The Evaluation of a Self-enumerated Scale of Quality of Life (CASP-19) in the Context of Research on Ageing: A Combination of Exploratory and Confirmatory Approaches

R. D. Wiggins · G. Netuveli · M. Hyde · P. Higgs · D. Blane

Abstract This paper describes the conceptual development of a self-enumerated scale of quality of life (CASP-19) and presents an empirical evaluation of its structure using a combination of exploratory and confirmatory factor analytic approaches across three different survey settings for older people living in England and Wales in the new millennium. All evaluations are conducted using MPlus which allows the analyst to evaluate the properties of the scale for a set of multivariate categorical items which are subject to item non-response. CASP-19 is a subjective measure of well-being derived from an explicit theory of human need spanning four life domains: control, autonomy, self-realisation and pleasure. Put formally, CASP-19 is a self-reported summative index consisting of 19 Likert scale items. The three survey settings include a postal survey of 263 people in early old age followed up from childhood when the respondents were first interviewed in the 1930's, the first wave (2002) of the English Longitudinal Study of Ageing (ELSA_1) and the eleventh wave of the British Household Panel Survey (BHPS_11) also conducted in 2002. These nationally representative surveys consisted of 9300 and 6471 respondents aged 55 years and older. The Boyd-Orr sample provides an exploratory context for the evaluation and ELSA_1 together with BHPS_11 provide the opportunity for confirmatory analyses of three measurement models. There is some support for the use of CASP-19 as a stand alone summative index. However, the analysis reveals that a shortened 12-item scale which combines the life domains 'control and autonomy' in a second order measurement model is the recommended model for analysts. The work was funded under the UK's Economic and

R. D. Wiggins (✉)
Department of Quantitative Social Science, The Faculty of Policy and Society, The Institute of Education, University of London, 20 Bedford Way, London, WC1H 0AL, UK
e-mail: r.wiggins@ioe.ac.uk

G. Netuveli · D. Blane
Imperial College of Science, Technology and Medicine, London, UK

M. Hyde
Sheffield Hallam University, Sheffield, UK

P. Higgs
University College, London, UK

Keywords Quality of life · Confirmatory and exploratory factor analysis · Shortened 12-item version of CASP-19 · Ageing

1 Introduction

Twenty-five years ago Baker and Intagliata (1982) remarked that there are almost 'as many definitions (*of quality of life measures*) as the number of people studying the phenomenon'. Following a number of excellent reviews of the definition and measurement of quality of life (Felce and Perry 1995; Cummins 1996; Diener and Suh 1997; Bowling 1997; Smith 2000; Bullinger 2002; Brown et al. 2004) that this remark has equal relevance today.

So why add to this array of quality of life measures? We do so because we feel that our measure is well supported by a strong theoretical argument for the basis measuring quality of life in the context of ageing. Secondly, in this paper we subject the properties of the measure to rigorous factor analytic assessment so as to test and refine the strength of our theory. The transparency of our analytical framework should provide a clear way forward for analysts who are keen to and review new and existing measures of quality of life. Our measure has been developed to take careful account of the need to have a definition of quality of life that stands alone from the factors that influence it, e.g. health or income. Our measurement approach is grounded upon the idea that the operational definition of quality of life is composed of related yet discrete domains. We reject the notion that a single item like 'How do you feel about your life as a whole?' (Andrews and Withey 1976) can be a barometer of life quality and follow the psychometric tradition that argues that multi-item scales are typically superior 'in respect of such matters as reliability, unidimensionality and specific wording bias' (McKennell 1977).

The strength of our underlying theory and subsequent evaluation drive the usefulness of our measure. It sits within an established tradition of defining subjective well-being as described by Diener (1994) and meets his three key criteria: 'First, it is subjective—it resides in the experience of the individual. Second, it is not just the absence of negative factors, but also includes positive measures. Third, it includes a global assessment rather than only a narrow assessment of one life domain' (p. 106).

In sum, our measure, CASP-19 is a theoretically grounded measure of quality of life consisting of 19 Likert scaled agreement items spanning four life domains: control (C), autonomy (A), self-realisation (S) and pleasure (P). The first letter of each domain label are joined together to create the acronym, CASP. Each life domain contains four or five items which are presented as statements to survey respondents (see Table 1). Each statement is assessed on a four point Likert scale as to the extent to which the description describes a person's feelings about their life (rated 'this applies to me: 'often', 'sometimes' 'not often' or 'never') applies to the respondent. The resulting scale scores are summed to form an index of quality of life where a high score indicates 'good' quality of life. For our overall measure of quality of life nineteen items form the scale hence, the label CASP-19.

The original scale was developed in the context of a postal follow-up to members of the Boyd-Orr sample in 2000 (Hyde et al. 2003). This unique sample was derived from a childhood survey of 1,352 families drawn from 16 locations in Great Britain by Sir John Boyd Orr and his team (Gunnel et al. 1996) for a study of family diet in pre-war Britain

Table 1 CASP-19 item wording arranged by domain categories

Item	Statement
Control	
1 C[a]	My age prevents me from doing the things I would like to do
2 C[a]	I feel that what happens to me is out of my control
3 C	I feel free to plan for the future
4 C[a]	I feel left out of things
Autonomy	
5 A	I can do the things I want to do
6 A[a]	Family responsibilities prevent me from doing the things I want to do
7 A	I feel that I can please myself what I do
8 A[a]	My health stops me from doing the things I want to do
9 A[a]	Shortage of money stops me from doing things I want to do
Pleasure	
10 P	I look forward to each day
11 P	I feel that my life has meaning
12 P	I enjoy the things that I do
13 P	I enjoy being in the company of others
14 P	On balance, I look back on my life with a sense of happiness
Self-realisation	
15 S	I feel full of energy these days
16 S	I choose to do things that I have never done before
17 S	I feel satisfied with the way my life has turned out
18 S	I feel that life is full of opportunities
19 S	I feel that the future looks good for me

[a] Item reverse coded for scoring

Note: In the original administration of the scale (Boyd-Orr) the 19 items are presented as a section entitled 'Your feelings about life'. Respondents are invited to indicate the extent to which each item applies to themselves on a four point Likert scale, 'often', 'sometimes', 'not often' or 'never'

between 1937 and 1939. During 1996, 99% of household records and 96% of the children's medical records were retrieved from the Rowlett Research Institute and entered into an electronic database at the Department of Social Medicine at the University of Bristol. Using the National Health Service Register, the Office for National Statistics successfully traced 85% of the children who participated in 1937–39 study. During 1997 David Blane and his colleagues (Berney and Blane 1997) conducted a follow-up survey of 293 of the original participants from the Boyd-Orr survey. These respondents were aged 5–14 years at the time of the original study and for whom complete records were available. The sample members were drawn from a stratified random sample of traced survivors and were shown by Berney and Blane to have social characteristics that were broadly comparable to the older resident population when compared to the 1991 decennial census. In 2000 these individuals were contacted again as part of the Economic and Social Research Council's Growing Older programme (www.growingolder.group.shef.ac.uk/). Subsequent analyses are based on an achieved sample of 263 aged between 65 and 74 years. The scale has recently been adopted in many national and cross-national studies of ageing, notably the two British surveys, the English Longitudinal Study of Ageing (ELSA) (Marmot et al. 2003) and the British Household Panel Survey (BHPS) both for its retirement module and to all adults aged over 16 years (Taylor et al. 2003) These major national studies form the sample sources for the statistical evaluation reported in this paper. The Boyd-Orr survey

(2003) now has a follow-up: Boyd-Orr 2004–5. CASP-19 is included in the cross-national survey, Health, Alcohol and Psychosocial factors in Eastern Europe (HAPIEE Study 2005) and the American Health and Retirement Survey. Recent translated versions of CASP-19 now include the Korean longitudinal Study of Ageing (KLoSA 2007) and a Malayalam version (Netuveli et al. 2007). At a regional or small study level, CASP-19 is a part of a community evaluation survey in Camden, London, UK; Wellbeing in Active Seniors, University of Westminster, London and Generations in Action, The Beth Johnson Foundation, Manchester. A shorter derivative, based on a provisional evaluation of the Boyd-Orr data, CASP-12 is included in the Study of Health Ageing and Retirement in Europe (SHARE, Börsch-Supan et al. 2005) which covers 11 European countries. In the evaluation that follows we first describe the theoretical development of the scale, its preliminary assessment using a combination of exploratory factor analysis and reliability analysis (McKennell 1977; Wiggins and Bynner 1993) using the first Boyd-Orr implementation of CASP-19 and subsequent confirmatory analyses (Maxwell 1977) across the first wave of ELSA (ELSA_1) and the retirement module in BHPS wave eleven (BHPS_11).

1.1 Theoretical Background for the Development of the Scale

Our measure of quality of life was developed in the context of a study of ageing at a time when changing social, economic and demographic circumstances of people in early old age, the 'young-old' required a rethink of the concept of what it was to be 'old' at the turn of the new millennium in Britain (ONS 1998; Phillipson 1998; Gilleard and Higgs 2000; Vincent 2003). In this historical period older people are no longer seen as 'dependent' upon the state for health needs or financial security but as socially active and increasingly free from work and family constraints (Young and Schuller 1991; Laslett 1996; Scase and Scales 2000). Following Laslett (1996) we will refer to this period as a distinct 'age' or stage in a person's life. The 'first' being a time of childhood dependency, the 'second' as a time of economic independence and the 'third age' a time to pursue a good quality of life. The 'fourth' age being characterised by declining health and frailty. We avoid the imposition of any chronological boundaries on these stereotypical descriptions of the life course simply because the continuum from mid-life independence to later life dependence varies so much across individuals. We will use the term 'early old age' and 'young-old' as alternatives to the terms 'third age' or 'third agers'.

It is this changing context of ageing in which the experiences of those in retirement are quite different from those of their predecessors (Gilleard and Higgs 2000) which informs the development of our measure. It draws upon a view of later life that is agentic and reflexive a time for self-realisation and pleasure (Hyde et al. 2003). Of course, this is not to argue that older people are now released form poverty or ill-health but to change the emphasis of description of life to one that it is positive and life affirming Higgs et al. (2003), Wiggins et al. (2007).

The arguments presented in Hyde et al. (2003) and Higgs et al. (2003) are predicated upon a position that quality of life in early old age has been under-theorised. Typically, measures of quality of life are based on proxies, such as health, which draw on a set of normative assumptions about what a particular condition implies for a person's quality of life without necessarily taking close account of a person's current life experience. The result is to substitute direct measures of quality of life with measures which describe the influences upon quality of life without measuring quality of life itself. As an alternative other researchers turned to taking careful subject-led accounts in the construction of

measures of quality of life (Bowling 1995; Farquhar 1995). Often these approaches are premised on asking subjects to rate some of the most important things in their lives. Whilst this method can be insightful and revealing it can also suffer from the weakness of the proxy measures in that the outcomes are often accounts of what constitutes a good quality of life as opposed to the bundle of feelings about freedom and constraint, the past, the present and the future.

The idea that a measure of quality of life should stand alone from the factors or influences that shape it was derived from Doyal and Gough's theory of human need that puts the biological and the social on an equal footing (Doyal and Gough 1991). This approach recognises that whilst there are common basic needs, like food and shelter, it is just as important to recognise that 'being human' is an active and reflective process (Giddens 1990; Turner 1995). This 'needs satisfaction' approach leads us to identify four conceptual domains of quality of life namely, 'control', 'autonomy', 'self-realisation' and 'pleasure'. Control and autonomy are natural prerequisites to the feelings associated with being able to participate in society and the extent to which these feelings of freedom can be realised is captured by the 'self-realisation' and 'pleasure' domains. The former represents the more reflexive nature of life and pleasure the sense of fun derived form the more active (doing) aspects of life. By separating the conceptual domains in this way we allow the theory to begin to describe the complexities of feelings about life. Under our theory for example, individuals can feel free but unable to seize the moment or have fun!

The final selection of the statements used to define CASP-19 were derived from our reading of the literature on quality of life, expert review[1] several focus groups and a number of one to one cognitive interviews. This process is fully described in Hyde et al. (2003). We now turn our attention to the evaluation of the operational scales.

1.2 The Role of Exploratory and Confirmatory Analysis

The evaluation of our operational definition of CASP-19 where 19 Likert scaled items are used to measure a person's quality of life undertakes a two-step journey. Firstly responses to items or statements are used to indicate the presence of an underlying concept or life domain, e.g. control. In classic measurement theory (Everitt 1984) this implies that any observed correlations between our items is due to their common dependence on an unobserved or latent variable. The latent variable or factor together with its corresponding set of indicators defines what is referred to as a 'measurement model'. For convenience these measurement models can be represented graphically. Certain conventions follow: concepts are represented by circles and indicators as rectangular boxes, lines arrows lead to the boxes defining the direction of dependency and arrows turn towards boxes or circles to indicate the presence of error. Figure 1 illustrates a measurement model for CASP-19 as a single unitary entity. All items load on to a single underlying factor which represents quality of life (or CASP-19). Confirmation of such a model would represent a popular option for users who simply want to work with a single summative index to compare population aggregates or subgroups.

From a substantive perspective our preferred description of the conceptual framework for CASP-19 is one in which the four life domains, control, autonomy, self-realisation and pleasure, map on to a single intrinsic or latent measure of quality of life. In this way, any

---

[1] Throughout the project we had regular meetings with our Advisory Panel consisting of Lee Berney, Anne Bowling, Maria Evandrou, Chris Gilleard, Katerina Hilari, Aubrey McKennell and Roger Thomas,
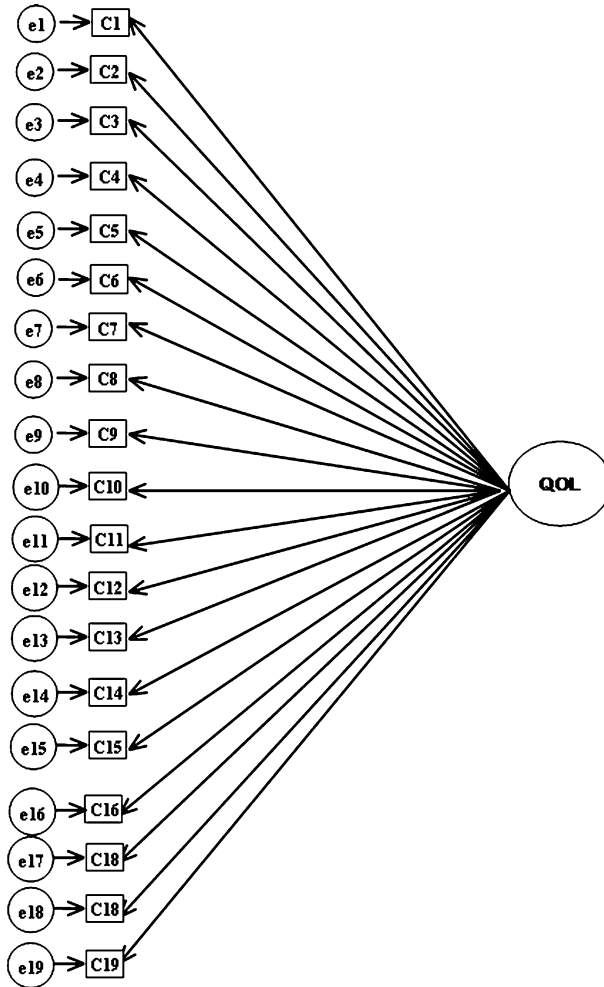
Fig. 1 A single factor measurement model for CASP-19

observed dependence amongst the four life domain scores could be argued to be due to their common dependence on an underlying measure of quality of life. In this second step of describing the model we have identified a structure for defining the inter-relationship between the domains or four factors, control, autonomy, self-realisation and pleasure. This conceptualisation is now referred to as a 'second order factor model' (Wiggins and Bynner 1993) and illustrated in Fig. 2 below.

In factor analytic tradition any model is one of several competing models (Everitt 1998; Bartholomew and Knott 1999). For instance, adding additional paths from concepts to indicators define alternative formulations. Where the researcher has a preferred formulation which precedes any empirical evaluation the analysis will be described as
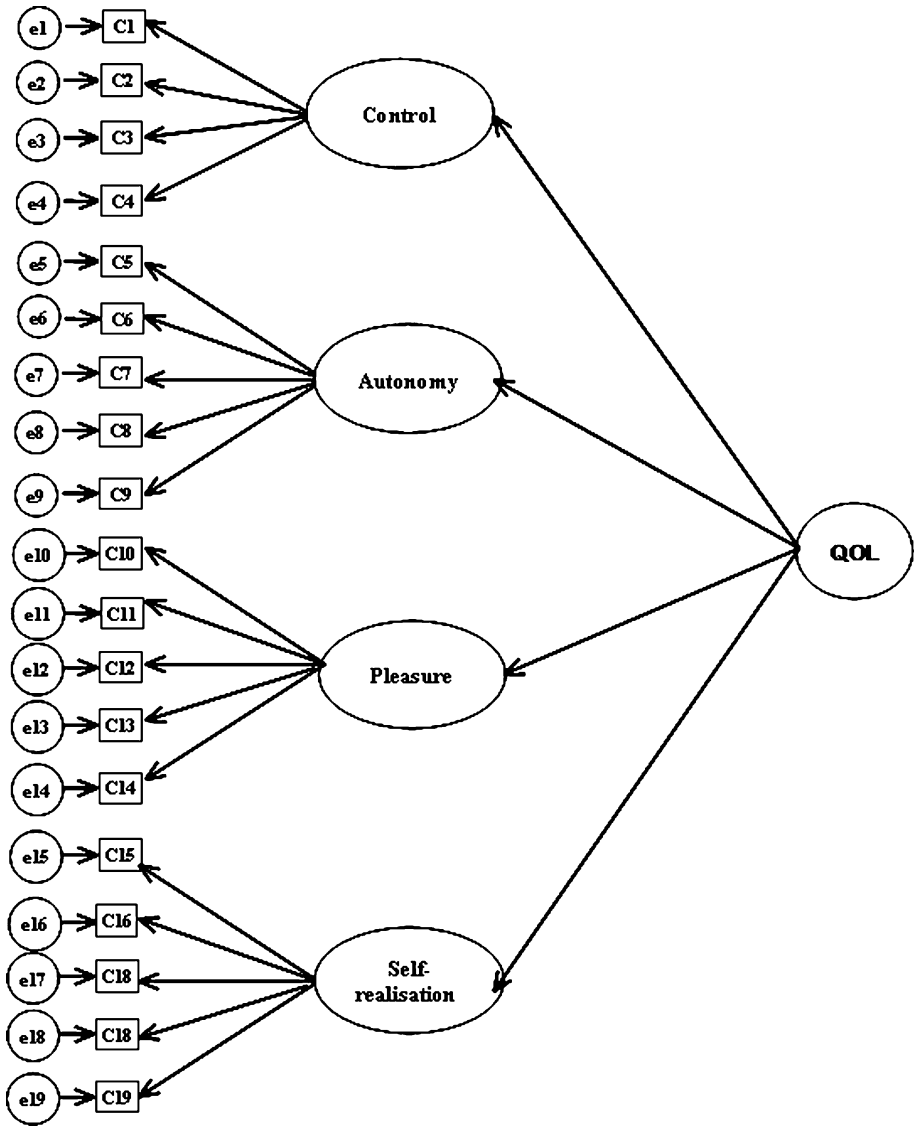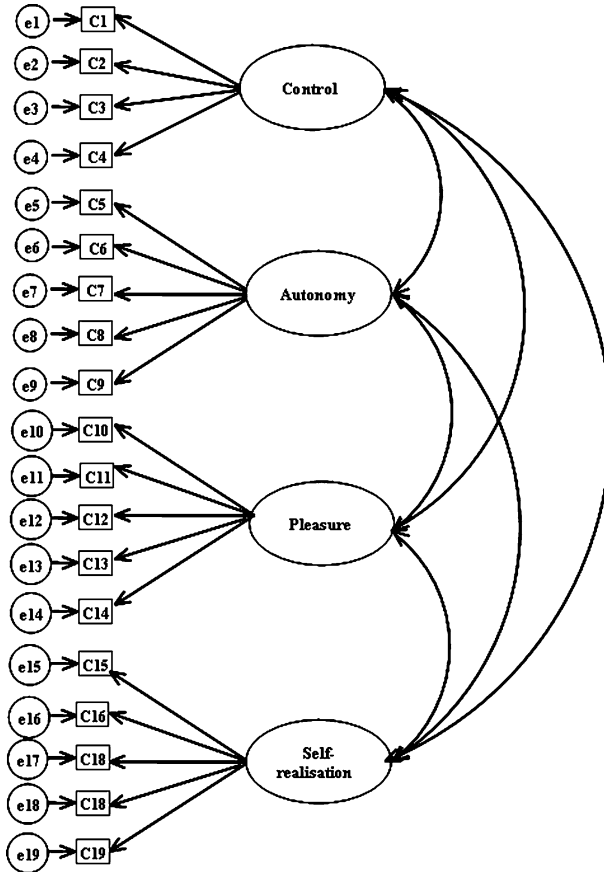
Fig. 2 Second order factor model for CASP-19

confirmatory. Otherwise, in the absence of a preferred model researchers may well adopt an exploratory approach. In such an instance the paths or connections between the latent variables which define the underlying structure and the manifest or indicator variables would be empirically determined.

In order to fully test our theory we consider another formulation in addition to the models already illustrated in Figs. 1 and 2. This third alternative is presented in Fig. 3 conveys a formulation where the researcher is most interested in the four life domain scores (control, autonomy, self-realisation and pleasure) and these four domains are simply allowed to inter-correlate rather than be dependent upon a single underlying factor (second

Fig. 3 A first order model for CASP-19

order) for quality of life. Essentially, this formulation is an intermediate alternative between the single factor formulation and the second order model and for this reason we present the results in the following order: the single factor model, the first order factor model and finally our theoretically preferred second order factor model. In factor analytic terms these models all represent competing accounts of the structure in the observed covariance or correlation matrix for the 19 CASP items. They are all confirmatory models in that various constraints are imposed on paths connecting latent and manifest variables as well as the inter-relationships between the latent variables themselves (Maxwell 1977). These three alternative formulations of the underlying structure can be assessed in empirical terms by their 'goodness of fit', i.e. how well the structure reproduces the observed correlations? To this extent there may not be a unique measurement model of quality of life. Our task in this paper is to examine firstly, whether or not our preferred model stands the test of empirical evaluation in relation to its close alternative formulations and secondly, whether a subsequent re-examination of the measurement properties lead to a more pragmatic formulation of the scale (i.e. a shorter version) which is both plausible and more parsimonious. The first part of the evaluation uses traditional confirmatory factor analysis (Maxwell 1977) and the second draws upon internal consistency analysis (McKennell 1977).

1.3 Analytical Strategy and Criteria for Evaluation

To map our theory of quality of life exactly on to a set of theoretical concepts suggest that the indicators or statements we choose to represent those concepts represent or convey the meaning of the underlying concepts as developed. To the extent that they are imperfect indicators of truth could be the result of the complex interaction of inadequate theory, poor piloting, imperfect operationalisation, respondent mood, personality, understanding and survey mode (Brenner et al. 1978; Biemer et al. 1991; Jabine et al. 1994; Eid and Diener 2004). For these reasons we have adopted a pragmatic approach to our evaluation of CASP-19. The rapid take-up of CASP-19 gave us a unique research opportunity to evaluate its properties across three research settings, the Boyd-Orr sample, the first wave of ELSA and wave 11 of the BHPS. The first research setting, namely Boyd-Orr, formed the context for our early development and testing of CASP-19. This seemed reasonable given its modest sample size (198 complete cases). As such the Boyd-Orr sample performed as a pilot study in our evaluation whereas, the inclusion of CASP-19 in BHPS and ELSA provided the opportunity to confirm our preferred structure across representative samples of the ageing population in 2002 for Britain (England, Wales and Scotland) and England. In both studies CASP-19 was administered as a self-report in the context of a face-to-face computer assisted interview. The assessment of CASP-19 described in Hyde et al. (2003) can be regarded a tentative appraisal of CASP-19 using a combination of exploratory factor analysis and internal consistency analysis (McKennell 1977). Once the pattern of item loadings confirmed the item membership of each domain, reliability analysis provided confirmation of the internal consistency of each domain (Cronbach's Alpha for control, autonomy, self-realisation and pleasure were, respectively: 0.60, 0.67, 0.73 and 0.78) and finally the intercorrelations between each of the four domain scores provided evidence to confirm their common dependence on a single underlying factor, 'quality of life'. Subsequent evaluations using ELSA_1 and BHPS_11 tested the three confirmatory approaches as illustrated in Figs. 1–3. Subsequent evaluation of the number items in life domain was then conducted in order to examine whether or not the final operational scale could be shortened. This latter objective is often seen to be attractive to survey methodologists who wish to economise on 'space' in questionnaires for self-enumeration scales without any resulting loss of breadth in measurement. The development of shortened versions of the General Health Questionnaire (Goldberg 1972; Goldberg and Williams 1998) from 60-items to 30-items to 12-items, similarly for the SF-36 physical and mental health measure which also has shortened 12-item versions (Ware et al. 1994, 1995, 2002) and the Depression-Happiness Scale (Joseph et al. 2004) from 25-items to 6-items all serve to emphasise this point.

Altogether our strategy maps on quite nicely to the practical scale evaluation procedures recommended by Muthén and Muthén (2005) where there is not necessarily a rigid distinction drawn between confirmatory and exploratory factor analysis. These authors describe the advantages of adopting an 'exploratory factor analysis' within a confirmatory framework. We consider our approach to be quite unique in that it has been possible to test the theoretical grounding which underwrites CASP-19 within two major national studies. This approach is quite distinct from evaluations which typically focus upon the evaluation of a well-known measure, for example the General Health Questionnaire (GHQ) and then carry out repeat exploratory analyses across different data sets (Campbell et al. 2003; Kalliath et al. 2004). We are primarily interested in testing the strength of our theory and, therefore, the appropriateness of our preferred model. We recognise that there will always be competing models to describe the associations in our data but prefer to test the

Table 2 Definitions and guide to evaluating the 'Goodness of fit indices' used to evaluate the latent structure of CASP-19

Index	Formula	Interpretation of fit
Deviance or CMIN/*df*	χ_1^2/df_1	>2.0 Not so good
Tucker Lewis index (TLI)	$\dfrac{(x_0^2/df_0)-(x_1^2/df_1)}{(x_0^2/df_0)-1}$	>0.90 Good
		>0.95 Very good
Comparative fit index (CFI)	$\dfrac{(\chi_0^2-df_0)-(\chi_1^2-df_1)}{(\chi_0^2-df_0)}$	>0.90 Good
		>0.95 Very good
Root mean square index of approximation (RMSEA)	$\sqrt{(\chi_1^2/df_1)}/(N-1)$	>0.10 Poor
		<0.05 Good

χ_0^2 Chi-squared for null model, df_0 degree of freedom for null model (all parameters set to zero)

χ_1^2 Chi-squared for specified model, df_1 degree of freedom for specified model

N Sample size

theoretical base of our model over empirical improvements in fit which may vary from one setting to another. In accord with Campbell and his colleagues (Campbell et al. 2003) 'it has to be recognised that confirmatory factor analysis cannot identify which model is the true model underlying the data. A number of different models would hypothetically fit the data well and the issue as to which is the 'better' model is one of parsimony and plausibility'.

The assessment of model fit in latent variable traditions has generated considerable debate as to which measures are most suitable (Marsh and Balla 1988; Bollen and Long 1993). Typically, measures of goodness of fit share a common objective to test how well a specified latent variable structure or measurement model actually reproduces the observed correlation or covariance matrix for the indicators used to define a scale. The difference of a function of the difference between the predicted and observed values is used to derive a chi-square value (or a function thereof) to assess fit. Unfortunately, raw chi-square values tend to be very large whenever a model is a poor fit or the sample size is large. To be consistent with a preferred set of criteria for assessing fit for the estimation algorithm adopted in these analyses (WLSMV) and recommended by Muthén and Muthén (2004) we have a adopted four criteria: a raw chi-square divided by its degrees of freedom (CMIN/df), the Tucker Lewis Index (TLI), the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The last three indices all adopt a convention of showing the relative improvement of some specified model (e.g. the second order factor model) compared to a null model where all of the parameters are set to zero. There are rules of thumb for judging the importance of any model and these are shown along with the algebraic summary of the indices in Table 2 below:

1.4 Sample Sources

To recap our sample sources derive from the Boyd Orr survey (Hyde et al. 2003) and subsequently abbreviated to 'Boyd Orr_2000', and respondents aged 55 years and above in the British Household Panel Survey wave 11 (BHPS_ 11) (Taylor et al. 2003), and the first wave of the English Longitudinal Study of Ageing (ELSA_1) (Marmot et al. 2003). All preliminary exploratory (or pilot) analyses were conducted for Boyd Orr_2000 followed by confirmatory analyses for ELSA_1 and BHPS_11. Analyses were conducted in Mplus

Table 3 Summary of the number of cases presented for analysis across the three research settings

Research Setting	Original sample size	Cases presented for analysis	Degree of missingness+ (%)[a]	Cases with *all* CASP items missing	% of original sample excluded from analysis
Boyd Orr-2000	294	198	0	33	14.4
ELSA_1	11,234	9,300	1.6	980	10.2
BHPS_11	7,480	6,471	1.2	513	7.9

[a] Percentage of empty cells in case by item (19) matrix

(Muthén and Muthén 2004) using full information maximum likelihood estimation (FIML) to handle any item non-response across CASP-19 (Asparouhov and Muthén 2003). Effectively, the application of the FIML procedure was constrained to those cases with responses to one or more CASP items. All individuals without any responses at all across CASP-19 items were omitted. This resulted in a very small degree of missingness to be handled ($<2\%$ for the matrices presented for analysis). Table 3 summarises the number of cases used in our analysis. FIML uses all of the available information in the case by item matrix for the CASP-19 scores. The procedure operates under the assumption that item non-response is assumed to be 'missing at random' (MAR). For further discussion of FIML in the context of latent variable modelling see Wiggins and Sacker (2002).

1.5 Resume of Evaluation

What follows is first a report of the preliminary analyses for CASP-19 using the Boyd Orr_2000 data and then confirmatory factor analyses for both ELSA_1 and BHPS_11. An overview of the performance of our measure of quality of life is provided in the context of our preferred second order factor model compared to single factor and first order models. All analyses use FIML to handle the problem of item non-response in CASP-19 and weighted least squares estimation (WLSMV)[2] as a recommended estimation procedure in Mplus for handling ordered categorical items as dependent variables. Finally, we present some recommendations for using CASP-19 or its shortened form, CASP-12.

2 Results

2.1 Preliminary Findings for the Boyd Orr_2000 Survey

Table 4 below provides a summary of taking each CASP-19 domain and subjecting the inter-item correlation matrices to separate exploratory factor analyses. In terms of the percentage of variance explained following a varimax rotation (Kaiser 1958) there is a modest degree of comfort that each domain can be considered in its own right. As mentioned above the values for Cronbach's Alpha for each sub-scale reach a degree of respectability (all above 0.60, Cronbach 1951). Subsequent exploratory factor analysis of the four domain scores provides evidence of a single (second order) factor to denote 'quality of life'.

---

[2] The full acronym refers to weighted least squares estimation using a diagonal weight matrix with standard errors and mean and variance adjusted chi-square test that uses a full weight matrix (see page 368 in Muthén and Muthén 2004).

Table 4 Results from an exploratory factor analysis treating each domain as a single factor for Boyd Orr_2000

Domain	Control	Autonomy	Self-realisation	Pleasure
% variation explained by a single factor	47.1	48.3	51.8	45.8
Cronbach's alpha	0.60	0.67	0.73	0.78

Being reasonably confident with our preferred second order model we set out to test the suitability of each of the three measurement models (as described in Figs. 1–3) across two national surveys, ELSA_1 and BHPS_11 for all individuals aged 55 years and above reporting one or more CASP items.

As expected the relatively large sample sizes render the deviances for both sources less than informative. Only, TLI indices suggest with some reassurance that there is again very little to choose between the first and second order models. What is remarkable is how similar the magnitude of the TLI and CFI criteria are for both national samples. Although the RMSEA coefficient behaves better for BHPS_11 compared to ELSA_1. In addition, simply correlating the magnitude of the factor loadings for each measurement model reveals very definite consistency across ELSA_1 and BHPS_11 (0.98 for all models). At this stage of the evaluation we have very little empirical evidence to prefer one measurement model over any other across these two national samples.

2.2 Further Evaluation of CASP-19 for the English Longitudinal Study of Ageing Wave 1 and the British Household Panel Survey Wave 11

The empirical evidence presented in Tables 5 and 6 above provides us with an argument to believe that there is no overwhelming reason to depart from the preferred measurement model presented in Fig. 2. However, the relative weakness of the CFI and RMSEA criteria suggest that as far as model fit is concerned we could do better! The obvious way to proceed is to examine the internal structure of the four conceptual domains and reflect upon our preferred model. Tables 7 and 8 provide a summary of the evaluations initially adopted in our appraisal of Boyd Orr (Table 4) by treating each life domain as a single factor.

It is clear form the estimated values of Cronbach's Alpha that the autonomy domain 'under performs'. Hence our next step was to carry out an internal consistency or reliability analysis to see whether or not the statistical fit across the two samples could be improved by reducing the number of items or indicators for each domain. In sum, the two items which least well correlated with the other items in each domain were candidates for exclusion. In practice, given that there were four items in the control domain we decided to only consider dropping one item from that domain and two items for the other domains in order to achieve a balanced 12-item version of CASP with three items per domain. Fortuitously, the candidate items for exclusion were the same for separate analyses of ELSA_1 and BHPS_11. These are listed in Table 9 below. Whilst the second order factor measurement model improved across all goodness of fit criteria the Cronbach Alpha coefficient for the autonomy scale remained stubbornly low at around 0.45 (see Tables 10 and 11 below).

From both a theoretical and empirical perspective the evidence for a four-domain model of quality of life was equivocal. A shortened, 12-item version[3] of the original scale had

---

[3] An earlier 12-item version of CASP was proposed on the basis of an analysis of Boyd Orr_2000 using AMOS (Arbuckle and Wotkhe 1999) and applied by Von dem Knesebeck et al. (2005) for the SHARE project mentioned in reference to Börsch-Supan et al. (2005). Ten of the items coincide with the version recommended here.

Table 5 Summary findings for confirmatory factor analysis of CASP-19 applying three measurement models in ELSA_1 (as illustrated in Figs. 1–3)

Measurement model	Deviance (CMIN/df)	TLI	CFI	RMSEA
Single factor	194.6	0.90	0.74	0.14
First order	150.3	0.92	0.80	0.12
Second order	177.4	0.91	0.76	0.13

Table 6 Summary findings for confirmatory factor analysis of CASP-19 applying three measurement models in BHPS_11 (as illustrated in Figs. 1–3)

Model	Deviance (CMIN/df)	TLI	CFI	RMSEA
Single factor	67.3	0.89	0.73	0.10
First order	53.5	0.92	0.79	0.09
Second order	59.6	0.91	0.76	0.09

Table 7 Results from an exploratory factor analysis treating each domain as a single factor for ELSA_1

Domain	Control	Autonomy	Self-realisation	Pleasure
% variation explained	48.3	35.5	54.2	59.8
Cronbach's alpha	0.63	0.51	0.78	.83

Table 8 Results from an exploratory factor analysis treating each domain as a single factor for BHPS_11

Domain	Control	Autonomy	Self-realisation	Pleasure
% variation explained	48.7	36.0	51.7	56.2
Cronbach's alpha	0.64	0.53	0.76	0.80

Table 9 Items excluded from CASP-19 following separate reliability analyses of each life domain across ELSA_1 and BHPS_11

Item	Statement
	Control
3 C	I feel free to plan for the future
	Autonomy
6 A[a]	Family responsibilities prevent me from doing the things I want to do
8 A[a]	My health stops me from doing the things I want to do
	Pleasure
13 P	I enjoy being in the company of others
14 P	On balance, I look back on my life with a sense of happiness
	Self-realisation
16 S	I choose to do things that I have never done before
17 S	I feel satisfied with the way my life has turned out

[a] Item reverse coded for scoring

Table 10 Goodness of fit criteria for a second order measurement model for a 12 item version of CASP across ELSA_1 and BHPS_11

Model fit indices	ELSA_1	BHPS_11
Deviance	73.9	24.9
CFI	0.938	0.942
TLI	0.969	0.971
RMSEA	0.084	0.059

Table 11 Cronbach's Alpha coefficients for key conceptual domains in CASP-12 across ELSA_1 and BHPS_11

Domain	ELSA_1	BHPS_11
Control	0.67	0.66
Autonomy	0.45	0.46
Self-realisation	0.77	0.76
Pleasure	0.80	0.77

Table 12 Model fit statistics for a second order measurement model based on three domains (control and autonomy are combined as a single domain) using CASP-19 and CASP-12 across ELSA_1 and BHPS_11

Sample and CASP version	Deviance	TLI	CFI	RMSEA
BHPS_11 CASP19	52.2	0.919	0.793	0.086
ELSA_1 CASP-19	147.0	0.923	0.803	0.119
BHPS_11 CASP12	36.8	0.957	0.913	0.072
ELSA_1 CASP12	110.2	0.954	0.910	0.103

improved the statistical fit for our preferred second order model but there were no strong grounds for having a distinct 'autonomy' domain. After further reflection and modelling we decided to combine the indicators for control and autonomy into a single domain. Our original operationalisation was not really sensitive enough to separate out the right of an individual to be free from the unwanted interference of others (autonomy) and their ability to actively intervene in their own environment (control) (Patrick et al. 1993). Items like 'I feel that what happens to me is out of my control' and 'I can do things I want to do' are best treated as indicators of a single domain. This revised second order measurement model performed better for CASP-12 across both samples (see Table 12). The alpha coefficient for 'control and autonomy' rose to 0.67.

3 Recommendations for Using the Scale

This evaluation of a self-completed multi-item scale to measure quality of life across two national samples of people aged 55 years and above in the UK has provided evidence to support its theoretical foundations. Our preferred model for identifying four domains of daily life as dependent upon a single unitary latent variable for quality of life is best adapted for analytical purposes by combining the original domains control and autonomy. Analyses may then proceed using combined summative indexes for 'control and

autonomy', 'self-realisation' and 'pleasure' as well as an overall index of quality of life. The latter agentic domains 'self-realisation' and 'pleasure' stood up well to assessment across all research settings. The shortened 12-item version of CASP has stronger measurement properties than the original CASP-19 measure and is recommended for future applications. However, whilst the single factor model (Fig. 1) was less robust than the first order and second order models Cronbach's Alpha for CASP-19 attained 0.87 for both samples and could well be used as a single summary index to compare population aggregates. CASP-12 is reproduced in the annex.

Annex

CASP-12 item wording arranged by domain categories

Original item number	Statements by domain
Control and autonomy	
1 CA[a]	My age prevents me from doing the things I would like to do
2 CA[a]	I feel that what happens to me is out of my control
4 CA[a]	I feel left out of things
5 CA	I can do the things I want to do
7 CA	I feel that I can please myself what I do
9 CA[a]	Shortage of money stops me from doing things I want to do
Pleasure	
10 P	I look forward to each day
11 P	I feel that my life has meaning
12 P	I enjoy the things that I do
Self-realisation	
15 S	I feel full of energy these days
18 S	I feel that life is full of opportunities
19 S	I feel that the future looks good for me

[a] Item reverse coded for scoring

References

Andrews, F. M., & Withey, S. B. (1976). *1976 Social indicators of well-being: Americans' perception of life quality*. New York: Plenum Press.

Arbuckle, J. L., & Wothke, W. (1999). *AMOS 4.0 user's guide*. USA: SmallWaters Corporation.

Asparouhov, T., & Muthén, B. O. (2003). Full-information-maximum-likelihood information estimation of general two-level latent variable models.

Baker, F. D., & Intagliata, J. (1982). Quality of life in the evaluation of community support systems. *Evaluation and Program Planning, 5*, 69–79.

Bartholomew, D., & Knott, M. (1999). *Latent variable models and factor analysis. Kendall's library of Statistics 7*, 2nd ed. London: Edward Arnold.

Berney, L., & Blane, D. (1997). Collecting retrospective data: Accuracy and recall after 50 years judged against historical records. *Social Science and Medicine, 45*, 1519–1522.

Biemer, P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., & Sudman, S. (eds.). (1991). *Measurement errors in surveys*. New York: Wiley.

Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*. Newbury Park, CA: Sage.

Börsch-Supan, A., Brugiavini, A., Jürges, H., Makenbach, J., Siegrist, J., & Weber, G. (eds.). (2005). *Health, ageing and retirement in Europe. First results from the survey of health, ageing and retirement in Europe (SHARE)*. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).

Bowling, A. (1995). What things are important in people's lives? A survey of the public's judgements to inform scales of health related quality of life. *Social Science and Medicine, Special Issue 'Quality of Life', 10*, 1447–1462.

Bowling, A. (1997). *Measuring health: A review of quality of life measurement scales*, 2nd ed. Buckingham: Open University Press.

Brenner, M., Marsh, P., & Brenner, M. (eds.). (1978). *The social contexts of method*. Croomhelm Ltd.

Brown, J., Bowling, A., & Flynn, T. N. (2004). Models of quality of life. A taxonomy and systematic review of the literature. Report commissioned by the European Forum on Population Ageing Research/Quality of Life, University of Sheffield (professor Alan Walker).

Bullinger, M. (2002). Assessing health related quality of life in medicine. An overview over concepts, methods and applications in international research. *Restorative Neurology and Neuroscience, 20*, 93–101.

Campbell, A., Walker, J., & Farrell, G. (2003). Confirmatory factor analysis of the GHQ-12: Can I see that again? *Australian and New Zealand Journal of Psychiatry, 37*, 475–483.

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

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research, 31*, 103–157.

Diener, E., & Suh, E. (1997). Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research, 40*, 189–216.

Cummins, R. A. (1996). The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research, 38*, 303–328.

Doyal, L., & Gough, I. (1991). *A theory of human need*. Hamphshire: Macmillan, Basingstoke.

Eid, M., & Diener, E. (2004). Global judgements of subjective well-being: Situational variability and long-term stability. *Social Indicators Research, 65*, 245–277.

Everitt, B. S. (1984). *An introduction to latent variable models*. London: Chapman and Hall.

Farquhar, M. (1995). Elderly people's definitions of quality of life. *Social Science and Medicine, Special Issue 'Quality of Life', 10*, 1439–1446.

Felce, D., & Perry, J. (1995). Quality of life: Its definition and measurement. *Research in Developmental Disabilities, 16*(1), 51–74.

Giddens, A. (1990). *The consequences of modernity*. Cambridge: Polity.

Gilleard, C., & Higgs, P. (2000). *Cultures of ageing*. London: Prentice Hall.

Goldberg, D. (1972). *The detection of psychiatric illness by questionnaire*. London: Oxford Press Monograph.

Goldberg, D., & Williams, P. (1988). *The user's guide to the general health questionnaire*. Windsor: NFER-NELSON.

Gunnell, D., Davey-Smith, G., Frankel, S., Nanchahal, K., Braddon, F., & Peters, T. (1996). Lifecourse exposure and later diseases: A follow-up study based on a survey of family diet and health in pre-war Britain (1937–1939). *Public Health, 110*, 85–94.

HAPiEE Study (2005). UCL Department of Epidemiology and Public Health. Health, Alcohol and Psychosocial factors in Eastern Europe. http://www.ucl.ac.ul/easteurope/happie_open.htm. Accessed 7/12/2005.

Higgs, P., Hyde, M., Wiggins, R. D., & Blane, D. (2003). Researching quality of life: The importance of the sociological dimension. *Social Policy and Administration, 37*, 239–252.

Hyde, M., Wiggins, R. D., Blane, D., & Higgs, P. (2003). A measure of quality of life in early old age: The theory, development and properties of a needs satisfaction model. *Aging and Mental Health 7*, 86–94.

Jabine, T., Straf, M., Tanur, J., & Tourangeau, R. (eds.). (1994). *Cognitive aspects of survey methodology: Building a bridge between the disciplines*. Washington, DC: National Academy Press.

Joseph, S., Linley, P. A., Harwood, J., Lewis, C. A., & McCollam, P. (2004). Rapid assessment of well-being: the short depression-happiness scale (SDHS). *Psychology and Psychotherapy: Theory, Research and Practice, 77*, 463–478.

Kalliath, T. J., O'Driscoll, M. P., & Brough, P. (2004). A confirmatory factor analysis for GHQ-12. *Stress and Health, 20*, 11–20.

Kaiser, H. F. (1958) The varimax criteria for analytic rotation in factor analysis. *Psychometrika, 23*, 187–200.

KLoSA (2007). http://klosa.re.kr/KLOSA/html_kor/klosaMeeting/KLoSA_2nd_meeting.pdf. Accessed 30/10/07.

Von dem Knesebeck, O., Hyde, M., Higgs, P., Kupfer, A., & Siegrist, J. (2005). Quality of life and well-being. In A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist, & G. Weber (eds.), *Health, ageing and retirement in Europe: First results from the survey of health, ageing and retirement in Europe* (pp. 199–203). Germany: Strauss, Mannheim.

Laslett, P. (1996). *A fresh map of life*, 2nd ed. London: MacMillan.

Marmot, M., Banks, J., Blundell, R., Lessof, C., & Nazroo, J. (2003). *Health, wealth and lifestyles of the older population in England: The 2002 English longitudinal study of ageing*. London: Institute for Fiscal Studies, University College.

Marsh, H. W., & Balla, J. R. (1988). Goodness of fit indices in confirmatory factor analysis: the effect of sample size. *Psychological Bulletin, 103*(3), 391–410.

Maxwell, A. E. (1977). *Multivariate analysis in behavioural research*. London: Chapman and Hall.

McKennell, A. (1977). Attitude scale construction. Chapter 7, Vol 1. *Exploring Data Structures*, The Analysis of Survey Data.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus. Statistical analysis with latent variables*. Los Angeles: User's Guide. Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2005). *Mplus short course notes day 1: Ttaditional latent variable modelling using Mplus*. Copyright Muthén & Muthén. http://www.statmodel.com.

Netuveli, G., Blane, D., & Bartley, M. (2007). Wellbeing and welfare states: Cross-national comparison of quality of life in market and transition economies. http://www.bath.ac.uk/econ-dev/wellbeing/conference2007/final-papers/5-irq/Netuveli Blane Bartley.pdf. Accessed 30/10/07.

Office for National Statistics (ONS) (1998). *Social focus on older people*. ONS, London.

Patrick, B. C., Skinner, E. U., & Connell, P. J. (1993). What motivates children's behaviour and emotion in the academic domain? *Journal of Personality and Social Psychology, 65*, 781–791.

Phillipson, C. (1998). *Reconstructing old age. New agendas in social theory and practice*. London: SAGE.

Scase, R., & Scales, J. (2000). *Fit and fifty*. Swindon, Wiltshire: Economic and Social Research Council.

Smith, A. E. (2000). Quality of life: A review. *Education and Ageing, 15*(3), 419–435.

Taylor, M. (ed.) with Brice, J., Buck, N., & Prentice-Lane, E. (2003). *British household panel survey user manual volume A. Introduction, Technical Report and Appendices*. Colchester, Essex: UK Data Archive, University of Essex.

Turner, B. (1995). Ageing and identity: some reflections on the somatisation of the self. In M. Featherstone, A Wernick (eds.), *Images of aging* (pp. 245–260). London: Routledge.

Vincent, J. (2003). *Old age*. London: Routledge.

Ware, J. E., Gandek, B., the IQOLA Project Group (1994). The SF-36Ⓡ Health Survey: Development and use in mental health research and the IQOLA Project. *International Journal of Ment Health, 23*(2), 49–73.

Ware, J. E., Kosinki, M., & Keller, S. D. (1995). *SF-12Ⓡ: How to score the SF-12Ⓡ physical and mental health summary scales*, 2nd ed. Boston: The Health Institute, New England Medical Center.

Ware, J. E., Kosinki, M., Turner-Bowker, D. M., & Gnadek, B. (2002). How to score Version 2 of the SF-12Ⓡ: the SF-12Ⓡ Health Survey (with Supplement documenting Version 1). Lincoln: QualityMetric Incorporated.

Wiggins, R. D., Bynner, J. (1993). 'Social attitudes', in Life at 33, Chapter 8. In: Ferri E (ed.), National Children's Bureau, London, pp 75–92.

Wiggins, R. D., & Sacker, A. (2002). A comparative evaluation of strategies to handle missing data in the context of structural equation modelling. Chapter 6, With Sacker, A. In G Marcoulides, I Moustaki (eds.), *Latent variable and latent structure models*. London: Erlbaum and Laurence.

Wiggins, R. D., Erzberger, C., Hyde, M., Higgs, P., & Blane, D. (2007) Optimal matching analysis using ideal types to describe the lifecourse: an illustration of how histories of work, partnerships and housing relate to quality of life in early old age. *International Journal of Social Research Methodology, 10*(4), 259–278.

Young, M., & Schuller, Y. (1991). *Life after work: The arrival of the ageless society*. London: Harper Collins.