of Life in Old Age.* Dordrecht: Springer, pp. 3-15.

Walker, A. & Naegele, G., 1999. *The Politics of Old Age in Europe.* Buckingham: Open University Press.

Walker, A. et al., 2010. *The Peter Townsend reader.* Bristol: Policy Press.

Walker, A., 1980. The Social Creation of Poverty and Dependency in Old Age. *Journal of Social Policy,* 9(1), pp. 49-75.

Walker, A., 1993. Poverty and Inequality in Old Age. In: *Ageing in Society: an Introduction to Social Gerontology.* London: Sage, pp. 280-304.

Walker, A., 2017. Why the UK Needs a Social Policy on Ageing. *Journal of Social Policy,* 47(2), pp. 253-273.

Walker, R., 1987. Consensual Approaches to the Definition of Poverty: Towards an Alternative Methodology. *Journal of Social Policy,* 16(2), pp. 213-226.

Walker, R., 2014. Conceiving of Poverty without Shame. In: R. Walker, ed. *The Shame of Poverty.* Oxford: Oxford University Press, pp. 85-97.

Walker, R., 2014. *The Shame of Poverty: Global perspectives.* Oxford: Oxford University Press.

Wall, E., 2012. *£7.3bn means-tested tax credits go unclaimed.* [Online] Available at: http://www.telegraph.co.uk/finance/personalfinance/tax/9423277/7.3bn-means-tested-tax-credits-go-unclaimed.html [Accessed 1 August 2016].

Webb, C. J. R. & Bywaters, P., 2018. Austerity, rationing and inequity: trends in children's and young peoples' services expenditure in England between 2010 and 2015. *Local Government Studies.*

Wei, T. et al., 2017. *R package "corrplot": Visualization of a Correlation Matrix.* [Online] Available at: https://cran.r-project.org/web/packages/corrplot/corrplot.pdf [Accessed 29 Dec 2018].

White, I. R., Royston, P. & Wood, A. M., 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine,* 30(4), pp. 377-399.

Wickham, H., 2009. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag.

Wickham, H., Francois, R., Henry, L. & Müller, K., 2017. *dplyr: A grammar of data manipulation,* s.l.: R package version 0.7.4: https://CRAM.R-project.org/package=dplyr.

Wickham, S., Whitehead, M., Taylor-Robinson, D. & Barr, B., 2017. The effect of a transition into poverty on child and maternal mental health: a longitudinal analysis of the UK Millennium Cohort Study. *Lancet Public Health,* Volume 2, pp. 141-148.

Widaman, K. F., Ferrer, E. & Conger, R., 2010. Factorial Invariance within Longitudinal Structural Equation Models: Measuring the Same Construct across Time. *Child Development Perspectives,* 4(1), pp. 10-18.

Wintour, P. & Watt, N., 2015. *Tory conference: Cameron's 'assault on poverty' pledge belied by new figures.* [Online] Available at: http://www.theguardian.com/politics/2015/oct/07/camerons-assault-on-poverty-pledge-undone-by-new-figures [Accessed 17 May 2016].

Wolf, E. J., Harrington, K. M., Clark, S. L. & Miller, M. W., 2013. Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement,* 76(6), pp. 913-934.

World Bank, 2016. *Poverty headcount ratio at $1.90 a day (2011 PPP) (% of population).* [Online] Available at: http://data.worldbank.org/indicator/SI.POV.DDAY?page=6 [Accessed 1 June 2016].

World Health Organisation, 2014. *Social Determinants of Mental Health.* Geneva: World Health Organisation.

Yang, Y. C. et al., 2016. Social relationships and physiological determinants of longevity across the human life span. *Proceedings of the National Academy of Sciences,* 113(3), pp. 578-583.

Zelinsky, T., Mysikova, M. & Vecernik, J., 2016. Occupational Mobility, Educational Mobility and Intergenerational Transmission of Disadvantages in Europe. *Ekonomicky casopis,* 64(3), pp. 197-217.

Zilanawala, A., 2016. Women's Time Poverty and Family Structure: Differences by Parenthood and Employment. *Journal of Family Issues,* 37(3), pp. 369-392.