THE UNIVERSITY
of EDINBURGH

Exploring what we know about retirement:

A systematic review of constructs and measures of adjustment and adaptation to retirement

and

A meta-analysis of the relationship between retirement and depression in later life

Claudia Coelho

Presented in Partial Fulfilment of the Requirements

for the Degree of Doctorate in Clinical Psychology

University of Edinburgh

January 2014

Word Count: 29783

D. Clin. Psychol. Declaration of own work

This sheet must be filled in (each box ticked to show that the condition has been met), signed and dated, and included with all assessments - work will not be marked unless this is done

Name: Claudia Coelho

Assessed work Case Study SSR Essay Question Paper Thesis

Title of work:

Exploring what we know about Retirement: A systematic review of constructs and measures of adjustment and adaptation to retirement *and* A meta-analysis of the relationship between retirement and depression in later life

I confirm that all this work is my own except where indicated, and that I have:

Read and understood the Plagiarism Rules and Regulations in the Programme Handbook	☑
Composed and undertaken the work myself	☑
Clearly referenced/listed all sources as appropriate	☑
Referenced and put in inverted commas any quoted text of more than three words (from books, web, etc)	☑
Given the sources of all pictures, data etc. that are not my own	☑
Not made undue use of essay(s) of any other student(s) either past or present (or where used, this has been referenced appropriately)	☑
Not sought or used the help of any external professional agencies for the work (or where used, this has been referenced appropriately)	☑
Not submitted the work for any other degree or professional qualification except as specified	☑
Acknowledged in appropriate places any help that I have received from others (e.g. fellow students, technicians, statisticians, external sources)	☑
Complied with other plagiarism criteria specified in the Programme Handbook	☑
I understand that any false claim for this work will be penalised in accordance with the University regulations	☑

Signature …Claudia Coelho… **Date** …31/01/14…

Please note:

a) If you need further guidance on plagiarism, you can: i/ Speak to your director of studies or supervisor; ii/ View university regulations at http://www.acaffairs.ed.ac.uk/Administration/GuidanceInformation/AcademicBestPractice/Plagiarism/Index.htm

b) Referencing for most assessed work should be in the format of the BPS style guide, which is freely available from the BPS web site

This is for Pete.

Acknowledgements

First, I wish to thank my supervisor, Dr Emily Newman, for her support and advice. I would also like to thank Dr Paul Hutton for his help at the final stages of the meta-analysis. My thanks also go to James, Fiona and Toby at Information Services (University of Edinburgh) for rescuing this thesis when it was nearly lost. I want to thank David, for being my friend, for all the wonderful chats and for his contribution to this thesis as an independent reviewer. Then there is Margi, Tara and Sarah, who I want to thank for keeping me in their minds and in their hearts. Above all, I want to thank Pete. Thank you for putting up with all this. Thank you for making it all all-right, all over again. Now, you and me, we get back to life!

Thesis Abstract

Objectives:

This thesis includes two distinct pieces of work. A systematic literature review (SLR) aimed to identify, organise and evaluate constructs and measures of adjustment and adaptation to retirement (AAR), as these are currently used in the literature. A meta-analysis (MA) sought to systematically evaluate the evidence of the relationship between retirement and self-reported symptoms of depression in later life.

Methods:

For the SLR, a comprehensive list of electronic databases, and additional sources, were searched (March-May 2013). The identified constructs of AAR were evaluated in relation to their content, clarity and frequency of use. The identified measures were evaluated in relation to pre-defined psychometric properties and frequency of use. For the MA, an equally comprehensive search strategy was used (December 2012-March 2013). A meta-analysis and systematic assessment of risk of bias were carried out on the studies eventually included.

Results:

In the SLR, 27 measures of AAR were put forward for evaluation, and 6 distinct clusters of constructs of AAR were identified. AAR is defined inconsistently in this literature. Most measures that are used to assess these constructs are adapted or imported from other contexts (e.g. mental health, well-being, life-satisfaction). The 7 measures that related specifically to AAR ('retirement satisfaction and role adjustment') lacked detailed psychometric information.

Eight non-randomised studies were included in the MA, 5 cohort studies and 3 cross-sectional studies. Studies were grouped and analysed according to these two design-type subgroups There was evidence of high dispersion of effect sizes, variable risk of bias and methodological and statistical heterogeneity between studies in both sub-groups – cohort (Q=640.728, df =4, p<0.001), cross-sectional (Q=76.611, df=2, p<0.001). Summary effects were therefore not meaningful. Sensitivity and sub-group analyses did not account for high heterogeneity of effect sizes.

Conclusions:

The SLR concluded that the variability in outcomes of research on AAR found in this literature may be underpinned, in part, by the different constructs and measures that are used. The 27 measures evaluated did not seem, at face-value, to measure the same construct; their psychometric properties also varied. The interpretation of outcomes, and comparisons between studies, in this area is hindered by this inconsistency.

The MA concluded that the relationship between retirement and self-reported depressive symptoms seems to be complex and variable. Effect-sizes of individual studies were small, non-significant and highly dispersed, and heterogeneity of true effects was high. These results may be limited by confounding factors in primary studies. This is discussed and contextualised in relation to the use of non-randomised studies in meta-analysis.

Word count: 412

Declaration of interest: None

Table of Contents

Tables

Figures

Thesis preface and structure

Content

This thesis is about research on retirement. Retirement is defined as a normative withdrawal from the work force in late adult life. However, from a psychological perspective, it is also a significant transition in late-life development.

The thesis is in two sections. The first is a systematic literature review that focuses on how adjustment and adaptation to retirement has been conceptualised and measured. The second is a meta-analysis of non-randomised studies that explores the relationship between retirement and self-reported symptoms of depression, in later life.

Both sections follow the general content guidelines for systematic reviews produced by the Cochrane Collaboration (Clarke, Oxman, et al., 2011), and specifically those guidelines for reviews of Non-Randomised Studies (Reeves, Deeks, Higgins, Wells, 2011). The meta-analysis also conforms to the Meta-Analysis of Observational Studies in Epidemiology Guidelines (Stroup, Berlin, et al., 2000).

Format

Both sections follow the publication guidelines of the journal *Clinical Psychology Review*. However, their full adaptation into the shape of a journal article would require adjustments in formatting and length, which, at this point, would impact on the readability of this piece of work as a thesis.

Additional material related to the systematic literature review (list of references and appendices) is included immediately after the main text of this section. Additional material related to the meta-analysis (as above) is included immediately after the main text of the second section. The guidelines for publication in *Clinical Psychology Review* (Appendix A) are included at the end of the full thesis.

References:

Clarke M., Oxman, A.D., Paulsen, E., Higgins, J.P.T., Green, S. (2011). Appendix A: Guide to the contents of a Cochrane Methodology protocol and review. In, Higgins, J.P.T., Green, S. (editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* (updated March 2011). Retrieved on 25th of September 2012 from www.cochrane-handbook.org

Reeves, B.C., Deeks J.J., Higgins, J.P.T., Wells, G.A. (2011). Chapter 13: Including non-randomized studies. In, Higgins, J.P.T., Green, S. (Editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A., Thacker, S.B. (2000). Meta-Analysis of Observational Studies in Epidemiology: A Proposal for Reporting. *Journal of the American Medical Association*, vol. 283(15), p. 2008-2015.

Exploring what we know about retirement: A systematic review of constructs and measures of adjustment and adaptation to retirement

Authors: Claudia Coelho[ab], Emily Newman[a]

[a] Section of Clinical and Health Psychology, School of Health in Social Science, University of Edinburgh, UK

[b] NHS Lothian

Address correspondence to:
Claudia Coelho,
Royal Edinburgh Hospital,
Morningside Terrace
Edinburgh

Word Count: 11290

Abstract

Objectives: To systematically review, organise and evaluate constructs and measures of adjustment and adaptation to retirement (AAR), currently used in the literature.

Methods: A comprehensive list of electronic databases, and additional sources, were searched from December 2012 to May 2013. The identified constructs of AAR were evaluated in relation to their content, clarity and frequency of use. The identified measures were evaluated in relation to pre-defined psychometric properties and frequency of use.

Results: 27 measures of AAR were evaluated, and 6 distinct clusters of constructs of AAR were identified. AAR is defined inconsistently in this literature. Most measures that are used to assess these constructs are adapted or imported from other contexts (e.g. mental health, well-being, life-satisfaction). The 7 measures that related specifically to AAR ('retirement satisfaction and role adjustment') lacked detailed psychometric information.

Conclusions: Outcomes of research on AAR vary greatly. This variability may be underpinned, in part, by the different constructs and measures that are used. The 27 measures evaluated here did not seem, at face-value, to measure the same construct. Their psychometric properties also varied. The interpretation of outcomes, and comparisons between studies, in this area is hindered by this inconsistency.

Declaration of interest: None

Word count (200 max.): 195

Key words (6 max.):

RETIREMENT, ADJUSTMENT, ADAPTATION, CONSTRUCT, MEASURE, SYSTEMATIC

1. Introduction

1.1 Retirement – definition and evolution

In most industrialised and developing societies, retirement from work constitutes a crucial transition in late adult life (e.g. Atchley, 1982a; Wang, Henkens, van Solinge, 2011; Kubicek, Korunka, Raymo, Hoonakker, 2011). Retirement is typically defined as '*later life withdrawal from the work force*' (e.g. Moen, 1996:131). It is, in itself, a relatively recent concept or social institution (Marshall and Taylor, 2005; Higo and Williamson, 2008; Warner, Hayward, Hardy, 2010). In Europe, the notion of retirement and the establishment of public pension systems (late 19th century) came as a form of intergenerational redistribution of wealth in support of older workers (Atchley, 1982b). In America, the establishment of retirement (and mandatory retirement ages), was only achieved later, in the 1930s (Warner et al., 2010). Retirement has since been transformed into a social institution and, therefore, a normative and expected event in an individual's life-cycle (Atchley, 1982b). In most industrialised or developing societies, now well established social and economic policies, entitle individuals to receive an old-age related pension, and thus exit the work force, commonly around the age of 65 or 60 (Organisation for Economic Co-operation and Development, OECD, 2011).

As a continuously evolving institution or process, more recently, the changing, extending and variable nature of contemporary work-life has rendered the above uniform definition of retirement somewhat simplistic (van Solinge and Henkens, 2009) – *'the actual decision to retire, as opposed to the right to obtain a pension, depends on more than the pensionable official retirement age'* (OECD, 2011:60)[1]. Age is no longer the single or main marker of retirement (Guillemard and Rein, 1993; Vickerstaff and Cox, 2005). Retirement has evolved into a variable and somewhat uncertain event (Han and Moen, 1999). For instance, labour policies in the UK and many European countries, during the 1980s and 90s, used early retirement incentives[2] to target workers in their 50s and 60s (Gautier, 2002; Cooke, 2006; Brugiavini, Croda, Mariuzzo, 2005). This created, an '*early retirement trap*' (Angelini, Brugiavini, Webber, 2009:464), a socio-economic problem of unused capacity in later life, which also had a deep impact on the financial sustainability of pension schemes, and, more widely, on contemporary constructs of age and ageing: '*as if a new age emerged (from around age 50 to 70) a time of life when people are old without being elderly'* (Gaultier, 2002:165). More recently, however, this trend is being reversed. During the last decade, most industrialised societies have adopted

[1] As such, a global characterisation of retirement timing is less than clear cut. The OECD (2011) and the UK's Office for National Statistics use an estimation of how many individuals are retired at any given point in time, through an indirect measure of '*withdrawal from the labour force'* (Mitchell and Guled, 2010:). This is derived from observed participation rate changes over a five-year period for successive cohorts of workers (by five-year age groups), aged 40 and over. '*It treats all of above the age of 40 as retired, if they are not in the labour force*' (OECD, 2011:60).

[2] As part of social and economic measures designed to counteract high youth unemployment.

measures designed to increase the employability of older workers and encourage them to remain in employment for longer (Leber and Wagner, 2007; Clarke, Marshall, Weir, 2012). These proposed reforms, broadly, aim to abolish a default retirement age, raise pension-eligibility age, increase the economic activity of older individuals, reduce state benefits or encourage a shift in responsibility of pension provision from the state to the individual (Shills, 2008; Brugiavini et al., 2005; Warner et al. 2010; The United Kingdom Government, 2011; Behncke, 2012)[3]. The wider determinants of this reversal are consistently linked to increased longevity, ageing of populations and the affordability of pensions (e.g. Loretto, 2010; Eichhorst, Gerard, et al., 2011), however, the discussion of these matters is outwith the scope of this review.

Contemporary retirement is an increasingly complex process, heterogeneous in timing, permanency and pathways (Clark and Quinn, 2002; De Vaus, Wells, Kendig, Quine, 2007; Nordenmark and Stattin, 2009; Warner et al., 2010). It encompasses multiple and transitional states of participation or cessation of work-life (Ekerdt, 2010), including: phased retirement; bridge employment, i.e. exit and re-entry into work, between career employment and complete retirement, (Cahill, Giandrea, Quinn, 2006); un-retirement, i.e. re-entry into the labour force, typically, after 2 years of early retirement (Higo and Williamson, 2008).

Despite this increasing institutional heterogeneity, a large proportion of workers in industrialised or developing societies expect (or even look forward) to retire. Retirement is still perceived, generally, as a normative transition (in contrast with job loss or unemployment) at a relatively defined time in late adulthood (Moen, 1996; Szinovacz and DeViney, 1999; Hardy, 2002; van Solinge and Henkens, 2007). Shared ideas about the age-appropriateness of transitions are necessarily socially constructed (e.g. Kohli and Künemund, 2002; Vickerstaff and Cox, 2005). Age markers are still, therefore, meaningful to individuals (Elder, 1994; Moen, 1996), and deviations from a more or less expected progression of working life tend to be constructed in an age-normative way, such as notions of "early" or "late" retirement (Kohli and Künemund, 2002; Shapiro and Yarborough-Hayes, 2008; van Solinge and Henkens, 2009; Schalk, van Veldhoven, et al., 2010).

1.2 Theoretical context – a field of multiple theories

The institutional heterogeneity of retirement described above finds a good fit in the theoretical context in the field of studies of retirement. Theories are not regularly used by authors in this field in a predictive or explanatory way (Bem and de Jong, 2002; Löckenhoff, 2012), fulfilling more of a contextualising function. The term "theories of retirement", therefore, is used here with this caveat.

[3] For instance, in the UK, the state pension eligibility age will gradually be raised for women from 60 to 65 by 2020, and thereafter, for both men and women, from 65 to 68 between 2024 and 2046 (Behncke, 2009). Recent forecasts place the date for this increase around the mid-2030s, and further increases up to 69 by the late 2040s (The United Kingdom Government, 2013).

Overall, these theories seem to be specific embodiments of larger theoretical frameworks on ageing and development. Broadly, these attempt to model a) the impact that retirement may have on individuals, or b) how individuals adapt or adjust to retirement. These will be addressed in turn.

Disengagement theory (Cumming and Henry, 1961, cited in Higo and Williamson, 2008), proposes that retirement is part of an expected and accepted process of individuals' disengagement with society, as ageing takes place. Retirement is conceptualised as a process of withdrawal from a familiar external world, with loss of identity and role, and potential crisis (Crawford, 1972). Activity theory (Havinghurst, 1961) was formulated as a response to disengagement theory. It suggests that, in the transition from middle to old age, individuals strive to maintain activities and attitudes as constant as possible (Havinghurst, 1961:8). The ideas of flexibility in the relinquishing of the worker role and the eventual adjustment or adaptation to other available roles (husband, carer, grand-mother), begin to take shape. Continuity theory considers maintenance of identities to be of critical importance (Higo and Williamson, 2008). Retirement is seen as a life-transition which can be potentially disruptive to individuals' internal and external structures (Atchley, 1989), and which triggers an adaptation response (Atchley, 1971). Atchley contributed significantly to the explicit modelling of retirement adjustment or adaptation, proposing a widely accepted, though not strongly empirically supported (Wang et al., 2011), stage model of adaptation to retirement. This emphasises retirement's temporal nature, hypothesising the following stages of adjustment: preretirement, honeymoon, disenchantment, reorientation, stability, and terminal stage (Atchley, 1976, cited in Wang, et al., 2011).

Role theory defines retirement as a role exit and entry, from worker to retiree, placing work, again, in a central spot in the construction of individuals' identity (Kim and Moen, 2002). It explicitly predicts that, on the one hand, the loss of work, and of the social environment associated with it, may leave individuals vulnerable to psychological distress. On the other hand, the disengagement with the work-role may '*serve to reduce role strain and overload, thereby enhance psychological well-being'* (Kim and Moen, 2002:213; Vandewater, Ostrove, Stewart, 1997). Role theory is currently a popular theoretical system in retirement literature, often dovetailed with life-course or life-span developmental perspectives (Quick and Moen, 1998; Kim and Moen, 2002; Elder, Johnson and Crosnoe, 2003). Together, these have provided a framework for the modelling of retirement as a transition, which is embedded within the individuals' (present and past) personal characteristics, pathways and social, temporal and environmental contexts (Szinovacz and Davey, 2004).

From a developmental perspective, retirement fits the construct of transition well (e.g. Szinovacz, 1980; Fouquereau, Fernandez, Fonseca, Paul, Uotinen, 2006; Silver, 2010), i.e. *"a discontinuity in a person's life of which he/she is aware and which requires new behavioural responses"* (Hopson and Adams, 1976:24, cited in Luhmann, Hoffman, Eid, Lucas, 2012:594). Retirement marks a discontinuity from a status (working) to another status (retired), which can be experienced as a slow, fast, continuous or discrete. Elder et al. (2003) and Löckenhoff (2012) consider that life-span

developmental models are useful as integrating frameworks, viewing retirement in very much the same way as other transitions (e.g. parenthood, bereavement, etc.) (Calvo and Sarkisian, 2011). A further and final addition to this framework suggests that the construct of 'resources' (perceived or actual) is key to understanding or modelling quality of adjustment or adaptation to retirement (Wang, 2007; Wang et al., 2011).

As an object of theory, retirement has, therefore, tended to be treated as a temporal phenomenon, implying different tasks to be engaged with at different times (Floyd, Haynes et al., 1992; Shultz and Wang, 2011). Retirement has become progressively understood as an individual and psychological phenomenon, conceptualised in terms of externalised action or internalised representation/processing, adjustment or adaptation (e.g. van Solinge and Henkens, 2007, Sargent, Bataille, Vough, Lee, 201; Shultz and Wang, 2011). As an adult life-transition, it is recognised as: a) multidirectional and involving a series of changes; b) a major life-event, with enduring consequences, triggering a series of internal reorganizations, namely, social roles, identity, goals, expectations and sources of pleasure/reward (Floyd et al., 1992); and c) contextualised in/by past and present individual experiences (Floyd et al., 1992).

1.3 Empirical context – issues of operationalization and measurement

Research on retirement has broadly developed along three directions, cutting across various fields of social and medical sciences (e.g. psychology, sociology, epidemiology or economics). The first direction has focused on how retirement, as a process or a decision, is predicted by a multitude of factors, such as physical and mental health, financial status, job/career characteristics, or larger socio-political and economic determinants (e.g. Atchley, 1979; Topa, Moriano, Depolo, Alcover and Morales, 2009; Sargent-Cox, Anstey, Kendig, Skladzien, 2012). The second, has explored the effect that retirement has on many aspects of individuals' lives, such as physical health, mental health, social functioning or marital quality (e.g. Kim and Moen, 2002; Szinovacz and Davey, 2004). The third direction has investigated variations and determinants of individuals' adjustment or adaptation to retirement (Atchley, 1979; Wang, 2007; van Solinge and Henkens, 2008). The present review is concerned with studies located in this latter direction of retirement research. It specifically focuses on how adjustment or adaptation to retirement have been operationalized and measured in this body of literature, using constructs like wellbeing or satisfaction. It is, therefore, important to define these terms, and briefly outline their use specifically in the retirement literature.

1.3.1 Wellbeing and neighbouring or component constructs

Wellbeing is a common outcome measure for research studies in a variety of areas (McDowell, 2010; Schrank, Bird, Tylee, Coggins, Rashid, Slade, 2013). World Health Organisation's (WHO) definition of health emphasises *'complete physical, mental and social well-being'* (WHO, 1946), aiming to

represent something broader of an individual's experience, beyond the presence or absence of distress. Wellbeing has multiple incarnations and associated measures (economic wellbeing, focusing on general quality of life, in reference to social determinants; physical wellbeing, measured by quality of life, in reference to illness or disorder). Psychological wellbeing is conceptualised as a subjective phenomenon, measured in reference to dimensions of affect or self-actualisation (Schrank, et al., 2013), or ideas of satisfaction, happiness and optimal functioning (McDowell, 2010). As a construct subjected to measurement, however, wellbeing poses complex problems (e.g. Neugarten, Havinghurst, Tobin, 1961;McDowell, 2010). This is an area of profound disagreement as to the specific structure or components of the construct and, consequently, its measurement (Gallagher, Lopez, Preacher, 2009).

Overall, psychological wellbeing seems to be defined in three related (i.e. not empirically independent), but, conceptually distinct ways: subjective or hedonistic wellbeing (Diener, 1984), related to the maintenance of pleasure and satisfaction; eudaimonistic wellbeing (Ryff, 1989), related to the pursuit of personal growth, meaning and self-actualisation; and social well-being (Keyes, 1998), related to optimal functioning within social worlds (Gallagher et al., 2009; McDowell, 2010). Subjective well-being (SWB) has been perhaps the most amply researched of these constructs in psychological literature (Diener, Eunkook, Lucas, Smith, 2009; Busseri and Sadava, 2011). Particularly, this has been used in research on adjustment or adaptation to life-events, illness, trauma, etc. (Kim-Prieto, Diener, Tamir, Scollon, Diener, 2006; Schrank et al., 2013). As it was proposed by Diener (1984), SWB represents a global, subjective evaluation that individuals make of their own lives, defined by three primary components – life satisfaction, presence of positive/pleasant affect, and absence of negative/unpleasant affect (Busseri and Sadava, 2011; Luhmann et al., 2012). However, the various constructs of wellbeing are frequently used interchangeably (e.g. SWB, happiness, satisfaction, psychological wellbeing or mental health). Kim-Prieto, et al. (2006), Gallagher, Lopez, Preacher (2009) and McDowell (2010) argue that this conceptual diversity is also related to the existence of various measures of well-being, which have relatively modest correlations between them (Kim-Prieto, et al., 2006). In the area of research on adjustment or adaptation to retirement, inevitably, the use of such constructs is equally variable. These are used to represent something broad about the retirement experience (e.g. satisfaction with retired life), or something quite specific (namely, the presence or absence of symptoms of mental ill-health) (e.g. Butterworth, Gill, et al., 2006; Wang, 2007; Wang and Bodner, 2007; Calvo, Haverstick, Sass, 2007).

1.3.2 Adjustment or adaptation

The concepts of adjustment and adaptation appear in various context of psychological research, to describe both a status and a process (Luhmann et al., 2012). Although these terms signify different constructs, these are used interchangeably in the literature on retirement. Therefore, from this point on in this review, the two terms – adjustment and adaptation – will be used as equivalent terms.

In studies of adjustment or adaptation to life-events (such as retirement), these terms characterise a state where an individual's '*level of subjective wellbeing exceeds a specific criterion'* (Luhmann et al., 2012:594). As a process, adjustment or adaptation is described as a trajectory of subjective wellbeing over time, in relation to an initial life-event (Kim-Prieto et al., 2006; Lucas, 2007) – the event triggers a psychological and/or physiological response, which decreases over time, as responsiveness diminishes, and the level of subjective wellbeing returns (or not) to its hypothesised pre-event level (Lucas, Clark, Georgellis, Diener, 2003; Lucas, 2007). Both long-term life circumstances and individual/sudden events can influence wellbeing (Biswas-Diener, Diener, 2001), however, the greatest adjustment or adaptation demands seem to be related to individual/sudden life-events (Kim-Prieto, et al., 2006).

Adjustment or adaptation to retirement, as the topic of interest in this review, seems to sit at the intersection of (at least) three wide trajectories of psychological constructs and research: a) research on ageing and the life-long developmental transitions that come with it (Baltes, 1987; Baltes, Staudinger and Linderberger, 1999); b) research on well-being, as a construct with the potential to be operationalised and measured; and c) research on specific issues of activity, work and occupation and their meaning to individuals' lives. Retirement, as a transition, is hypothesised to trigger significant changes in levels of well-being (Atchley, 1976, cited in Wang et al., 2011; Wang, 2007). Although there is great interest (both theoretical and practical) in better understanding adjustment to retirement, predictably, there are very diverse constructs, operationalisations and measurements of this process and its outcomes (Robbins, Payne, Chartrand, 1990; Price, 2003). Indicators such as retirement satisfaction, quality, happiness, wellbeing, or morale have all been used in this sense. Attempts to define and to measure outcomes of adjustment or adaptation in retirement are therefore diverse and problematic (Neugarten et al., 1961).

1.4 Rationale for this review

Retirement, it is widely accepted, carries consequences for psychological wellbeing (e.g. Fouquereau et al., 2006; Wang et al., 2011). It is also accepted that, at this point in time, when radical changes to the institution of retirement (namely, the extension of working lives) are taking place, an understanding of this process, its correlates and consequences is important (e.g. Westerlund, Vahtera, et al., 2010; Behncke, 2012). However, as many researchers in this field have observed, retirement is not particularly well understood (e.g. Christ and Lee et al., 2007; Alavinia and Burdorf, 2008; Shultz and Wang, 2011). An interesting feature of many partial reviews of primary studies over six decades of research, is the very unclear and contradictory characterisation of the effects, consequences or impact of retirement (e.g. Kim and Moen, 2001; Mein, Martikainen, et al., 2003; Pinquart and Schindler, 2007; Calvo and Sarkisian, 2011; Oksannen, Vahtera, et al. 2011; Coelho, Newman, Huxtable, 2014). A recent meta-analysis of studies focusing on adjustment to a variety of life-events,

including retirement, and associated changes specifically in relation to SWB, also suggested that retirement has variable influence in relation to the different components of SWB (Luhmann, et al., 2012).

Some authors have argued that, in this area of research, the complexity of this transition has been, at times, traded-in for empirical simplicity or straightforwardness of measurement (George, 1993; Moen, 1996; Calvo and Sarkisian, 2011, Horner, 2012). This, they argue, has overlooked substantive differences and variations in the experience of retirement (e.g. the voluntariness of retirement), which can be operationalised as predictors or moderators, in an attempt to explain some of this variability. Other researchers have argued that the variability in primary outcomes is linked to fundamental differences in the methodology or design of studies in this area (e.g. Ross and Drentea, 1998; Warr, Butcher, Robertson, Callinan, 2004; Pinquart and Schindler, 2007; Behncke, 2012; Coelho et al., 2014). One of these methodological issues has been the inconsistency found between studies on: a) the definition of outcome constructs related to the experience of retirement; b) the operationalisation of constructs into measures. Although there is growing interest in understanding the process and outcomes of adjustment or adaptation to retirement, these inconsistencies place obstacles related to the comparability of data, the interpretation of outcomes, and their potential to be systematically reviewed (Hofer and Piccinin, 2009). The issue can be simply put: if what is being measured, and how it is being measured, is unclear or variable, how can meaningful conclusions be drawn?

To understand how adjustment or adaptation to retirement have been conceptualised and measured in this literature is, therefore, key to understanding this evidence base, and make judgments about validity and utility of such constructs and measures. An integrative systematic review of constructs and measures of adjustment or adaptation to retirement has not been done to date. The present review proposes to contribute to this area in this way, aiming to identify and critically evaluate how adjustment or adaptation have been conceptualised and measured in retirement research studies.

2. Objectives

The review aimed to:

a) describe, organise and evaluate the constructs of adjustment and adaptation to retirement, as used in this literature.

b) identify how adjustment and adaptation to retirement have been measured in this literature, and briefly evaluate these measures' psychometric properties and conceptual clarity.

3. Methods

3.1 Criteria for considering studies for this review

3.1.1 Types of studies

This systematic review considered all empirical primary studies (i.e. not theoretical or review studies), which included a measurement/quantification (i.e. not a qualitative description, or anecdotal material) of any kind of concept/construct of adjustment or adaptation to retirement, either as a primary or secondary outcome. Studies were considered, independently of research design, on the basis of using a measurement instrument to assess adjustment or adaptation to retirement. This fitted the aim of this review, to systematically characterise and evaluate the way adjustment or adaptation to retirement is constructed, operationalized and measured in this literature. As such, studies were not excluded on the basis of low methodological or reporting quality (Tugwell, Petticrew, et al., 2010; King, Haagsma, Delfabbro, Gradisar, Griffiths, 2013).

Retirement was defined as the state or process that follows the normative exit from the work force in late adulthood (i.e. not job loss or unemployment). The presence of a construct of adjustment or adaptation to retirement was understood as an explicit reference to these phenomena as outcome variable(s). A measure of adjustment or adaptation was defined as any self-report measure (e.g. questionnaire, inventory, single-question, scale or subscale), which the authors of the study *explicitly use* to represent and quantify either: a) a state of adjustment or adaptation to retirement (e.g. at a single measurement point); or b) a process of adjustment or adaptation to retirement (e.g. at more than one measurement point). Following Schrank, et al.'s (2013) methodology, studies were included if they examined adjustment or adaptation to retirement using constructs and measures derived from psychological and mental health research. Therefore, studies that examined this solely in economic, physical health, physical/social functioning or in mortality terms were excluded. Studies were also excluded if the measurement instrument could not be identified.

There were no restrictions in relation to the period of inclusion of studies, and both published and unpublished studies were eligible. However, selective language limits were imposed, and only studies in any of the following languages were considered: English, Portuguese, Spanish, French and Italian. These limits are based on the author's available language resources, and, introduce a potential cultural bias (Lipsey and Wilson, 2001; Elwood, 2007).

3.1.2 Types of participants

Studies were considered if they included individuals: a) of both genders, b) who were retired or about to retire from the workforce (as defined above), c) who were representative of general community-dwelling populations (i.e. not of specific populations or populations with specific physical and mental health, cognitive or occupational needs or characteristics); and c) who are over the age of 40 (the lower limit of OECD's estimation of the effective age of retirement, OECD, 2011).

3.1.3 Types of outcome measures

The review included studies that provided at least one specifically stated outcome measure of adjustment or adaptation to retirement. Studies were included if adjustment or adaptation to retirement were explicitly identified as primary or secondary outcomes (so as not to exclude studies which may primarily focus on other outcomes), and if instruments (namely self-report instruments) were explicitly selected as measures of these constructs, i.e. if authors provided a rationale or statement that linked the measure to the construct.

Table 3.1 – Summary of inclusion and exclusion criteria for studies

Inclusion criteria
1. Primary studies reporting on quantitative data (no research design limits);
2. Studies which include measurement/quantification of any kind of construct of adjustment or adaptation to retirement, either as a primary or secondary outcome;
3. Studies which include an appropriate definition of retirement or retired status, as the state that follows the normative, permanent or temporary exit from the work force in late adulthood;
4. Studies which include an explicit and appropriate definition of adaptation or adjustment to retirement as an outcome variable, defined according to psychological or mental health literature and research;
5. Studies which provide at least one explicitly identified measure of adaptation or adjustment to retirement, using any of the following types of self-report instruments: questionnaire, inventory, single-question, scale or subscale;
6. Studies which include as participants individuals: a) of both genders, b) who were retired or about to retire from the workforce, c) who are representative of general community-dwelling populations; and c) who are over the age of 40;
7. Studies which are written in English, Portuguese, Spanish, French, or Italian;
8. Studies which are published or unpublished.
Exclusion Criteria
1. Studies which report on qualitative data, review data;
2. Studies which do not identify the measures that were used;
3. Studies which examine adaptation and adjustment in solely economic, physical, physical/social functioning or mortality terms;
4. Studies which include participants selected from specific populations (in relation to their retirement process and experiences): individuals with learning disabilities; individuals with pre-existing chronic or acute physical conditions (e.g. heart disease, stroke, diabetes, multiple sclerosis, cancer, chronic respiratory conditions, back pain, etc.); individuals with pre-existing chronic or acute mental illnesses (e.g. psychosis, personality disorder, etc.); individuals with pre-existing diagnosis of cognitive impairment or dementia; individuals who served in the armed forces or police services; individuals who were professional or elite sports men and women; individuals in the prison system.

3.2 Search methods for identification of studies

As most of the studies considered for this review fell into the category of observational, non-randomized studies, the challenges that these pose to identification in systematic review were taken into account in the design of this review's search strategy (Stroop, Berlin et al., 2000; Higgins and Deeks, 2011; Higgins, Ramsay, et al., 2013). Thus, the strategy prioritised sensitivity over specificity (Petticrew, Roberts, 2006; Elwood, 2007; Lefebvre, Manheimer, Glanville, 2011), and included electronic databases, other specific publication sources (i.e. governmental or institutional websites) and grey literature. All search terms were in English and included: retirement, work-cessation, bridge-employment, employment; adaptation, adjustment, wellbeing; psychological, subjective, happiness, quality, satisfaction, mental, health, illness, distress, affect, depression, retired, retirees, old, older, elder. The literature search was conducted between March and May 2013 (full search strategy and sources in Appendix 1).

3.2.1 Electronic searches

Data search included electronic databases from various disciplines (health, psychology, social sciences and business), namely: CENTRAL, MEDLINE, EMBASE, CINAHL Plus, PsycINFO, ASSIA, Social Sciences Citation Index, Global Health, ABI/INFORM Complete (see Appendix 1).

3.2.2 Searching other resources

3.2.2.1 Grey literature

Grey literature was searched, encompassing a variety of sources, namely: Conference Proceeding Citation Index, ETHOS, US Office of Public Health and Science Publications, Department of Work and Pensions (UK), International Labour Organisation, OECD iLibrary, Open Grey, Global Health Library (WHO). In addition to these sources, past or ongoing studies of ageing and/or retirement, and their host institutions, were also searched, including: English Longitudinal Study of Ageing (ELSA), Whitehall Study I and II (UK), China Health and Retirement Longitudinal Study (CHARLS), Survey of Health Ageing and Retirement in Europe (SHARE), Health and Retirement Study (US National Institute on Aging). (Appendix 1).

3.2.2.2 Hand-searching

Individual journals (digital version), which regularly publish on retirement, were consulted (since 1980) including: The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, Gerontologist, Aging and Mental Health, Social Science and Medicine, Epidemiology, Journal of Vocational Behaviour. Reference lists of relevant publications or reviews of the literature were also searched.

3.3 Data collection and analysis

3.3.1 Selection strategy and unit of analysis

Publications or records were selected by the main author, using a funnelling approach over successive sifts of data (Figure 3.1), applying the inclusion criteria systematically. The successive sifts approach was an iterative process in which the same pool of data was screened over and over again, progressing towards the final data set. Publications or records were appraised for inclusion, first, by screening of title and abstracts (Sifts 1 to 4) and, second, by reading the full-text (Sift 5). For both objectives of this review, the unit of analysis was the record of use of a measure of adjustment or adaptation to retirement, using the same selection and data management methodology as Sanderson, Tatt, Higgins (2007), Smith and Alloy (2009), Williams, Watts, Wade (2012), King et al. (2013) and Schrank et al. (2013). Records were included even if these corresponded to the same study or data-set. In this area, it is relatively frequent for authors to publish multiple reports on a study or dataset. However, because these can focus on analyses of different variables, and in order not to unwittingly exclude data, all records were included.

3.3.2 Data extraction and management

Data extraction was again done by the first author, using a purpose-built form (Appendix 2). Data extraction was guided by the objectives of the review. The first objective was to was to organise and evaluate the constructs of adjustment and adaptation to retirement described in each record, following the methodology used by Smith and Alloy (2009), Tilghman-Osborne, Cole, Felton (2010) and Schrank et al. (2013). The second objective, was to identify and understand how adaptation and adjustment to retirement has been measured in this literature. The measures used were descriptively listed, and the frequency of their use in the retrieved records presented. The specific measures were then critically evaluated in relation to their conceptual clarity and selected psychometric properties (Table 3.2). The criteria for critical evaluation followed those proposed by Terwee, Bota, et al. (2007), Sklar, Groessl, O'Connell, Davidson, Aarons (2013) and informed by Kerlinger and Lee (2000).

Figure 3.1 – Flow-chart of record identification, screening, selection and inclusion (Moher, Liberati, Tetzlaff, Altman, 2009)

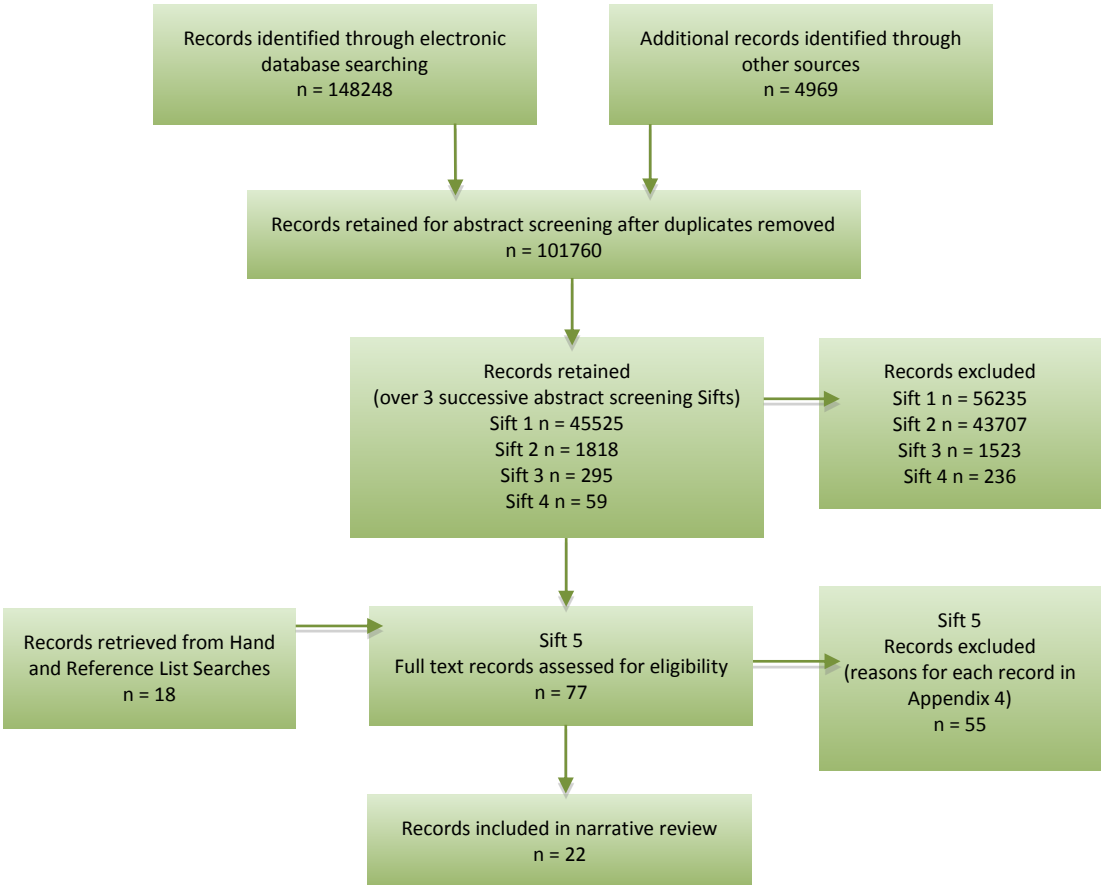


Table 3.2 – Criteria for evaluation of psychometric properties of measures

Property Type	Property Definition	Criteria for Evaluation	Notation
Conceptual Clarity	The extent to which the construct being measured is adequately described and evidence or theory-based	A clear narrative description of the construct being measured is provided with theoretical, model or evidence support.	
Content Validity	The extent to which the domain of interest is comprehensively sampled by the items in the instrument.	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and item selection.	
Internal Consistency	The extent to which individual items of the instrument are consistent with each other.	Scales measuring a unitary construct expected to have high internal consistency. Multifactorial scales expected to have lower internal consistency. Either α reliability or split half reliability values are > .70.	
Convergent Validity	The extent to which the instrument relates to existing instruments of the same construct.	Convincing arguments for selection of 'gold standard' measure of the same construct, and when compared to this, correlation of at least 0.7 is obtained	+ Present - Absent ? Cannot Tell NA Not Applicable
Construct Validity	The extent to which the instrument fits the theoretical construct that it relates to.	Factor analysis is used, and results fit the proposed structure of the instrument. Inter-correlations between items of the same factor are > 0.7. Scores on the instrument correlate with scores on other measures, in a way that is theoretically consistent or predicted.	
Test-Retest Reliability	The extent to which an instrument produces consistent or reproducible outcomes	Test-retest reliability studies produces a weighted Kappa of ≥ .70	
Interpretability	The extent to which qualitative meaning can be assigned to quantitative scores	Mean and Standard Deviation scores are presented for at least two relevant subgroups.	

3.3.3 Assessment of quality or risk of bias in included studies

The aim of this review was to systematically characterise, organise and evaluate the way adaptation or adjustment to retirement is operationalized and measured in this literature. Given this specific aim, reviews of this kind do not require formal assessment of methodological quality of included studies or risk of bias (e.g. Sanderson et al., 2007; King et al., 2013).

3.3.4 Dealing with missing data

For all publications or records, at the coding stage, to improve completeness of the data set, additional information was sought from other sources (namely, publications relating to properties of the measurement instruments, correspondence with authors [see Appendix 3]), where this was feasible and in the context of available resources.

4. Results

4.1 Description of studies

4.1.1 Results of the search – included and excluded records

The search strategy for this review (in Appendix 1) yielded a very large number of records, at a first stage (Figure 3.1 outlines the search process and outcomes). Because of this, a sequential approach of selection 'sifts' was used, and, at a final stage, 77 records were identified for full-text reading. At this point, 55 records were excluded and 22 records retained.

4.1.1.1 Excluded records

Information on excluded records, with respective reasons for exclusion, is provided in Appendix 4. These studies did not meet inclusion criteria for two main reasons: a) the absence of a link, reference or definition of any construct of adjustment or adaptation to retirement, b) deficient identification in the study of measures of these constructs.

4.1.1.1 Included records

The unit of analysis in this review was the record of use of a measure of adjustment or adaptation to retirement (e.g. as used by Schrank, 2013). As such, the 22 records that were retained correspond to 19 individual studies, which report on the use of 27 identifiable measures of adjustment or adaptation to retirement[4]. Specifically, 3 pairs of records (Potocnik, et al. [2010] and Potocnik, et al. [2013]; Reitzes and Mutran [2004] and Reitzes and Mutran [2006]; van Solinge and Henkens [2005] and van Solinge and Henkens [2008]) correspond to 3 individual studies only. Systematic review methodologies developed for the assessment of the effect of interventions or exposures (e.g. Lipsey and Wilson, 2001; Higgins and Deeks, 2011) define, accordingly, the unit of interest to be the study, not the report. However, in the case of this review, given its objectives, this logic was reversed. The interest was on the report of use of a measure of adjustment or adaptation to retirement. The existence of more than one record for an individual study (i.e. more than one journal publication) is very common. It is also common for authors to report on different aspects of their data analysis, or on different sets of variables, in different reports. The decision was made to retain all records of a study in the data set. This was done to prevent the exclusion of constructs or measures which could be of interest.

[4] The use of physical health measures was recorded in the analysis of these records, but not explored further, as the focus of this review was on psychological constructs of adjustment or adaptation to retirement. As outlined below (e.g. Table 4.3) measures or indexes of physical health were recorded as part of composite measures of adjustment or adaptation to retirement.

The retained 22 records were put forward for review. Information on the characteristics of the studies that the records correspond to is provided in Appendix 5 (Table of characteristics of included studies). These studies were either cross-sectional or longitudinal/cohort in design, and varied in relation to sample size.

4.2 Main findings

4.2.1 Constructs of adjustment or adaptation to retirement

The analysis of constructs of adjustment or adaptation to retirement, as they are used in these records, resulted in a possible framework of six clusters of constructs (outlined in Table 4.1). Overall, these define adjustment or adaptation to retirement in terms of: 1) retirement-specific satisfaction; 2) role adjustment; 3) general wellbeing, life-satisfaction and related terms; 4) stress and coping with crisis; 5) mental health or distress; and 6) compounds of dimensions (with variable sets of components). The links between these construct clusters and theoretical frameworks was also variable, but, seemed to include most of the main theories/models outlined in the introduction to this review. It is interesting to note that most studies that took a compound of dimensions approach used Atchley's model of adjustment or adaptation as their theoretical frame.

4.2.1.1 Variability of constructs

It is clear from Table 4.1 that there is considerable variability in how adjustment or adaptation to retirement is defined in this set of studies (which was predictable). However, there was also enough consistency to allow the outlining of construct clusters. A considerable number of studies used the idea of adjustment or adaptation to retirement as a function of retirement satisfaction, or satisfaction with life in retirement. This, as outlined in Table 4.3, is linked to the use of specifically developed measures (e.g. Floyd, et al., 1992; van Solinge and Henkens, 2008), or the adaptation of life-satisfaction measures into the retirement context (e.g. Herve, et al., 2012). Also, this construct cluster seems to be based, both, on theories of retirement and wider literature on adjustment or adaptation, life-span development and wellbeing.

Although, the construction of adjustment or adaptation to retirement just in terms of mental health or distress was used only by Wang (2007), all the studies that used a compound of dimensions approach included mental health (or 'psychological well-being') as a component. In addition to this, these typically included in their definition a reference to life-satisfaction and other aspects of individual functioning and resources (namely, health status, social and financial resources).

Finally, only a few of these studies used theories to generate constructs and hypotheses. Van Solinge and Henkens (2008) specifically used theory to guide the elaboration and testing of differential

constructs of adjustment and satisfaction (explored further below). In turn, Gall, et al. (1997), Reitzes and Mutran (2004) and Wang (2007) directly tested the extent to which Atchley's model of adjustment to retirement predicted patterns of adjustment in their samples. Although, they initially used different constructs and measures of adjustment, their outcomes (which generally find support for the model) are presented later in the results section of their studies mainly in terms of psychological wellbeing. It is somewhat difficult to differentiate if this is a conceptual choice (i.e. this is how the authors chose to ultimately define adjustment or adaptation to retirement), or a choice based on selective reporting of outcomes.

Table 4.1 – Constructs of adjustment or adaptation and associated measures

Construct Cluster	Construct	Adjustment or adaptation to retirement is viewed as…	Typical corresponding theoretical frame/model in studies	Studies where present
Retirement-specific satisfaction	Satisfaction with life in retirement, or retirement satisfaction	1. A function of satisfaction/contentment with life in retirement or after retirement. 2. Positive retirement experiences. 3 Ability to develop a new satisfying life-style and role as a retired person. 4. A parallel/identical process to development of satisfaction/dissatisfaction with work.	. Continuity theory . Atchley's model of retirement adjustment . Role theory . Resource theory . Life-span developmental theories . General well-being and life-satisfaction models . Successful aging models . Job satisfaction model	Beck (1982) Donaldson et al. (2010) Dulin et al. (2011) Floyd et al. (1992) Fonseca (2007) Potocnik et al. (2010) Potocnik et al. (2013) Reitzes and Mutran (2004) Reitzes and Mutran (2006) van Solinge and Henkens (2008) van Solinge and Henkens (2005) Wong and Earl (2009)
Role adjustment	Ability to change role successfully after retirement	1. A function of detachment from the work role and coping with loss. 2. A function of acceptance of alternative roles to that of worker.	. Role theory . Disengagement theory . Activity theory	Crawford (1972) George and Maddox (1977) van Solinge and Henkens (2008) van Solinge and Henkens (2005)
General well-being, life satisfaction & related terms	Subjective well-being, wellbeing, or life- satisfaction	A function of changes in SWB or life-satisfaction, used as equivalent terms.	. Life-span developmental theories	Herve et al. (2012)
Stress and coping	Coping, distress management after the crisis of retirement	A function of the ability to cope with and adapt to stressful life events.	. Crisis orientation models of adjustment and adaptation, emphasising coping with stress	George and Maddox (1977) Palmore et al. (1979)
Mental health or distress	Mental health or psychological distress	The absence of symptoms of mental ill health or psychological distress.	. Continuity theory . Role theory . Resource theory . Life-span developmental theories	Wang (2007)
Multi-dimensional	Compound of dimensions	A positive outcome in a set of life domains, including: physical health, mental health, activities, social context, finances, life-satisfaction, wellbeing.	. Atchley's model of adjustment to retirement . Resource theory	Braithwaite et al. (1986) Gall et al.(1997) Isaksson and Johansson (2000) Mattila et al. (1989) Nuttman-Shwartz (2004)

4.2.1.2 Clarity of constructs

The level of clarity with which the constructs were described in these reports was also variable (Table 4.2). A number of studies provided clear, theoretically rich and consistent definitions of adjustment or adaptation to retirement (namely, Donaldson, et al., 2010; Floyd, et al., 1992; George and Maddox, 1977; van Solinge and Henkens, 2008; Wang, 2007). However, in those reports that provided less clear definitions, there were four identifiable ways in which this was found lacking. The first, reports did not elaborate beyond the labelling of the construct as adjustment or adaptation. Although the label may have been used consistently throughout the paper (e.g. Braithwaite, et al., 1986; Crawford, 1972; Dulin, et al., 2011; van Solinge and Henkens, 2005; Wong and Earl, 2009), there was uncertainty about how the measures related to the underlying phenomenon being quantified. Second, in some cases, what authors defined as adjustment or adaptation to retirement could only be inferred 'backwards', through the measures that were used (e.g. Mattila, et al., 1989; Isaksson and Johansson, 2000). Although this can be seen more of an issue of reporting quality, again, it leads to a similar kind of uncertainty. The third, was the lack of a definition altogether (Nuttman-Shwartz, 2004), which again could be a problem of reporting quality. The fourth, was the diversity of constructs offered within one report (e.g. Beck, 1982; Gall, et al., 1997; Isaksson and Johansson, 2000), and the lack of a clear link between the construct used and the measures then selected to quantify it. It is interesting to note that this issue, as well as the issue of 'backward' inference described above, was typical of studies that took a compound of dimensions approach to adjustment or adaptation to retirement.

Table 4.2 – Clarity of constructs of adjustment or adaptation to retirement

Study ID and Reference	Clear	Unclear or variable
1. Beck (1982)		●
2. Braithwaite et al. (1986)		●
3. Crawford (1972)		●
4. Donaldson et al. (2010)	●	
5. Dulin et al. (2011)		●
6. Floyd et al. (1992)	●	
7. Fonseca (2007)	●	
8. Gall et al. (1997)		●
9. George and Maddox (1977)	●	
10. Herve et al. (2012)	●	
11. Isaksson and Johansson (2000)		●
12. Mattila et al. (1989)		●
13. Nuttman-Shwartz (2004)		●
14. Palmore et al. (1979)	●	
15. Potocnik et al. (2010)	●	
16. Potocnik et al. (2013)	●	
17. Reitzes and Mutran (2004)	●	
18. Reitzes and Mutran (2006)	●	
19. van Solinge and Henkens (2008)	●	
20. van Solinge and Henkens (2005)		●
21. Wang (2007)	●	
22. Wong and Earl (2009)		●
Total	**12**	**10**

Note: This judgment is based on the presence, anywhere in the report, of a construct, or at least an operationalisation of adjustment or adaptation to retirement; this is considered clear if it is described in terms that are intelligible and consistent.

4.2.2 Measures of adjustment or adaptation to retirement

The analysis of the measures employed to quantify adjustment or adaptation to retirement focused, both, on the frequency with which each measure is used, their link to construct clusters above (Table 4.3) and a brief evaluation of their psychometric properties (Table included in Appendix 6). The following narrative part of the review will focus specifically on measures corresponding to the retirement-specific satisfaction or role adjustment cluster, and the work satisfaction/dissatisfaction cluster. The decision to narrow the focus of the narrative exploration at this point is related to the fact that the psychometric properties of the measures included in the mental health and well-being and life-satisfaction clusters have been very widely and competently researched elsewhere in psychological and medical literature. These measures are, nonetheless, an important part of this data-set. These were, therefore, included in the selection phase and systematically described and evaluated according to the pre-specified psychometric criteria in Appendix 6. However, to narratively explore these measures at length here would be outwith the space limits of this article, and repetitive in relation to recent review efforts (e.g. McDowell, 2010; Carleton, Thibodeau, et al., 2013).

4.2.2.1 Constructs, measures and frequency

Variability is, again, present in the choice of measures across studies. This is likely to be a function of the dispersion in time in this set of records (from 1977 to 2013). Studies will, of course, make use of what measures are available to them at the time. However, as shown in Table 4.3, most measures are not used in more than one or two studies. This raises issues of commensurability in review efforts focusing, for instance, on the effect of retirement, as an exposure. This diversity is also present in relation to measures of mental health, which may be determined by choices that are exogenous to the specific analysis or report in the selected publication. This research area is one in which large bodies of survey data are used by authors for different analytical aims. At times, publications present a secondary analysis of a data set, which was not specifically created for the objectives of the secondary analysis. Wang (2007) is a good example of this. This record used data from the American Health and Retirement Study, and, because of this, a measure of mental health (CES-D 8) was adapted into a single measure of adjustment or adaptation to retirement. The author does note that this was not an ideal choice (in terms of construct specificity and validity), but a functional choice.

Another feature of the choice of measures in this data set was the use of non-specific measures to retirement. Potocnik, et al. (2010 and 2013) and Herve, et al. (2012) seem to have co-opted, shortened or adapted measures that had been developed for other constructs, into the realm of the retirement experience. Namely, Potocnik et al. (2010 and 2013) used a measure of work satisfaction/dissatisfaction (Bussing, Bissels, Fuchs, Perrar, 1999), though highly adapted for their purposes. This was theoretically framed and appropriately operationalised in Potocnik, et al. (2013), but not in Potocnik, et al. (2010). However, the level of changes that Potocnik et al. (2013) introduced in the adaptation of the measure for their purposes, makes the consideration of the initial properties of

this measure irrelevant in this context (as it is so different from the original form). Herve, et al. (2012), on the other hand, included a measure of boredom-proneness, as part of a composite of measures designed to assess satisfaction or wellbeing in retirement. Although this does not seem, at face-value, incongruous to the notion of psychological adjustment or adaptation, the issue of conceptual justification remains unaddressed in this study.

Finally, like in the case of Herve, et al. (2012), most measures in this set of records are used as part of composites of measures, that is, more than one indicator is defined to quantify adjustment or adaptation to retirement. This seems to fit a more general sense that these phenomena of adjustment or adaptation to retirement include changes in more than one domain of human experience.

Table 4.3 – Frequency of use of each measure of adjustment or adaptation to retirement

Construct Cluster	Measure	Studies that use the measure	as single measure	as part of composite measure	Total n
Retirement-specific satisfaction or Role adjustment	*Ad-hoc* single-item/question Typically directly asking about satisfaction with retirement, adjustment with retirement, well-being or life-satisfaction in retirement or related terms.	Beck (1982) (2 item) Braithwaite et al. (1986) (unclear number) Crawford (1972) (unclear number) Gall et al. (1997) (1 item) Herve et al. (2012) (2 items) Isaksson and Johansson (2000) (1 item) Palmore et al. (1979) (1 item)	• • 	 • • • • •	7
	Retirement Satisfaction Inventory Main Ref. Floyd et al. (1992)	Floyd et al. (1992) Fonseca (2007)	• •		2
	Retirement Adjustment Questionnaire Main ref.: Wells et al. (2006)	Donaldson et al. (2010) Wong and Earle (2009)	• •		2
	Expected adjustment to retirement subscale (adapted) Main ref. Taylor and Shore (1995)	Dulin et al. (2011)	•		1
	Retirement Descriptive Index Main ref. Smith, Kendall, Hulin (1969)	Gall et al. (1997)		•	1
	Retirement Adjustment and Satisfaction Scales Main Ref. van Solinge and Henkens (2007)	van Solinge and Henkens (2008) van Solinge and Henkens (2005)	• •		2
	Positive Attitudes Towards Retirement Main Ref. Atchley and Robinson (1982)	Reitzes and Mutran (2006) Reitzes and Mutran (2004)	• •		2
Work satisfaction/dissatisfaction	Job-Satisfaction Questionnaire Short Form (adapted) Main Ref. Bruggemann (1976) (German), also in Bussing, Bissels, Fuchs, Perrar (1999)	Potocnik (2010) Potocnik (2013)	• •		2
	Boredom Proneness Scale (French adaptation) Main Ref. Gana and Akremi (1998) also in Farmer and Sundberg (1986)	Herve et al. (2012)		•	1
Mental health or distress	Centre for Epidemiological Study of Depression Scale – 8 (CES-D 8) Main Ref. Radloff (1977)	Wang (2007)	•		1
	General Health Questionnaire – 12 (GHQ12) Main Ref. Goldberg, Garter, et al. (1997)	Isaksson and Johansson (2000) Potocnik et al. (2010)	 	• •	2
	General Health Questionnaire – 36 (GHQ36) Main Ref. Goldberg (1978)	Mattila et al. (1989)		•	1
	Symptom Checklist 90 (SCL-90) Main Ref. Derogatis, Lipman, Covi (1973)	Gall et al. (1997)		•	1
	Mental Health Inventory (MHI) Main Ref. Veit and Ware (1983)	Nuttman-Shwartz (2004)		•	1

General well-being, life-satisfaction or related terms	Kutner Morale Scale Main Ref.: Kutner, Fanshel, Togo and Langner (1956)	George and Maddox (1977)	•		1
	Lawton's Philadelphia Geriatric Centre Morale Scale Main Ref. Lawton (1975)	Mattila et al. (1989)		•	1
	Affect Balance Scale Main Ref. Bradburn (1969)	Palmore et al. (1979)		•	1
	Satisfaction With Life Scale Main Ref. Diener, Emmons, Larsen , Griffin (1985)	Herve et al. (2012)		•	1
	The Ryff's Scales of Psychological Well-Being Main Ref: Ryff and Keyes (1995)	Herve et al. (2012)		•	1
	Life-Satisfaction Index A Main Ref. Neugarten, Havinghurst, Tobin (1961)	Mattila et al. (1989)		•	1
Physical health recorded as part of Multidimensional construct	Physical health measures/indexes	Braithwaite et al. (1986) Gall et al. (1997) Herve et al. (2012) Isaksson and Johansson (2000) Mattila et al. (1989) Nuttman-Shwartz (2004) Palmore et al. (1979)		• • • • • • 	6
		Total	**9**	**12**	

Notes:
All single-item/questions were collapsed into one 'Single-item/questions' category, which includes all single-item/questions used, independent of content;
All physical health measures/indexes were collapsed into one 'Physical health measures/ indexes' category, which includes all health measures/indexes used, independent of content;
Judgment of measure used as 'part of composite' only in the specific measurement of the construct of adjustment or adaptation to retirement, as it is defined by each study in the data set.

4.2.2.2 Properties of included measures

The following narrative part of this review, as stated above, will focus mainly on the specific measures that formed the clusters of *Retirement-specific satisfaction* or *Role adjustment*, and *Work satisfaction/dissatisfaction* measures. However, it will also briefly focus on single-item/questions, as these form an important part of the measurement selections in this data set.

4.2.2.2.1 Single-item/questions

Looking at this category of measures, these were included in 7 of the records in this review, again, as part of measurement compounds. These tend to ask participants to rate or scale their level of satisfaction, happiness or wellbeing in relation to retirement, retirement process or life after retirement. Where a verbatim reproduction of the questions is presented in the record, these tend to be variations around: '*taking things altogether, would you say you're very happy, somewhat happy, somewhat unhappy, or very unhappy these days*?' (Beck, 1982:616); '*was adjustment [to retirement] difficult, somewhat difficult or not at all difficult*' (Braithwaite et al., 1986:494); '*in general, how satisfying do you find the way you're spending your life today*?' (Gall et al., 1997:112). These questions are worded in a relatively similar way to single-item measures of general well-being or satisfaction (McDowell, 2010). Questions of this type are commonly used in surveys, as they have the advantage of a quick assessment, and, Kim-Prieto, et al. (2005) and McDowell (2010) consider that measurement properties of single-item/questions on general satisfaction or wellbeing are of better quality than what would be expected. McDowell (2010) reports that these measures are relatively stable in the short-term (but not in the long term), and show high convergent validity in relation to multi-item measures (Diener, 1984; Diener, Emmons, Larsen, Griffin, 1985; McDowell, 2010). However, Diener (1984) points out some of the difficulties that reliance on a single item poses for measurement, namely: confounding introduced by the specific wording/interpretation of the question, celling effects, and lack of specificity.

4.2.2.2.2 Retirement-specific satisfaction and role adjustment cluster

A characteristic of the *Retirement-specific satisfaction* or *Role adjustment* cluster of measures was the relative lack of detailed psychometric information in relation to some of these instruments. As it is possible to see in Appendix 6, information on the Wells, deVaus, Kendig, Quine, Petralia (2006) Retirement Adjustment Questionnaire used in Wong and Earl (2009) and Donaldson et al., (2010) is, at this point, inaccessible (this is not published formally, referenced in these studies to a document published online only). The properties included in Appendix 6 are surmised from brief information included in Donaldson et al. (2010) and Wong and Earl (2009) (namely, internal consistency estimates $\alpha = .81$). Although the main author of the original questionnaire was contacted, the original document could not be obtained. The same issue arose for the Retirement Descriptive Index (Smith, Kendall, Hulin, 1969), the Positive Attitudes Towards Retirement tool (Atchley and Robinson, 1982) and the Expected Adjustment to Retirement subscale (Taylor and Shore, 1995). These are instruments that are not widely available or not widely used in the literature. Although, multiple references for these

instruments were sought, the psychometric information that was available was, nonetheless, very sparse. Therefore, these do not allow further meaningful comment on their measurement properties.

The Retirement Satisfaction Inventory (Floyd, et al., 1992), also used by Fonseca (2007) was specifically developed as a measure of satisfaction with life in retirement. This is is used in both studies as a single measure of retirement adjustment. It stands out in this dataset for its conceptual clarity and measurement transparency. Floyd, et al.'s (1992) publication presents the two studies involved in the development of the measure and reports on psychometric data (see Appendix 6). However, this inventory's properties (namely, construct validity, convergent validity and test-retest reliability) did not fully meet the quality thresholds defined by this review's evaluation criteria. Equally, the Retirement Adjustment and Satisfaction Scales developed by van Solinge and Henkens (2008) (and used in prototypical form by the same authors in 2005), again, represent an effort of conceptually-driven measurement that stands out in this data set. These scales were based on the definition of two differentiated constructs of adjustment and satisfaction, based on the developmental tasks the authors considered key to each process, respectively: coping with loss of the worker role (and associated social losses); and the development of alternative and satisfactory roles and life. This seems to be an inherently psychological approach to these phenomena, which is then tested empirically in a relatively large sample (n = 778). Although the internal consistency of the two scales was acceptable (α = .80 and .65, respectively), the papers did not present further psychometric information. Therefore, the quality of these scales in assessing the presumed two different constructs is not established.

4.2.2.2.3 Work satisfaction/dissatisfaction cluster

As mentioned above, the Job-Satisfaction Questionnaire Short Form (Bussing, et al., 1999) was excluded from further evaluation because of its very partial and adapted use in Potocnik et al. (2010 and 2013). However, the remaining measure in this cluster, the Boredom Proneness Scale (Farmer and Sundberg, 1986) in its French adaptation (Gana and Akremi, 1998), was assessed. Herve et al. (2010) used this adapted measure as part of a composite, in a way that is conceptually very unclear and unsupported by the design purpose and previous use of this measure (Farmer and Sundberg, 1986). The psychometric information in both publications is not completely consistent. Namely, while Gana and Akremi (1998) found acceptable levels of construct validity in their French validation studies, the same was not obtained by Farmer and Sundberg (1986).

5. Discussion

5.1 Summary of main results

This review aimed to characterise how adjustment or adaptation to retirement is constructed and measured in this literature. It included 22 records of use of measures of adjustment or adaptation to retirement (corresponding to 19 individual studies), which report on 27 identifiable measures of these constructs. The results offered a detailed insight into what is being measured and how.

As described in the introduction, this is an area of research that has produced very variable and inconsistent evidence in relation to the impact of retirement on individuals' lives. Part of this variability has been attributed to methodological diversity across studies (e.g. Pinquart and Schindler, 2007; Topa et al., 2009), specifically, inconsistencies in definition of outcome constructs and selection of measures (e.g. Ross and Drentea, 1998). This review sought to explore and understand how consistent or comparable constructs and measures of adjustment or adaptation to retirement are in this literature. Without this understanding, it is argued, the ability to design and systematically review outcome research in this area (i.e. interpret outcome data) will be limited.

5.1.1 Variability of constructs of adjustment and adaptation to retirement

As described in the results, these constructs were defined and operationalised in very variable ways across studies. Overall, this variation was organised into 6 identifiable construct clusters (Table 4.1): retirement-specific satisfaction; role adjustment; general wellbeing, life-satisfaction and related terms; stress and coping with crisis; mental health/distress; and compounds of dimensions (with variable sets of components). Additionally, variability was also found within studies, namely, in the 5 studies in this review that used compounds of dimensions (and measures) of adjustment or adaptation to retirement (Table 4.1). Rather than offering a consistent integration of outcomes of the different components of the construct that were initially proposed, these studies seemed to use these compounds as sets of measures from which to later chose an 'effect' (in the report of outcomes).

This variability between and within studies is not a surprising observation. A preliminary overview of the literature had already suggested this (e.g. Wang, 2007; van Solinge and Henkens, 2008; Wang, et al. 2011). However, this review sought to go beyond this and provide an organising framework for these disparate (but related) constructs. The aim was, therefore, to make a first and tentative effort towards clarification and systematisation of the specific meanings of outcome research in this area. This is not the same as reducing complexity in approaching this phenomenon. This review argues that the systematisation offered allows a more transparent or deciphered view of retirement outcome

research. Specifically, it may allow a possible way of understanding part of the reasons why outcomes of research into the impact or effect of retirement diverge so widely (Kim-Prieto, et al., 2005).

5.1.2 Variability of measures

The measures found in this review included those designed for the assessment of mental health, wellbeing, stress and coping, work satisfaction, boredom-proneness or retirement satisfaction (Table 4.3). The specific measurement properties for each one were also explored (Appendix 6). This variability can be assumed to be related to authors' choices, and specific research traditions/approaches. In the review dataset, measurement choices were well supported theoretically or empirically in some studies (e.g. van Solinge and Henkens, 2008 is a good example of this), but, in others this is not addressed. This leads to relative difficulty in understanding what exactly is being measured (i.e. what does the self-reported change in retirement satisfaction score actually mean?). The variability of outcome measures, however, can also be related to factors that are exogenous to the specific study or analysis being carried out. Some of the reports included in this review represented secondary analyses of data collected for wider purposes, e.g. Wang (2007). This author acknowledged that that his choice of CES-D (measure of symptoms of depression) as a way to operationalise adjustment or adaptation to retirement was not ideal, and dictated by the data that he was using. This aspect is relevant to this review in that it suggests that conceptual and measurement refinements and validity in the area of research on retirement may suffer because of the choice to forgo these concerns over the attractive availability of large bodies of data.

5.2 Quality, completeness and applicability of evidence

5.2.1 Quality

The comments in this section are related to reporting quality only. This was assessed only during the face-value characterisation of studies (Appendix 5), as no formal assessment was conducted. Reporting quality was found to be variable across studies. However, the studies in this review spanned four decades, during which standards for reporting quality in research also evolved – namely, the development of guidelines for the reporting of observational studies (e.g. Tooth, Ware, Bain, Purdie, Dobson, 2005; von Elm, Altman et al., 2007). Taking this into account, nonetheless, the variable quality and completeness of information within reports was not always a function of the age of the study or publication – Dulin, et al. (2011) is a good example of very incomplete reporting, whereas Floyd, et al. (1982) are a good example of complete reporting. The issue of missing information posed difficulties for this review.

5.2.2 Completeness and applicability

The dataset that was reviewed here allowed a characterization of, both, constructs and measures used in the retirement literature to define and quantify adjustment or adaptation to retirement. In that strict sense, it allowed the fulfilment of both review objectives. However, as outlined above, this is an area that intersects a large amount of publications and records in diverse research areas. Therefore, any statement related to the completeness of the data would be, without further quantitative analysis, speculative. Given this characteristic of the field, a) the very inclusive identification of studies, and b) stringent selection criteria and process, were used to attempt to enhance both the sensitivity of the search and the specificity of the selection. In this sense, this review drew on a broad literature base, which provides some level of external validity to the findings. It is also an independent effort (to the extent that this is possible), with no specific allegiance to theoretical or research traditions. Equally, the use of 'gold standard' procedural guidelines (Higgins and Green, 2011; Reeves et al., 2011) offer some support to the reliability of the process and outcomes. However, as there are no comparable reviews in this area of research, it is difficult to make any further statements related to completeness of the evidence.

Equally careful considerations are made in relation to the applicability of the evidence. The verification and description of the variability of constructs and measures (across and within studies), the organisation of six conceptual clusters and the evaluation of the measurement properties of the measures used so far, can potentially help researchers in this field to thoughtfully chose measures that fit more closely the constructs of adjustment or adaptation to retirement that they endorse. Moreover, it may emphasise how relevant it is to clearly define and justify how adjustment and adaptation to retirement is understood in their research, so that outcomes can be adequately and usefully interpreted. For instance, in cohort studies of ageing and retirement, it would be useful and more theoretically consistent (e.g. with continuity models, or life-long development perspective) to link measures to specific tasks of adjustment to retirement (for instance, using van Solinge and Henkens's [2008] ideas of phases of adjustment to loss of worker role/life, and later finding satisfaction with retired role/life). Compared to the use of a general mental health functioning scale, this kind of characterisation would provide a much more detailed, informative and useful view of what a) demands are put on individuals at what time during the retirement transition, and b) how difficulties in adjustment or adaptation to retirement can be more specifically formulated and supported.

5.3 Potential biases in the review process

The risk of bias is inherent to any review, more so, in narrative reviews (Higgins, et al., 2011). As addressed above, there were several issues that suggest limitations to the robustness of the inferences drawn from this review.

5.3.1 Characteristics of the review process

There were several points of decision-making in the review process. Three of these decisions are focused here. First, the decision to focus strictly on explicitly used constructs and measures of adjustment or adaptation to retirement in the literature, introduced bias from the start. This was done to provide some homogeneity to the dataset, so that studies could be comparable, i.e. an attempt to manage the 'apples and oranges' problem (Sharpe, 1997; Lipsey and Wilson, 2001). Equally, this was done to enhance the possibility of identifying some conceptual framework (Schrank, et al., 2013). It thus required records to use and label these constructs explicitly, potentially excluding related constructs and measures that could inform the review objectives further. Second, the review included literature that was published in a limited set of languages. Although this was more inclusive than just English, it excluded an identified (during data searches) rich body of evidence written in other languages (namely, Scandinavian languages, Dutch, German and Chinese languages). This will limit the inferences that were drawn to specific geographical, social and research settings (Lipsey and Wilson, 2001). Finally, the decision to include all identifiable measures of the construct of adjustment or adaptation to retirement, although designed to enhance the completeness of this data set, brought about difficulties in the characterisation and evaluation of the properties of these measures. In some cases, because these measures related to vast bodies of literature (namely, established and widely used/evaluated mental health measures, e.g. CES-D or GHQ), in other cases, because these were rarely used or referred to in the data (namely, very narrow measures of retirement satisfaction like the Positive Attitudes Towards Retirement Tool [Atchley and Robinson (1992)]).

5.3.2 Characteristics of the literature

The large amount of formal and informal publications in this area was noticeable. There also seemed to be an identifiable trend for secondary analyses of data, namely, data resulting from large surveys (in the USA, UK and Australia, for instance). This introduces a somewhat misleading perception of diversity and independence of evidence and design characteristics in this field of research[5]. Also, the spread of relevant methodological information across several publications/records introduced further difficulties, as the authors refer to design characteristics already described in previous papers or less accessible study protocols. Although this is a common and understandable practice (related to publication word limits, etc.), it leads to difficulty in understanding what exactly is being measured.

In a related way, the variability of concepts and variables in this body of literature was also problematic, in requiring an inclusive set of initial search terms (and potential initial dispersion of the review focus). The construct of adjustment or adaptation is perhaps particularly vulnerable to multiple interpretations and operationalisations. As addressed above, the study of retirement is obviously a

[5] For a review of implications for the study of the effects of retirement as an exposure, see Coelho et al. (2014).

multidisciplinary area. The relevant body of literature for this review was, therefore, very spread-out, introducing challenges related to the identification of studies and data.

6. Conclusions

This systematic review included 22 records of use of measures of adjustment or adaptation to retirement (corresponding to 19 individual studies). Overall, it described, organised and evaluated 27 measures of adjustment or adaptation to retirement. There are various limitations to the external and internal validity of this review, however, it can be argued that it stands as a first comprehensive effort in evaluating and linking-up both constructs and measures of adjustment and adaptation to retirement. This review argued that a) the detailed characterization of the variability of constructs and measures (found across and within studies), b) the organisation of 6 conceptual clusters, and c) the evaluation of the measurement properties of the measures used so far, can: 1) potentially help researchers in this field to thoughtfully chose measures that fit more closely the constructs of adjustment or adaptation to retirement that they endorse; 2) highlights the critical need for researchers in this area to clearly define and justify how adjustment and adaptation to retirement is understood in their research, so that their outcomes can be adequately interpreted. Ultimately, however, this review ends with a question: given the heterogeneity found in how adjustment or adaptation to retirement was defined and operationalised across and within studies, how useful is this as a research construct in this area (and for our understanding) of retirement as a life-transition?

Highlights (see Appendix A):

- Outcomes of research on adjustment and adaptation to retirement vary greatly.
- This variability is potentially, in part, underpinned by the variable constructs and measures used.
- The 27 measures evaluated do not seem, at face-value, to measure the same construct.
- The ability to understand and compare outcomes between studies in this area is hindered by this.

References

Key:
Papers included in the Systematic Review are signalled with **
Papers *related to measures evaluated* in the Systematic Review are signalled with #

Alavinia, S.M., Burdorf, A. (2008). Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. *International Archives of Occupational and Environmental Health*, vol. 82, p. 39-45

Angelini, V., Brugiavini, A., Webber, G. (2009). Aging and unused capacity in Europe: is there an early retirement trap? *Economic Policy*, vol. 24(59), p. 463-504.

Atchley, R.C. (1971). Retirement and leisure participation: continuity or crisis? *The Gerontologist*, vol. 11(1), p. 13-17.

Atchley, R.C. (1979). Issues in retirement research. *The Gerontologist*, vol. 19(1), p. 44-54.

Atchley, R.C. (1982a). Retirement: leaving the world of work. *Annals of the American Academy of Political and Social Science*, vol. 464, p. 120-131.

Atchley, R.C. (1982b). Retirement as a social institution. *Annual Review of Sociology*, vol. 8, p. 263-287.

Atchley, R.C. (1989). A Continuity theory of normal aging. *The Gerontologist*, vol. 29(2), p. 183-190.

Baltes, P. (1987). Theoretical propositions of life-span developmental psychology: on the dynamics between growth and decline. *Developmental Psychology*, vol. 23(5), p. 611-626.

Baltes P.B., Staudinger U.M., Lindenberger, U. (1999). Lifespan psychology: theory and application to intellectual functioning. *Annual Review of Psychology*, vol. 50, p. 471-507.

** Beck, S.H. (1982). Adjustment to and satisfaction with retirement. *Journal of Gerontology*, vol. 37(5), p. 616-624.

Behncke, S. (2012). Does retirement trigger ill health? *Health Economics*, vol. 21, p. 282-300.

Bem, S., de Jong, H.L. (2002). *Theoretical issues in psychology: an introduction*. London: Sage Publications.

Biswas-Diener, R., Diener, E. (2001). Making the best of a bad situation: satisfaction in the slums of Calcutta', *Social Indicators Research*, vol. 55, p.329-352.

** Braithwaite, V.A., Gibson, D.M., Bosly-Craft, R. (1986). An exploratory study of poor adjustment styles among retirees. *Social Science and Medicine*, vol. 23(5), p. 493-499.

Brugiavini, A., Croda, E., Mariuzzo, F. (2005). Labour force participation of the elderly: unused capacity? In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *Health, Ageing and Retirement in Europe: First Results from the Survey of Health, Ageing and Retirement in Europe,* p. 236-240. Mannheim: Mannheim Research

Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/SHARE_FirstResultsBookWave1.pdf

Busseri, M. A., Sadava, S.W. (2011). A review of the tripartite structure of subjective well-being: implications for conceptualization, operationalization, analysis, and synthesis. *Personality and Social Psychology Review*, vol. 15(3), p. 290-314.

Bussing, A., Bissels,T., Fuchs, V., Perrari, K.-M. (1999). A dynamic model of work satisfaction: qualitative approaches. *Human Relations*, vol. 52, p. 999-1028.

Butterworth, P., Gill, S.C., Rodgers, B., Anstey, K. J, Villamil, E., Melzer, D. (2006). Retirement and mental health: analysis of the Australian National Survey of Mental Health and Wellbeing. *Social Science and Medicine*, vol. 62, p. 1179-1191.

Cahill, K. E., Giandrea, M. D., Quinn, J. F. (2006). Retirement patterns from career employment. *The Gerontologist*, 46(2), p. 514–523.

Calvo, E., Haverstick, K., Sass, S.A. (2007). *What makes retirees happier: a gradual or 'cold turkey' retirement?* Center for Retirement Research at Boston College Document CRRWP2007-18. Retrieved on the 10th of February 2013 from http://www.bc.edu/crr

Calvo, E., Sarkisian, N. (2011). *Retirement and well-being: examining the characteristics of life-course transitions*. Retrieved on the 10th of February 2013 from www.politicaspublicas.udp.cl

Carleton, R.N., Thibodeau, M.A., Teale, M.J.N., Welch, P.G., Abrams, M.P., Robinson, T.,. Asmundson, G.J.G. (2013). The Center for Epidemiologic Studies Depression Scale: a review with a theoretical and empirical examination of item content and factor structure. *PLoS ONE*, 8(3): e58067. doi:10.1371/journal.pone.0058067

Christ, S.L., Lee, D.J., Fleming, L.E., LeBlanc, W.G., Arheart, K.L., Chung-Bridges, K., Caban, A.J., McCollister, K.E. (2007). Employment and occupation effects on depressive symptoms in older Americans: does working past the age of 65 protect against depression? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(6), p. S399-S403.

Clark, R. L., Quinn, J. F. (2002). Patterns of work and retirement for a new century. *Generations*, 26(2), p. 17-24.

Clarke, P, Marshall, V.W., Weir, D. (2012). Unexpected retirement from full time work after age 62: consequences for life satisfaction in older Americans. *European Journal of Ageing*, vol. 9, p. 207–219.

Coelho, C., Newman, E., Huxtable, D. (2014). A meta-analysis of the relationship between retirement and depression in later life. Unpublished Doctoral Thesis, University of Edinburgh.

Cooke, M. (2006). Policy changes and the labour force participation of older workers: Evidence from six countries. *Canadian Journal on Aging/La Revue Canadienne du Vieillissement*, vol. 25(4), p. 387-400.

** Crawford, M. (1972). Retirement as a psycho-social crisis. *Journal of Psychosomatic Research*, vol. 16, p. 375-380.

Cyr, J.J., Mckenna-Foley, J.M., Peacock, E. (1985). Factor structure of the SCL-90-R: Is there one? *Journal of Personality Assessment,* vol. 49(6), p. 571-578.

Davies, A.R., Sherbourne, C.D., Peterson, J.R., Ware, J.E. (1998). *Scoring manual: adult health status and patient satisfaction measures used in RAND's Health Insurance Experiment*. Santa Monica: RAND Corporation.

De Vaus, D., Wells, Y., Kendig, H., Quine, S. (2007). Does gradual retirement have better outcomes than abrupt retirement? Results from an Australian panel study. *Ageing and Society*, vol. 27(5), p. 667-682.

Derogatis, L. R., Cleary, P. A. (1977). Confirmation of the dimensional structure of the SCL-90: a study in construct validation. *Journal of Clinical Psychology*, vol. 33, p. 981-989.

Dick, H.R., Friedsam, H.J. (1964). Adjustment of residents of two homes for the aged. *Social Problems*, vol. 11(3), p. 282-290.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, vol. 95, p. 542-575.

Diener, E., Emmons, R., Larsen, R.J., Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, vol. 49(1), p. 71-75.

Diener, E., Eunkook, M.S., Lucas, R.E., Smith, H., L. (2009). Subjective well-being: three decades of progress. *Psychological Bulletin*, vol. 125(2), p. 276-302.

** Donaldson, T., Earl, J.K, Muratore, A.M. (2010). Extending the integrated model of retirement adjustment: incorporating mastery and retirement planning. *Journal of Vocational Behavior*, vol. 77, p. 279–289.

Dorfman, L. (1989). Retirement preparation and retirement satisfaction in the rural elderly. *Journal of Applied Gerontology*, vol. 8, p. 432-450.

** Dulin, P.L, Stephens, C., Alpass, F., Hill, R.D., Stevenson, B. (2011). The impact of sociocontextual physical and lifestyle variables on measures of physical and psychological wellbeing among Māori and nonMāori: the New Zealand Health, Work and Retirement Study. *Ageing and Society*, vol. 31, p. 1406-1424.

Eichhorst, W., Gerard, M., Kendzia, M.J., Mayrhuber, C., Nielsen, C., Rünstler, G., Url, T. (2011). *Pension systems in the EU – contingent liabilities and assets in the public and private sector*. Retrieved on the 2nd of March 2013 from http://www.europarl.europa.eu/activities/committees/studies.do?language=EN.

Ekerdt, D.J. (2010). Frontiers of research on work and retirement. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 65B(1), p. 69-80.

Elder, G.H. (1994). Time, human agency, and social change: perspectives on the life course. *Social Psychology Quarterly*, vol. 57(1), p, 4-15.

Elder, G.H., Johnson, M.K., Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer, J.T and Shanahan, M.J. (Eds.), Handbook of the life course, p. 3-19. New York: Kluwer.

Elwood, M. (2007). *Critical Appraisal of Epidemiological Studies and Clinical Trials, 3rd Edition*. Oxford: Oxford University Press.

Farmer, R., Sundberg, N.D. (1986). Boredom proneness – the development and correlates of a new scale. *Journal of Personality Assessment*, vol. 50(1), p. 4-17.

** Floyd, F.J., Haynes, S.N., Doll, E.R., Winemiler, D., Lemsky, C., Burgy, T.M., Werle, M., Heilman, N. (1992). Assessing retirement satisfaction and perceptions of retirement experience. *Psychology and Aging*, vol. 7(4), p. 609-621.

** Fonseca, A.M. (2007). Determinants of successful retirement in a Portuguese population. *Reviews in Clinical Gerontology*, vol. 17(3), p. 219-224.

Fouquereau, E., Fernandez, A., Fonseca, A.M., Paul, M.C., Uotinen, V. (2005). Perceptions of and satisfaction with retirement: a comparison of six European Union countries. *Psychology and Aging*, vol. 20(3), p. 524-528.

** Gall, T.L., Evans, D.R., Howard, J. (1997). The retirement adjustment process: changes in the wellbeing of male retirees across time. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 52B(3), p. 110-117.

Gallo, W.T., Bradley, E.H., Siegel, M., Kasl, S.V. (2000). Health effects of involuntary job loss among older workers: findings from the Health and Retirement Survey. *Journal of Gerontology: Psychological Sciences and Social Sciences*, vol. 55B(3), p. 131-140.

Gana, K., Akremi, M. (1998). L'échelle de disposition à l'ennui (EDE): adaptation Française et validation du Boredom Proneness Scale (BP). *L'année psychologique*, vol. 98(3), p. 429-450.

Gautier, X. (2002). Emploi, retraites et cycles de vie. *Retraite et Société*, vol. 37, p. 163-207.

Gallagher, M.W., Lopez, S.J., Preacher, K.J. (2009). The hierarchical structure of well-being. Journal of Personality, vol. 77(4), p. 1025-1050.

George, L. K. (1993). Sociological perspectives on life transitions. *Annual Review of Sociology*, vol. 19, p. 353-73.

** George, L.K., Maddox, G.L. (1977). Subjective adaptation to loss of the work role: a longitudinal study. *Journal of Gerontology*, vol. 32(4), p. 456-462.

Gilhooly, M.L.M. (1984). The impact of care-giving on care-givers: factors associated with the psychological well-being of people supporting a dementing relative in the community. *British Journal of Medical Psychology*, vol. 57, 35-44.

Gill, S.C., Butterworth, P., Rodgers, B., Anstey, K.J. , Villamil, E., Melzer, D. (2006). Mental health and the timing of men's retirement. *Social Psychiatry and Psychiatric Epidemiology*, vol. 41, p. 515-522.

Goldberg, D.P., Gater, R., Sartorius, N., Ustun, T.B., Piccinelli, M., Gureje, O., Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, vol. 27, p. 191-197.

Goldberg, D.P., Oldehinkel, T., Ormel, J. (1998). Why GHQ threshold varies from one place to another. *Psychological Medicine*, vol. 28, p. 915-921.

Guillemard, A.M., Rein, M. (1993). Comparative patterns of retirement: recent trends in developed societies. *Annual Review of Sociology*, vol. 19, p. 469-503.

Han, S.K., Moen, P. (1999). Clocking out: temporal patterning of retirement. *American Journal of Sociology*, vol. 105, p. 191-236.

Hanisch, K.A., Hulin, C.L. (1990). Job attitudes and organizational withdrawal: an examination of retirement and other voluntary withdrawal behaviors. *Journal of Vocational Behavior*, vol. 37, p. 60-78.

Hardy, M. A. (2002). The transformation of retirement in twentieth-century America: from discontent to satisfaction. *Generations*, vol. 26(2), p. 9-16.

Havinghurst, R.J. (1961). Successful aging. *The Gerontologist*, vol. 1, p. 8-13.

Helmes, E., Goffi, R.D., Chrisjohn, R.D. (2010). Confirmatory analysis of the Bradburn Affect Balance Scale and its relationship with morale in older Canadian adults. *Canadian Journal on Aging / La Revue canadienne du vieillissement*, vol 29 (2), p. 259-266.

** Hervé, C., Bailly, N., Joulain, M., Alaphilippe, D. (2012). Comparative study of the quality of adaptation and satisfaction with life of retirees according to retiring age. *Psychology*, vol. 3(4), p. 322-327.

Higgins, J.P.T., Green, S. (Editors) (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J. P. T., Deeks, J.J. (2011). Chapter 7: Selecting studies and collecting data. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Altman, D.G., Sterne, J.A.C. (2011). Chapter 8: Assessing risk of bias in included studies. In, Higgins, J.P.T., Green, S. (Editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Ramsay, C., Reeves, B.C., Deeks, J.J., Shea, B., Valentine, J.C., Tugwell, P., Wells, G. (2013) Issue relating to study design and risk of bias when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 12-25.

Higo, M., Williamson, J.B. (2008). Retirement. In, D. Carr, R. Crosnoe,, M. E. Hughes, A. M. Pienta (Editors), *Encyclopedia of the life course and human development vol. 3 (Social Sciences)*, p. 328-336. Detroit: MacMillan Reference.

Hofer, S.M., Piccinin, A.M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods*, vol. 14(2), p. 150-164.

Holi, M. (2003). *Assessment of psychiatric symptoms using the SCL-90*. Unpublished Doctoral Thesis, Helsinki University.

Hooker, K., Ventis, D.G. (1984). Work ethic, daily activities, and retirement satisfaction. *Journal of Gerontology*, vol. 39(4), p. 478-484.

** Isaksson, K., Johansson, G. (2000). Adaptation to continued work and early retirement following downsizing: long-term effects and gender differences. *Journal of Occupational and Organizational Psychology*, vol. 73, p. 241-256.

Kerlinger, F.N., Lee, H.B. (2000). *Foundations of behavioural research (Fourth Edition).* Fort Worth: Harcourt College Publishers.

Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly*, 61, 121–140.

Kim, J. E., Moen, P. (2001). Is retirement good or bad for subjective well-being? *Current Directions in Psychological Science*, vol. 10(3), p. 83-86.

Kim, J.E., Moen, P. (2002). Retirement transitions, gender and psychological wellbeing: a life course ecological model. *Journals of Gerontology Series B: Psychological Science and Social Sciences*, vol. 57(3), p. 212-222.

Kim-Prieto, C., Diener, E., Tamir, M., Scollon, C., Diener, M. (2006). Integrating the diverse definitions of happiness: a time-sequential framework of subjective well-being. *Journal of Happiness Studies*, vol. 6, p. 261-300.

King, D.L., Haagsma, M.C., Delfabbro, P.H., Gradisar, M., Griffiths, M.D. (2013). Toward a consensus definition of pathological video-gaming: a systematic review of psychometric assessment tools. *Clinical Psychology Review*, vol. 33, p. 331-342.

Kohli, M., Künemund, H. (2002). La fin de carrière et la transition vers la retraite: Les limites d'âge chronologiques sont-elles un anachronism? *Retraite et société*, vol. 36, p. 81-107.

Kohout, F.J., Berkman, L.F., Evans, D.A., Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D Depression Symptoms Index. *Journal of Aging and Health*, vol. 5, p. 179-193.

Kubicek, B., Korunka, C., Raymo, J.M., Hoonakker, P. (2011). Psychological wellbeing in retirement: the effects of personal and gendered contextual resources. *Journal of Occupational Health Psychology*, vol. 16(2), p. 230-246.

Larson, R. (1978). Thirty years of research on the subjective well-being of older Americans. *Journal of Gerontology*, vol. 33(1), p. 109-125.

Lawton, M.P. (1975). The Philadelphia Geriatric Centre Morale Scale: a revision. *Journal of Gerontology*, vol. 30(1), p. 85-89.

Lawton, M.P. (2003). *Lawton's PGC Morale Scale* [Morale Scale created by M. Powell Lawton (1923-2001) while at the Polisher Research Institute of the Philadelphia Geriatric Center

(now known as the Abramson Center for Jewish Life)]. Retrieved on the 6th of November 2013, from http://www.abramsoncenter.org/PRI/ (Scales page).

Leber, U., Wagner, A. (2007). *Early and phased retirement in European companies: establishment survey on working time 2004–2005*. Retrieved on the 2nd of March 2013 from http://www.eurofound.europa.eu/publications/htmlfiles/ef0688.htm

Lefebvre, C., Manheimer, E., Glanville, J. (2011). Chapter 6: Searching for studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Lipsey, M.W., Wilson, D.B. (2001). *Practical Meta-Analysis*. Thousand Oaks: Sage Publications, Inc.

Löckenhoff, C.E. (2012). Understanding retirement: the promise of life-span developmental frameworks. *European Journal of Ageing*, vol. 9, p. 227-231.

Lohmann, N. (1977). Correlations of life satisfaction, morale and adjustment measures. *Journal of Gerontology*, vol. 32(1), p. 73-75.

Loretto, W. (2010). Work and retirement in an ageing world: the case of older workers in the UK. *21st Century Society*, vol. 5(3), p. 279-294.

Lucas, R.E. (2007). Adaptation and the set-point model of subjective well-being: does happiness change after major life events? *Current Directions in Psychological Science,* vol. 16(2), p. 75-79.

Lucas, R.E., Clark, A.E., Georgellis, Y., Diener, E. (2003). Re-examining adaptation and the set-point model of happiness: reactions to changes in marital status, *Journal of Personality and Social Psychology*, vol. 84(3), p. 527-539.

Luhmann, M., Hoffman, W., Eid, M., Lucas, R.E. (2012). Subjective wellbeing and adaptation to life-events: a meta-analysis. *Journal of Personality and Social Psychology*, vol. 102(3), p. 592-615.

Marshall, V.R., Taylor, P. (2005). Restructuring the life course: work and retirement. In, Johnson, M.L., Bengston, V.L., Coleman, P.G. and Kirkwood, T.B.L. (Editors), *The Cambridge Handbook of Age and Ageing*, p. 572-582. Cambridge: Cambridge University Press.

** Mattila, V.J., Joukamaa, M.I., Salokangas, R.K.R. (1989). Retirement, aging, psychosocial adaptation and mental health: findings of the TURVA project. *Acta Psychiatrica Scandinavica*, vol. 80. p. 356-367.

McDowell, I. (2010). Measures of self-perceived well-being. *Journal of Psychosomatic Research*, vol. 69, p. 69-79.

McDowell, I., Praught, E. (1982). On the measurement of happiness: an examination of the Bradburn Scale in the Canada Health Survey. *American Journal of Epidemiology*, vol. 116(6), p. 949-958.

Mein, G., Martikainen, P., Hemingway, H., Stansfels, S., Marmot, M. (2003). Is retirement good or bad for mental and physical health functioning? Whitehall II longitudinal study of civil servants. *Journal of Epidemiology and Community Health*, vol. 57, p. 46-49.

Mitchell, H., Guled, G. (2010). *Average age of withdrawal from the labour market: A methodology update*. Retrieved on the 3rd of January 2013 from www.ons.gov.uk/ons/rel/pensions/average-age-of-withdrawal-from-the-labour-market/2010/index.html

Moen, P. (1996). A life course perspective on retirement, gender and wellbeing. *Journal of Occupational Health Psychology*, vol. 1(2), p. 131-144.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6), p. e1000097.

Moriwaki, S.Y. (1974). The Affect Balance Scale: a validity study with aged samples. *Journal of Gerontology*, vol. 29(1), p. 73-78.

Morris, J.N., Sherwood, S. (1975). A retesting and modification of the Philadelphia Geriatric Center Morale Scale. *Journal of Gerontology*, vol. 30(1), 77-84.

Neugarten, B.L., Havighurst, R.J., Tobin, S.S. (1961). The measurement of life satisfaction. *Journal of Gerontology*, vol. 16, p. 134-143.

Nordenmark, M., Stattin, M. (2009). Psychosocial wellbeing and reasons for retirement in Sweden. *Ageing & Society*, vol. 29, p. 413-430.

** Nuttman-Shwartz, O. (2004). Like a high wave: adjustment to retirement. *The Gerontologist*, vol. 44(2), p. 229-236.

Oksannen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjosten, N., Virtanen, M., Kawachi, I., Kivimaki, M. (2011). Is retirement beneficial for mental health? Antidepressant use before and after retirement. *Epidemiology*, vol. 22(4), p.553-559.

Organisation for Economic Co-operation and Development (OECD) (2011). Pensionable years, in *Society at a Glance 2011. OECD Social Indicators*. OECD Publishing. Retrieved on 31st of December 2012 from http://dx.doi.org/10.1787/soc_glance-2011-14-en

** Palmore, E., Cleveland, W.P., Nowlin, J.B., Ramm, D., Siegler, I.C. (1979). Stress and adaptation in later life. *Journal of Gerontology*, vol. 34(6), p. 841-851.

Petticrew, M., Roberts, H. (2006). *Systematic Reviews in the Social Sciences: a practical guide*. Oxford: Blackwell Publishing.

Pinquart, M., Schindler, I. (2007). Changes of life satisfaction in the transition to retirement: a latent class approach. *Psychology and Aging*, vol. 22(3), p. 442-455.

** Potocnik, K., Tordera, N., Peiró, J.M. (2010). The influence of the early retirement process on satisfaction with early retirement and psychological well-being. *International Journal of Aging and Human Development*, vol. 70(3), p. 251-273.

** Potocnik, K., Tordera, N., Peiró, J.M. (2013).Truly satisfied with your retirement or just resigned? pathways toward different patterns of retirement satisfaction. *Journal of Applied Gerontology*, vol. 32, p. 164.

Price, C.A. (2003). Professional women's retirement adjustment: the experience of reestablishing order. *Journal of Aging Studies*, vol. 17, p. 341-355.

Quick, H. E., Moen, P. (1998). Gender, employment, and retirement quality: a life-course approach to the differential experiences of men and women. *Journal of Occupational Health Psychology*, vol. 3, p. 44-64.

Reeves, B.C., Deeks J.J., Higgins, J.P.T., Wells, G.A. (2011). Chapter 13: Including non-randomized studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

** Reitzes, D.C., Mutran, E.J. (2004). The transition to retirement: stages and factors that influence retirement adjustment. International Journal of Aging and Human Development, vol. 59(1), p. 63-84.

** Reitzes, D.C., Mutran, E.J. (2006). Lingering identities in retirement. *The Sociological Quarterly*, vol. 47, p. 333-359.

Robbins, S.B., Payne, E.C., Chartrand, J.M. (1990). Goal instability and later life adjustment. *Psychology and Aging*, vol. 5(3), p. 447-450.

Ross, C. E., Drentea, P. (1998). Consequences of retirement activities for distress and sense of personal control. *Journal of Health and Social Behavior*, vol. 39(4). p. 317-334.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of eudaimonic well-being. *Journal of Personality and Social Psychology*, vol. 57, p. 1069-1081.

Ryff, C.D., Keyes, C.L.M. (1995). The Structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, vol. 69(4), p. 719-727.

Sanderson, S., Tatt, I.D., Higgins, J.P.T. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *International Journal of Epidemiology*, vol. 36, p. 666-676.

Sargent, L.D., Bataille, C.D., Vough, H.C., Lee, M.D. (2011). Metaphors for retirement: unshackled from schedules. *Journal of Vocational Behaviour*, vol. 79, p. 315-324.

Sargent-Cox, K.A., Anstey, K.J., Kendig, H., Skladzien, E. (2012). Determinants of retirement timing expectations in the United States and Australia: A cross-national comparison of the effects of health and retirement benefit policies on retirement timing decisions. *Journal of Aging & Social Policy*, vol. 24(3), p. 291-308.

Schalk, R., van Veldhoven, M., de Lange, A.H., de Witte, H., Kraus, K., Stamov-Roßnagel, C., Tordera, N., van der Heijden, B., Zappalà, S., Bal, M., Bertrand, F., Claes, R., Crego, A., Dorenbosch, L., de Jong, J., Desmette, D., Gellert, F.J., Hansez, I., Iller, C., Kooij, D., Kuipers, B., Linkola, P., van den Broeck, A., van der Schoot, E., Zacher, H. (2010). Moving

European research on work and ageing forward: overview and agenda. *European Journal of Work and Organizational Psychology*, vol. 19(1), p. 76-101.

Schiaffino, K.M. (2003). Other measures of psychological well-being: The Affect Balance Scale (ABS), General Health Questionnaire (GHQ-12), Life Satisfaction Index-A (LSI-A), Rosenberg Self-Esteem Scale, Satisfaction with Life Scale (SWLS), and State-Trait Anxiety Index (STAI). *Arthritis & Rheumatism (Arthritis Care & Research)*, vol. 49(5S), p S165-S174.

Schrank, B., Bird, V., Tylee, A., Coggins, T., Rashid, T., Slade, M. (2013). Conceptualising and measuring the well-being of people with psychosis: systematic review and narrative synthesis. *Social Science & Medicine*, vol. 92, p. 9-21.

Shapiro, A., Yarborough-Hayes, R. (2008). Retirement and older men's health. *Generations*, vol 32(1), p. 49-53.

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: why validity issues in meta-analyses will not go away. *Clinical Psychology Review*, vol. 17(8), p. 881-901.

Shills, T. (2008). Early retirement in Germany, the Netherlands, and the United Kingdom: a longitudinal analysis of individual factors and institutional regimes. *European Sociological Review*, vol. 24(3), p. 315-329.

Shultz, K.S., Wang, M. (2011). Psychological perspectives on the changing nature of retirement. *American Psychologist*, vol. 66(3), p. 170-179.

Silver, M.P. (2010): Women's retirement and self-assessed well-being: an analysis of three measures of well-being among recent and long-term retirees relative to homemakers. *Women & Health*, vol. 50, p. 1-19.

Sklar, M., Groessl, E.J., O'Connell, M., Davidson, L., Aarons, G.A. (2013). Instruments for measuring mental health recovery: a systematic review. *Clinical Psychology Review*, vol. 33, p. 1082-1095.

Smith, J.M., Alloy, L.B. (2009). A roadmap to rumination: a review of the definition, assessment, and conceptualization of this multifaceted construct. *Clinical Psychology Review*, vol. 29, p. 116-128.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A., Thacker, S.B. (2000). Meta-Analysis of Observational Studies in Epidemiology: A Proposal for Reporting. *Journal of the American Medical Association*, vol. 283(15), p. 2008-2015.

Szinovacz, M. E. 1980. Female retirement: Effects on spousal roles and marital adjustment. *Journal of Family Issues*, vol. 1, p. 423-440.

Szinovacz, M.E., DeViney, S. (1999). The retiree identity: gender and race differences. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 54B(4), p. S2O7-S218.

Szinovacz, M.E., Davey, A. (2004). Honeymoons and joint lunches: effects of retirement and spouse's employment on depressive symptoms. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* vol. 59B(5), p. 233-245.

Taylor , Shore (1995). Predictors of planned retirement age: an application of Beehr's model. *Psychology and Aging*, vol. 10(1), p. 76-83.

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

The United Kingdom Government (2011). Department of Business Innovation and Skills: Phasing out the default retirement age: government response to consultation. Retrieved on the 8th of April 2013 fromhttp://www.bis.gov.uk

The United Kingdom Government (2013). *HM Treasury: Autumn Statement*. Retrieved on the 5th of December from https://www.gov.uk/government/news/autumn-statement-2013-date-announced-by-chancellor-as2013

Tilghman-Osborne, C., Cole, D.A., Felton, J.W. (2010). Definition and measurement of guilt: implications for clinical research and practice. *Clinical Psychology Review*, vol. 30, p. 536-546.

Tooth, L., Ware, R., Bain, C., Purdie, D.M., Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology*, vol. 161, p. 280-288.

Topa, G., Moriano, J.A., Depolo, M., Alcover, C.M., Morales, J.F. (2009). Antecedents and consequences of retirement planning and consequences: a meta-analysis and model. *Journal of Vocational Behaviour*, vol. 75, p. 38-55.

Tugwell, P., Petticrew, M., Kristjansson, E., Welch, V., Ueffing, E., Waters, E., Bonnefoy, J., Morgan, A., Doohan, E., Kelly, M.P. (2010). Assessing equity in systematic reviews: realising the recommendations of the Commission on Social Determinants of Health. *British Medical Journal*, vol. 341, p. c4739.

Turvey, C., Wallace, R.B., Herzog, R. (1999). Depression: a revised CES-D measure of depressive symptoms and a DSM-based measure of major depressive episodes in the elderly. *International Psychogeriatrics,* vol. 11(2), p. 139-148.

** van Solinge, H., Henkens, K. (2005). Couples' adjustment to retirement: a multi-actor panel study. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 60B(1), S11–S20.

van Solinge, H., Henkens, K. (2007). Involuntary retirement: the role of restrictive circumstances, timing and social embeddedness. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* vol. 62B(5), p. S295-S303.

** van Solinge, H., Henkens, K. (2008). Adjustment to and satisfaction with retirement: two of a kind? *Psychology and Aging*, vol. 23(2), p. 422-434.

van Solinge, H., Henkens, K. (2009). Living longer, working longer? The impact of subjective life expectancy on retirement intentions and behaviour. *European Journal of Public Health*, vol. 20(1), p. 47-51.

Vandewater, E.A., Ostrove, J.M., Stewart, A.J. (1997). Predicting women's well-being in midlife: the importance of personality development and social role involvements. *Journal of Personality and Social Psychology*, vol. 72, 1147–1160.

Veit, C.T., Ware, J.E. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology*, vol. 51(5), p. 730-742.

Vickerstaff. S., Cox, J. (2005). Retirement and risk: the individualization of retirement experiences? *Sociological Review*, vol. 53(1), p. 77-95.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines of reporting observational studies. *Epidemiology*, vol. 18(6), p. 800-804.

** Wang, M. (2007). Profiling retirees in the retirement transition and adjustment process: examining the longitudinal change patterns of retirees' psychological well-being. *Journal of Applied Psychology*, vol. 92(2), p. 455-474.

Wang, M., Bodner, T.E. (2007). Growth mixture modeling identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, vol. 10(4), p. 635-656.

Wang, M., Henkens, K., van Solinge, H. (2011). Retirement adjustment: a review of theoretical and empirical advancements. *American Psychologist*, vol. 66(3), p. 204-213.

Warner, D.F., Hayward. M.D., Hardy, M.A. (2010). The retirement life course in America at the dawn of the twenty-first century. *Population Research Policy Review*, vol. 29, p. 893–919.

Warr, P., Butcher, V., Robertson, I., Callinan, M. (2004). Older people's well-being as a function of employment, retirement, environmental characteristics and role preference. *British Journal of Psychology*, vol. 95, p. 297-324.

Werneke, U., Goldberg, D.P., Yalcin, I., Ustun, B.T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, vol. 30, p. 823-829.

Westerlund, H., Vahtera, J., Ferrie, J.E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M., Zins, M., Kivimaki, M. (2010). Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. *British Medical Journal*, BMJ 2010:341:c6149.

Williams, S.E., Watts, T.K.O., Wade, T.D. (2012). A review of the definitions of outcome used in the treatment of bulimia nervosa. *Clinical Psychology Review*, vol. 32, p. 292-300.

World Health Organisation (1948). Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946. *Official Records of the World Health Organization*, no. 2, p. 100. Retrieved on the 31st of October 2013 from http://www.who.int/about/definition/en/print.html

** Wong, J.Y., Earl, J.K. (2009). Towards an integrated model of individual, psychosocial, and organizational predictors of retirement adjustment. *Journal of Vocational Behavior*, vol. 75, p. 1-13.

SLR Appendices

Appendix 1

Data Search Strategy

1. Bibliographic databases and other electronic Searches – White and Grey Literature

Topical area, geo. location, white (W) or grey (G)	Database	Date	Search Commands (all)	Imported for deduplication and selection
General **W**	The Cochrane Non-Randomised Studies Methods Group Specialised Register	04.03.13 Not available	--	--
W	The Cochrane Central Register of Controlled Trials (CENTRAL)	04.03.13	Retire* Work* AND (Old* OR elder* OR senior*) Employ* AND (Old*OR elder* OR senior*)	W 3475
W	The Centre for Reviews and Dissemination databases of research in health and social care (all – DARE, NHSEED, HTA)	04.03.13	Retirement Work AND (Older OR Elder OR Senior)	W 379
W	The Campbell Collaboration Library	04.03.13	Retirement Employment Work	W 318
G	Open Grey	10.03.12	Retire*	G 545
W	Science Citation Index *and* Social Sciences Citation Index *and* Conference Proceeding Citation Index	10.03.12	1 Retir* 2 Retir* AND (wellbeing* OR well-being OR health* OR adjust* OR adapt* OR happy OR happiness OR satisfy* OR subjective OR quality*) 3 Retir* AND (depress* OR distress* OR ill* OR psych* OR mental*) 4 S2 AND S3 5 (Work* OR Employ*) AND (Old* OR elder* OR senior*) AND (health* OR adapt* OR adjust* OR ill* OR distress* OR wellbeing* OR well-being* OR happiness OR quality OR satisf* OR subjective OR depress* OR mental* OR psych*) 6 S5 limit to (sociology or gerontology or psychiatry or geriatrics gerontology or social sciences	

			interdisciplinary or nursing or psychology or health policy services or psychology developmental or psychology multidisciplinary or psychology clinical or psychology social)	W 24486
W	ScienceDirect	10.03.13	1 Retir* 2 Retir* Ti, Ab, KW 3 Retir* AND (Wellbeing* OR well-being* OR subjective OR adjust* OR adapt* OR happy OR happiness OR satisf* OR “quality of life”) 4 Limit S3 to ("health care, public health, lancet, society, hospital, patient, woman, mental health")	W 8953
W	BioMed Central	09.03.13	1 Retire* 2 Work* AND (Old* OR elder* OR senior*) AND (Health* OR adapt* OR satisf*OR adjust*OR wellbeing* OR well-being* OR quality OR subjective OR illness* OR disease* OR distress* OR anxi* OR depress* OR mental*)	W 12251
W	Scopus	09.03.13	1 Retir* 2 Retir* Ti, Abs, Kw 3 Retir* AND (Wellbeing* OR well-being* OR subjective OR adjust* OR adapt* OR happy OR happiness OR satisf* OR “quality of life”) 4 Work* AND (Old* OR elder* OR senior*) 5 S3 limit to limit-to (subjarea, "soci") or limit-to (subjarea, "psyc") or limit-to(subjarea, "nurs")	W 12556
G	British Library for Development Studies	09.03.12	Retirement KW	G 158
Health and Psychology **W**	MEDLINE (US Library of Medicine)	24.03.13	1 Map term Retire exp Retirement/px, sn [Psychology, Statistics & Numerical Data] 2 Retir$ Ti 3 Retir$ AND (Health$ OR dapat$ OR adjust$ OR wellbeing$ OR well-being$ OR happy$ OR happiness$ OR satisfy$ OR subjective OR depress$ OR psych$ OR mental$ OR disorder$ OR “quality of life”) 4 Work$ OR Employ$ AND (Old$ OR elder$ OR senior$) 5 S3 AND S4	W 14573
W	EMBASE	23.03.13	1 Retire Map Term / Exp Retirement OR medical research OR decision making 2 exp work/ OR exp work disability/ OR exp work resumption/ 3 Map term Employ / exp employment 4 Retir$ AND (wellbeing$ OR depress$ OR Psych$ OR mental$ OR wellbeing$ OR well-being$ OR satisf$ OR	

			subjective OR adapt$ OR adjust$ OR happiness$ OR "quality of life" OR disorder$ OR distress$) NOT Village$ 5 Work$ AND Old$ OR elder$ OR senior$ 6 S4 AND S5	W 7937
W	CINAHL Plus	23.03.13	1 retire exp 2 Retir* AND (wellbeing* OR well-being* OR adjust* OR adapt* OR satisfy* OR qualit* OR subjective OR depress* OR health* OR Illness* OR mental* OR psych* OR distress*) 3 (Work* OR Employ*) AND (Old* OR elder* OR senior*)	W 16333
W	PsycINFO *and* PsycArticles	23.03.13	1 exp retirement/ exp Lifestyle/ or exp Employment Status/ or exp Retirement/ or exp Age Differences/ or exp Job Satisfaction/ or exp Life Changes/ or exp Employee Attitudes/ or exp Intention/ 2 Wellbeing* OR well-being* OR adjust* OR adapt* OR happiness* OR satisfy* OR depress* OR illness* OR "quality of life" OR distress OR mental* OR disease* OR psych* 3 S1 AND S2 4 (Old* OR elder* OR senior*) 5 S2 AND S4 6 S3 AND S4	W 4872
W	Psychology and Behavioural Sciences Collection	23.03.13	1 Retir* 2 Narrow by subject (Narrow by Subject: - older people -- mental health, public health, retirement planning, mental health, geriatric psychiatry, older people – health, retirees, old age, mental depression, quality of life, statistics, aging, geriatrics, older people, medical care, retirement, older people – care) 3 Retir* AND (Wellbeing* OR well-being OR happiness OR satisfy* OR depress* OR Health* OR subjective OR Illness* OR quality OR adjust* OR adapt* OR distress OR mental* OR disease* OR psych*)	W 6832
W	Global Health *and* CAB	23.03.13	1 exp retirement/ retired people or psychology or occupational health or "quality of life" or health or women or public health or retirement or sociology).sh. or man.od. or work satisfaction.sh. or mental health.sh. or risk factors.sh. 2 Old* OR elder* OR senior* 3 S1 AND S2	

			4 Retir* AND (Wellbeing* OR adjust* OR adapt* OR happiness OR satisfy* OR depress* OR well-being* OR Health* OR Illness* OR quality OR distress* OR mental* OR disease* OR psych*)	W 1183
G	Global Health Library (WHO) Geographical sub-indices: Africa (AFRO), Americas (AMRO/PAHO), Eastern, Mediterranean (EMRO), Europe (EURO), South-East Asia (SEARO), Western Pacific (WPRO), Latin America and Caribbean (LILACS)	24.03.12	Retire	G 647
G	NHS Evidence – NICE	01.04.13	Retirement	G 5
G	SIGN	01.04.13	Retirement	G 1
G	King's Fund Publications	01.04.13	Retire Work	G 100
Social Sciences and Social Care **W**	ASSIA	01.04.13	1 Retir* 2 Retir* AND (Wellbeing* OR adjust* OR adapt* OR happiness OR satisfy* OR depress* OR well-being* OR Health* OR Illness* OR quality OR distress* OR mental* OR disease* OR psych*) 3. (Work* OR Employ* OR retir*) AND (Wellbeing* OR adjust* OR adapt* OR happiness OR satisfy* OR depress* OR well-being* OR Health* OR Illness* OR quality OR distress* OR mental* OR disease* OR psych*) AND (Old* OR elder* OR senior*)	W 10830
W	Sociological Abstracts *and* Social Services Abstracts *and* IBSS (international Bibliography of Social Sciences)	01.04.13	1 Retir* 2 Retir* AND (Wellbeing* OR adjust* OR adapt* OR happiness OR satisfy* OR depress* OR well-being* OR Health* OR Illness* OR quality OR distress* OR mental* OR disease* OR psych*) 3. (Work* OR Employ*) AND (Wellbeing* OR adjust* OR adapt* OR happiness OR satisfy* OR depress* OR well-being* OR Health* OR Illness* OR quality OR distress* OR mental* OR disease* OR psych*) AND (Old* OR elder* OR senior*) Ab	W 21845
Work and Business **W**	ABI/INFORM Complete	31.03.13	1 Retir* 2 Retir* AND (wellbeing* OR well-being OR health* OR adjust* OR adapt* OR satisf* OR quality OR happiness OR subjective OR illness* OR distress OR depress* OR psych* OR mental*) Limit to Ab	W 5127

Regional **G**	European Union: European Commission Libraries Catalogue Europeana The European Library	01.04.13	Retirement Ab Retire Ab Retirement AND Health Retirement Retirement AND Health Sb	G 45
W	US Library of Congress Online Catalogue	01.04.13	Retire? AND Health	W 1015
G	US Centres for Disease Control and Prevention	01.04.13	Retirement AND Mental Health	G 2390
G	Department of Work and Pensions (UK)	01.04.13	Retirement and Health	G 6
G	Institute for Fiscal Studies (UK) Search under ELSA	-	-	-
G	Office for National Statistics UK	01.04.13	Retirement Ti	G 13
Theses and Dissertations **G**	ETHOS (UK and Ireland)	14.04.13	Retire Retirement	G 183
G	Dissertations and Theses (Worldwide) (check if each Uni. has it online)		Retir* Ab Retir* AND (wellbeing* OR well-being OR health* OR adjust* OR adapt* OR satisf* OR quality OR happiness OR subjective OR illness* OR distress OR depress* OR psych* OR mental*) Ab	G 500
G	Index to Theses (UK and Ireland)		Retirement	G 367
G	Edinburgh Research Archive	14.04.13	Retirement	G 8

2. Published, unpublished and on-going studies (searches between 21.04.13 and 04.05.13) all Grey Literature

Region/Type of source	Details	New Source	1st Yield
Global	The World Mental Health Survey Initiative	Website	0
UK	English Longitudinal Study of Ageing (ELSA) Wave 6 (current) (2002-current)	Website	5
	Whitehall Study I and II(UK)	Website	0
	British Household Panel Survey	Via UoE	1
	United Kingdom Household Longitudinal Study (UKHLS) or *Understanding Society*	Via UoE	0
	Health Survey for England	Website	0
	1958 National Child Development Study	Website	0
	Scotland and European Health for All Database (2006)?	Website (via KN)	0
Australia and NZ	New Zealand Health Work and Retirement Study	Website	5
	New Zealand Longitudinal Study of Ageing (NZLSA)	Website	0

	Australian National Survey of Mental Health and Wellbeing 2007	Website	0
	The Household, Income and Labour Dynamics in Australia Survey (HILDA)	Website	2
	The Melbourne Longitudinal Studies on Healthy Ageing Program (MELSHA)	Website	0
	The Florey Adelaide Male Ageing Study (FAMAS) 2002-2005	Website	0
	Canberra Longitudinal Study 1990-2002	Website	0
Asia	Korean Longitudinal Study of Ageing (KLOSA)	Website	0
	China Health and Retirement Longitudinal Study (CHARLS)	Website	0
	Beiging Longitudinal Study of Health Ageing	No website	0
	Longitudinal Ageing Study in India (LASI)	Website	0
	Singapore Longitudinal Ageing Studies	Website	0
Europe	Survey of Health Ageing and Retirement in Europe (SHARE)	Website	5
	European Community Household Panel (1994-2001)	Website	0
	European Prospective Investigation into Cancer and Nutrition (EPIC)	Website	0
Ireland	TILDA – Irish Longitudinal Study on Ageing	Website	0
France	Veillissement Sante Travail (VISAT) 1996-??	Website	0
	Enquête santé, travail et vieillissement (ESTEV) 1990-1995	No website	0
	GAZEL Cohort (open cohort) 1989 – on-going	Website	0
	The Constances Cohort – an open epidemiological lab	No website	0
	AMI cohort – Health and Aging in Elderly Farmers	No website	0
Germany	German Socio-Economic Panel (GSOEP) 1984-2011 (on-going)	Website	0
	Berlin Ageing Study (BASE)	Website	0
	Leipzig Longitudinal Study of the Aged (LEILA 75+)	No website	0
Italy	Italian Longitudinal Study on Aging (ILSA)	No website	0
	Italian National Research Council Targeted Project on Ageing	No website	0
Holland	Longitudinal Aging Study Amsterdam (LASA) 1992 – on-going	Website	1
	GLOBE Study – health inequalities	No website	0
Sweden	The Swedish National Study of Aging and Research in Kungsholmen (SNACK)	Website	0
	The Stockholm Birth Cohort	Website	0
	The Stockholm Public Health Cohort	No website	0
Finland	The Helsinki Ageing Study	No website	0
	TURVA project – adjustment to retirement	No website	0
USA			
Institutes an Research Centres	US National Institute on Aging (NIA)	Website	1
	14 NIA Demography of Aging Centres	Institutional Websites	0
	Sloan Centre on Aging and Work Boston College	Website	0
	The Job Stress Network	Website	
Studies	US Health and Retirement Study	Website	0
	Baltimore Longitudinal Study of Aging (1958-current)	Website	0
	Georgia Centenarian Study	Website	0

	The Seattle Longitudinal Study 1956-2005	Website	0
	Cornell Retirement and Wellbeing Study 1952 – ??	No website	1
	Detroit Area Studies 1951 – ??	No Website	0
	The Kaiser-Permanente Retirement Study	No website	0
	Normative Aging Study	No website	0
Canada	Canadian Study of Health and Aging	Website	0
	Fredericton 80+ Study	Website	0
	The Victoria Longitudinal Study	Website	0
Central & South America	Mexican Health and Aging Study (MHAS) 2001-2005 (planned surveys 2012 and 2014)	Website	0
	Puerto Rican Elderly: Health Conditions (PREHCO) Project	Website	0
	Health, Wellbeing and Ageing in Latin America and the Caribbean (SABE)	Website	0

Appendix 2

Data Extraction and Evaluation Form

Full Reference		**STUDY ID**

A.

Data Category	Data Sub-category	Verbatim Support	Additional comments
Study objectives	--		
Study population (age interval; proportion of men) and N	--		
Study design	--		
Construct of Adjustment or Adaptation	Definition		
	Clear vs unclear or variable?*		
	Reference to model/theory related to Retirement?		
	Related to wider literature on Adjustment or Adaptation?		
	Hypotheses based on model/theory? (Yes/No)		
Measurement Instrument	Name		
	Reference(s) provided in record		
	General properties: Number of Items Subscales? Name		
	Psychometric Properties (Sub-Form)		
Main findings summary	--		

* This judgment is based on the presence, anywhere in the report, of a construct or, if this is missing, at least an operationalisation of adjustment or adaptation to retirement; this is considered clear if it is described in terms that are intelligible and consistent throughout.

B.

Property Type	**Property Definition**	**Criteria for Evaluation**	**Notation**
Conceptual Clarity	The extent to which the construct being measured is adequately described and evidence or theory-based	A clear narrative description of the construct being measured is provided with theoretical, model or evidence support.	+ Present - Absent ? Cannot Tell NA Not Applicable
Content Validity	The extent to which the domain of interest is comprehensively sampled by the items in the instrument.	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and item selection.	
Internal Consistency	The extent to which individual items of the instrument are consistent with each other.	Scales measuring a unitary construct expected to have high internal consistency. Multifactorial scales expected to have lower internal consistency. Either α reliability or split half reliability values are > .70.	
Convergent Validity	The extent to which the instrument relates to existing instruments of the same construct.	Convincing arguments for selection of 'gold standard' measure of the same construct, and when compared to this, correlation of at least 0.7 is obtained	
Construct Validity	The extent to which the instrument fits the theoretical construct that it relates to.	Factor analysis is used, and results fit the proposed structure of the instrument. Inter-correlations between items of the same factor are > 0.7. Scores on the instrument correlate with scores on other measures, in a way that is theoretically consistent or predicted.	
Test-Retest Reliability	The extent to which an instrument produces consistent or reproducible outcomes	Test-retest reliability studies produces a weighted Kappa of ≥ .70	
Interpretability	The extent to which qualitative meaning can be assigned to quantitative scores	Mean and Standard Deviation scores are presented for at least two relevant subgroups.	

Appendix 3

Email Correspondence with Authors

To:
Y.Wells@latrobe.edu.au;
...
Dear Professor Wells,

I'm writing to you as the main author of the flowing document:
Wells, Y., deVaus, D., Kendig, H., Quine, S., & Petralia, W. (2006). Healthy Retirement Project: Technical Report [Electronic version]. Retrieved August 8, 2008. Available from: http://www.latrobe.edu.au/alpc/projects/hrp.pdf.

I am currently undertaking a Systematic Literature Review (SLR) on available measures of adjustment to retirement. This is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have tried to access this document, as I would be very interested in including your measure (as it is cited by a few papers that I've reviewed) it in my SLR. However, the location of the document no longer seems to be active. I was wondering if it would be at all possible to ask you for access to this document?

Thank you so much for taking the time to read this and for your help,

All the best,
Claudia.

Claudia Coelho
Specialist Clinical Psychology Practitioner,
Edinburgh Clinical Psychology Services for Older People
National Health Service (NHS) Lothian/University of Edinburgh

To:
COELHO Claudia;
...
You replied on 17/12/2013 07:14.

Dear Claudia,

Thank you for your query.
Yes, La Trobe has changed its website so all the links have changed.
I'd be happy to send you a copy.
I am not in the office until later today but will send you one.
If I forget, please feel free to remind me.

Regards
Yvonne

Appendix 4

Table of Excluded Studies

Full Reference	Reasons for exclusion
Alpass, F. (2008). *Health, Work and Retirement Survey: Summary report for the 2006 data wave. Work and Retirement.* Retrieved on the 7th of April 2013 from http://hwr.massey.ac.nz/resources/Work_Fiona%20Alpass.pdf	The term 'anticipation of retirement adjustment' is used. The measure used is not clear or stated in the report.
Alpass, F.A., Towers, A., Stephens, C.A., Fitzgerald, A.E., Stevenson, B., Davey, J. (2007). Independence, well-being, and social participation in an aging population. *Annals of the New York Academy of Sciences*, vo. 1114, p. 241-250.	The term 'retirement adjustment' is used throughout. However, though results are reported, the measure used is not identified anywhere in the paper.
Anderson, W.F., Cowan, N.R. (1956). Work and retirement: influences on the health of older men. *The Lancet*, vol. 29, p. 1344-1347	The term and concept of happiness are used. There is no identifiable measure, although results are reported. No explicit link with adjustment or adaptation to retirement.
Atchley, R.C. (1976). Selected social and psychological differences between men and women in later life. *Journal of Gerontology*, vol 31(2), p. 204-211.	The term and concepts of 'wellbeing' and 'attitude towards retirement' are used. The latter is measured with a single-item instrument, which uses the words 'like' or 'dislike' (being retired) No explicit link with adjustment or adaptation to retirement
Barfield, R.E., Morgan, J.N. (1978). Trends in Satisfaction with Retirement. *The Gerontologist,* vol. 18(1), p. 19-23	The term and concept of 'satisfaction with retirement' is used consistently throughout. It is measured with a single-item scale. Satisfaction with retirement is the singular construct being used and measured. Therefore, no explicit link with adjustment or adaptation to retirement.
Bellis, M.A., Lowey, H., Hughes, K., Deacon, L., Stansfield, J., Perkins, C. (2012). Variations in risk and protective factors for life satisfaction and mental wellbeing with deprivation: a cross-sectional study. *BMC Public Health*, vol. 12, p. 492-508.	Uses concepts of Mental Well-Being (MWB) and 'life satisfaction' MWB was measured using the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS). Life satisfaction measured with a single-item instrument. No explicit link with adjustment or adaptation to retirement
Bender, K.A. (2012). An analysis of well-being in retirement: the role of pensions, health, and 'voluntariness' of retirement. *The Journal of Socio-Economics*, vol. 41, p. 424-433	The term and concept of 'subjective wellbeing' is used throughout. It is measured with a single-item scale, from the HRS. No explicit link with adjustment or adaptation to retirement
Bonsang, E., Klein, T.J. (2011). Retirement and Subjective Well-Being. Retrieved on the 10th of February 2013 from http://ftp.iza.org/dp5536.pdf	The term and concept of 'subjective wellbeing' is used throughout. It is measured with a set of purpose built likert-scales. No explicit link with adjustment or adaptation to retirement
Börsch-Supan, A., Jürges, H. (2007). *Early retirement, social security and well-being in Germany*. Retrieved on the 10th of February 2013 from http://mea.mpisoc.mpg.de/uploads/user_mea_discussionpapers/ztx3fj26kmz9eqfa_134-2007.pdf	The term and concept of 'subjective wellbeing' is used throughout. It is measured with a set of purpose built likert-scales. No explicit link with adjustment or adaptation to retirement
Braithwaite, V.A., Gibson, D.M. (1987). Adjustment to retirement: what we know and what we need to know. *Ageing and Society*, vol. 7, p.1-18.	Theoretical paper only.
Burr, A., Santo, J.B., Pushkar, D. (2011). Affective well-being in retirement: the influence of values, money, and health across three years. *Journal of Happiness Studies*, vol. 12, p. 17-40.	The term and concept of 'affective wellbeing' is used throughout. It is measured with the Positive and Negative Affect Scale. No explicit link with adjustment or adaptation to retirement

Butterworth, P., Gill, S.C., Rodgers, B., Anstey, K.J., Villamil, E., Melzer, D. (2006). Retirement and mental health: analysis of the Australian national survey of mental health and wellbeing. *Social Science and Medicine*, vol. 62, p. 1179-1191.	Wellbeing is used here as a descriptor of presence/absence of mental health symptoms. The Composite International Diagnostic Interview (CIDI) is the only outcome related to 'wellbeing'. No explicit link with adjustment or adaptation to retirement
Calasanti, T. (1996). Gender and life satisfaction in retirement: an assessment of the male model. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 51B(1), p. S18-S29.	The term and concept of 'life-satisfaction in retirement' is used consistently throughout. It is measured with a validated instrument. Life-satisfaction in retirement is the singular construct being used and measured. Therefore, no explicit link with adjustment or adaptation to retirement.
Calvo, E., Haverstick, K., Sass, S.A. (2007). *What makes retirees happier: a gradual or 'cold turkey' retirement?* Center for Retirement Research at Boston College Document CRRWP2007-18. Retrieved on the 10th of February 2013 from http://www.bc.edu/crr	The term and concept of happiness is used. It is measured with an *ad hoc* selection of questions from the Health and Retirement Study (HRS, US) questionnaire. No explicit link with adjustment or adaptation to retirement.
Calvo, E., Haverstick, K., Sass, S.A. (2009). Gradual retirement, sense of control, and retirees' happiness. *Research on Aging,* vol. 31(1), p. 112-135	The term and concept of happiness is used. It is measured with an ad hoc selection of questions from the HRS questionnaire. No explicit link with adjustment or adaptation to retirement.
Calvo, E., Sarkisian, N. (2011). *Retirement and well-being: examining the characteristics of life-course transitions*. Retrieved on the 10th of February 2013 from www.politicaspublicas.udp.cl	The term and concept of 'wellbeing' is used throughout. Psychological wellbeing is measured with the CES-D (in addition to indicators of economic, social and physical wellbeing). No explicit link with adjustment or adaptation to retirement
Charles, K.K. (2002). *Is retirement depressing? Labour force inactivity and psychological wellbeing in later life.* Working Paper, National Bureau of Economic Research. Retrieved on the 9th of March 2013 from http://www.nber.org/papers/w9033	The term and concept of 'subjective wellbeing' is used throughout. Unclear measurement of the construct, though using HRS (can be inferred). No explicit link with adjustment or adaptation to retirement
Clarke, P, Marshall, V.W., Weir, D. (2012). Unexpected retirement from full time work after age 62: consequences for life satisfaction in older Americans. *European Journal of Ageing*, vol. 9, p. 207–219.	The term and concept of 'life-satisfaction' is used consistently throughout. It is measured with a validated instrument, from the HRS. Life-satisfaction is the singular construct being used and measured. Therefore, no explicit link with adjustment or adaptation to retirement.
Crowley, J.E. (1986). Longitudinal effects of retirement on men's well-being and health. *Journal of Business and Psychology*, vol. 1(2), p. 95-113.	The term and concept of 'subjective wellbeing' is used throughout. It is measured with various validated scales, for the affective and cognitive dimensions of SWB. No explicit link with adjustment or adaptation to retirement
De Vaus, D., Wells, Y., Kendig, H., Quine, S. (2007). Does gradual retirement have better outcomes than abrupt retirement? Results from an Australian panel study. *Ageing and Society*, vol. 27(5), p. 667-682.	The term and construct of 'wellbeing' and 'outcomes of the retirement transition' is used. It is measured using multiple measures/ constructs (e.g. self-esteem, optimism, life satisfaction, marital cohesion) No explicit link with adjustment or adaptation to retirement
Dorfman, L. T. Health Conditions and Perceived Quality of Life in Retirement. *Health & Social Work*, vol. 20(3), p. 192-200.	The term and concept of 'quality of life in retirement' is used consistently throughout. The author is explicit in saying that this is measured by 'retirement satisfaction'. It is measured with a validated instrument to measure retirement satisfaction. Quality of life in retirement is the singular construct being used and measured. Therefore, no explicit link with adjustment or adaptation to retirement.
Dorfman, L.T. (1992). Academics and the transition to retirement. *Educational Gerontology*, vol. 18(4), p. 343-363.	The term retirement satisfaction is used consistently. It uses a specific population, with specific characteristic (academics), no consideration of issues of generalisation No link to adjustment and adaptation to retirement considered
Elgarresta, I.L., de Miguel, M.S., Arruabarrena, L. R. (2009). Diferentes formas de aceder a la jubilación y su	Life satisfaction is used as an indicator of psychological health.

relación con la salud psicológica [Different retirement trajectories and their relationship with psychological wellbeing]. *Revista Española de Gerontologia*, vol. 44(6), p. 311-316.	Measured using the Philadelphia Geriatric Center Morale Scale. No explicit link with adjustment or adaptation to retirement.
Fouquereau, E., Fernandez, A., Fonseca, A.M., Paul, M.C., Uotinen, V. (2005). Perceptions of and satisfaction with retirement: a comparison of six European Union countries. *Psychology and Aging*, vol. 20(3), p. 524-528.	The terms retirement satisfaction and life-satisfaction are used consistently. It is measured using Retirement Satisfaction Inventory (Floyd et al., 1992) No link to adjustment and adaptation to retirement considered
Fouquereau, E., Fernandez, A., Mullet, E. (2001). Evaluation of determinants of retirement satisfaction among workers and retired people. *Social Behaviour and Personality*, vol. 29(8), p. 777-786.	The term retirement satisfaction is used consistently. It is measured using multiple measures/ constructs No link to adjustment and adaptation to retirement considered
Gall, T.L., Evans, D.R. (2000). Preretirement expectations and the quality of life of male retirees in later retirement. *Canadian Journal of Behavioural Science*, vo. 32(3), p. 187-197.	The term and concept of 'quality of life' is used. It is measured with a validated measurement instrument. Some initial discussion in the Introduction of successful adjustment to retirement and pre-retirement expectations, 'quality of life' is the singular construct being used and measured. Therefore, no explicit link is made with adjustment or adaptation to retirement.
George, L.K., Fillenbaum, G.G., Palmore, E. (1984). Sex differences in the antecedents and consequences of retirement. *Journal of Gerontology*, vol. 39(3), p. 364-371.	The term and concept of 'subjective wellbeing' and life satisfaction is used throughout. These are measured with: four-item scale measuring satisfaction with life as whole; multiple measures of subjective well-being and related concepts. No explicit link with adjustment or adaptation to retirement
Hao, Y. (2008). Productive activities and psychological well-being among older adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 63B(2), p. S64-S72	Psychological Wellbeing is used to signify presence/absence of depressive symptoms. Only outcome related to psychological wellbeing is the CES-D No explicit link with adjustment or adaptation to retirement
Herzog, A.R., House, J.S., Morgan, J.N. (1991). Relation of work and retirement to health and well-being in older age. *Psychology and Aging*, vol. 6(2), p. 202-211.	Wellbeing is used to signify presence/absence of mental health symptoms. Well-being variables include an index of physical health, two indexes of mental health consisting of a depression index and an index of life satisfaction, and two indexes of cognitive functioning consisting of a cognitive impairment and a verbal aptitude measure.
Horner, E.M. (2012). Subjective well-being and retirement: analysis and policy recommendations. *Journal of Happiness Studies*. Published Online on 1st December 2012, DOI 10.1007/s10902-012-9399-2	The term and construct of 'subjective wellbeing' is used throughout It is measured using validated measures of SWB/quality of life (CASP 12) and life satisfaction. No explicit link with adjustment or adaptation to retirement is made.
Latif, E. (2011). The impact of retirement on psychological well-being in Canada. *The Journal of Socio-Economics*, vol. 40, p. 373-380.	Psychological Wellbeing is used throughout Only outcome related to psychological wellbeing is a single-item instrument (using the term 'happy/unhappy') No explicit link with adjustment or adaptation to retirement
Marshall, V.W., Clarke, P.J., Ballantyne, P.J. (2001). Instability in the retirement transition: effects on health and well-being in a Canadian study. *Research on Aging*, vol. 23, p. 379-409.	The term and construct of 'wellbeing' is used throughout Only outcome related to wellbeing is a single-item instrument (life satisfaction) No explicit link with adjustment or adaptation to retirement
McGoldrick, A.E., Cooper, C.L. (1994). Health and ageing as factors in the retirement experience. *European Work and Organisational Psychologist*, vol. 4(1), p. 1-20.	The term and concept of 'satisfaction with retirement' is used throughout. It is measured with non-validated likert-type scales. No explicit link with adjustment or adaptation to retirement.
McMunn, A., Nazroo, J., Wahrendorf, M., Breeze, E., Zaninotto, P. (2009). Participation in socially productive activities, reciprocity and wellbeing in later life: baseline results in England. *Aging and Society*, vol. 29, 765-782.	The term and construct of 'wellbeing' is used throughout It is measured using three validated instruments (quality of life, life satisfaction, and depression). No explicit link with adjustment or adaptation to retirement
Moen, P., Erickson, W.A., Agarwal, M., Fields, V., Todd, L. (2000). *The Cornell Retirement and Well-Being Study: Final Report*. Ithaca, New York: Bronfenbrenner Life Course Center, Cornell University. Retrieved on	The term and construct of 'wellbeing' is used throughout It is measured using three instruments (life satisfaction, psychological wellbeing [mastery, attitudes

the 4th of March 2013 from http://worlddatabaseofhappiness.eur.nl/hap_bib/freetexts/moen_p_2000.pdf	towards ageing, self-esteem] and depression). No explicit link with adjustment or adaptation to retirement
Nordenmark, M. (1999). Non-financial employment motivation and well-being in different labour market situations: a longitudinal study. *Work Employment Society*, vol. 13(4), p. 601-620	The term mental wellbeing is used, but, not predominantly. It is measured using the GHQ No link to adjustment and adaptation to retirement considered
Nordenmark, M., Stattin, M. (2009). Psychosocial wellbeing and reasons for retirement in Sweden. *Ageing & Society*, vol. 29, p. 413-430.	Terms post-retirement wellbeing and psychosocial wellbeing are used consistently. Index of psychosocial wellbeing 'constructed' out of multiple individual questions. No link to adjustment and adaptation to retirement considered
Oliveira, C., Torres, A.R.R., Simões de Albuquerque, E. (2009). Análise do bem estar psicossocial de aposentados de Goiânia [Analysis of psychosocial well-being of retired men in Goiânia]. *Psicologia em Estudo*, vol. 14(4), p. 749-757	The term and construct of 'psychosocial wellbeing' is used throughout It is measured using a purpose built, validated instrument. No explicit link with adjustment or adaptation to retirement
Pinquart, M., Schindler, I. (2007). Changes of life satisfaction in the transition to retirement: a latent class approach. *Psychology and Aging*, vol. 22(3), p. 442-455.	The term and concept of 'life satisfaction' and trajectories of life satisfaction is used. It is measured with a single-item scale. Some initial discussion in the Introduction of adjustment to retirement and its determinants, however, 'life-satisfaction' is the singular construct being used and measured. Therefore, no explicit link is made with adjustment or adaptation to retirement.
Price, C.A, Balaswamy, S. (2009). Beyond health and wealth: predictors of women's retirement satisfaction. *International Journal of Aging and Human Development*, vol. 68(3), p. 195-214	The term and concept of retirement satisfaction is used consistently throughout. It is measured by Retirement Satisfaction Inventory (Floyd et al., 1992) Retirement satisfaction is the singular construct being used and measured. No explicit link is made with adjustment or adaptation to retirement.
Quick, H. E., Moen, P. (1998). Gender, employment, and retirement quality: a life-course approach to the differential experiences of men and women. *Journal of Occupational Health Psychology*, vol. 3, p. 44-64.	The terms and concepts of 'retirement quality' and satisfaction in retirement are used interchangeably in the Introduction. It is measured with two single-item instruments. Retirement quality is the construct being used and measured. Therefore, no explicit link is made with adjustment or adaptation to retirement.
Reis, M., Pushkar-Gold, D. Retirement, personality, and life satisfaction: a review and two models. *Journal of Applied Gerontology*, vol. 12(2), p. 261-282.	Theoretical paper only.
Robbins, S.B., Payne, E.C., Chartrand, J.M. (1990). Goal instability and later life adjustment. *Psychology and Aging*, vol. 5(3), p. 447-450.	The study is a validation study of the Goal Instability Scale on a group of retirees, it is considered to be a predictor of life satisfaction in later life. However, it is not a measure of adjustment and a distinct construct.
Robinson, O.C., Demetre, J.D., Corney, R. (2010). Personality and retirement: exploring the links between the Big Five personality traits, reasons for retirement and the experience of being retired. *Personality and Individual Differences*, vol. 48, p. 792–797	The terms and concepts of adjustment to retirement, satisfaction with retirement, life satisfaction and enjoyment of retirement are used inter used interchangeably in the Introduction and Results sections. Retirement experience questionnaire and life satisfaction are measured with two instruments/scales, though the first in not validated. Retirement experiences and life satisfaction are the constructs being used and measured. Therefore, no explicit link is made with adjustment or adaptation to retirement.
Shultz, K.S., Morton, K.R, Weckerle, J.R. (1998). The influence of push and pull factors on voluntary and involuntary early retirees' retirement decision and adjustment. *Journal of Vocational Behavior,* vol. 53, p. 45-57.	The term and concept of 'life satisfaction', 'satisfaction with retirement' and adjustment is used initially. However, only results for the first two measures are reported, but the measurement instruments have not been identified anywhere in the paper. No elaboration ore results re. adjustment or adaptation.

Silver, M.P. (2010): Women's retirement and self-assessed well-being: an analysis of three measures of well-being among recent and long-term retirees relative to homemakers. *Women & Health*, vol. 50, p. 1-19	The term and construct of 'wellbeing' is used throughout It is measured using three instruments (depressive symptoms, finantial worries and bad health). No explicit link with adjustment or adaptation to retirement
Smith, D.B., Moen, P. (2004). Retirement satisfaction for retirees and their spouses: do gender and the retirement decision-making process matter? *Journal Of Family Issues*, Vol. 25 No. 2, March 2004 262-285	The term and concept of 'retirement satisfaction' is used. It is measured with a single-item scale. Some initial discussion in the Introduction of adjustment to retirement and its determinants, however, 'retirement-satisfaction' is the singular construct being used and measured. Therefore, no explicit link is made with adjustment or adaptation to retirement.
Szinovacz, M.E. (1980). Female retirement: effects on spousal roles and marital adjustment. *Journal of Family Issues*, vol. 1(3), p. 423-440.	Term of 'retirement satisfaction' is used, as well as some initial discussion of patterns of adjustment (and gender specific characteristics) No identifiable measure used; the author calls this study 'exploratory'; though statistics are reported, it becomes eventually clear in the results section that this is a qualitative study.
Szinovacz, M.E., Davey, A. (2004). Honeymoons and joint lunches: effects of retirement and spouse's employment on depressive symptoms. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 59B(5), p. 233-245.	The terms transition and wellbeing is used in this paper, related to the individual and to the couple. Depressive symptoms are the main dependent variable and outcome measure Despite some initial discussion of adjustment phases, no explicit link with adjustment or adaptation to retirement is made.
Szinovacz, M.E., Davey, A. (2005). Retirement and Marital Decision Making: Effects on Retirement Satisfaction. *Journal of Marriage and Family*, vol. 67(2), p. 387-398	The term and concept of retirement satisfaction is used consistently throughout. It is measured by single-item scale from HRS. Retirement satisfaction is the singular construct being used and measured. No explicit link is made with adjustment or adaptation to retirement.
Vaillant, G.E., DiRago, A.C., Mukamal, K. (2006). Natural history of male psychological health, XV: retirement satisfaction. *American Journal of Psychiatry*, vol. 163, p. 682-688.	The term and concept of satisfaction with retirement is used throughout. It is measured by a set of likert-type scales, using both enjoyment with/and satisfaction with retirement as perms. Retirement satisfaction is the singular construct being used and measured. No explicit link is made with adjustment or adaptation to retirement.
Wahrendorf, M., Siegrist, J. (2010). Are changes in productive activities of older people associated with changes in their well-being? Results of a longitudinal European study. *European Journal of Ageing*, vol. 7, p. 59-68.	The term and construct of 'wellbeing' is used throughout It is measured using a measure of quality of life (CASP 12) No explicit link with adjustment or adaptation to retirement
Warr, P., Butcher, V., Robertson, I., Callinan, M. (2004). Older people's well-being as a function of employment, retirement, environmental characteristics and role preference. *British Journal of Psychology,* vol. 95, p. 297-324.	The term and construct of 'wellbeing' is used throughout It is measured using a purpose built, validated composite measure of life satisfaction No explicit link with adjustment or adaptation to retirement
Wells, Y., De Vaus, D., Kending, H., Quine, S. (2009). Health and wellbeing through work and retirement transitions in mature age: understanding pre–post and retrospective measures of change. *International Journal of Aging And Human Development*, vol. 69(4), p. 287-310, 2009	Constructs of wellbeing and change in wellbeing are used throughout, first in ageing studies then in relation to a sample of retirees. It tests multiple indices of wellbeing (physical, psychological, social) No link to adjustment and adaptation to retirement considered
Zenger, M., Brähler, E., Berth, H., Stöbel-Richter, Y. (2011). Unemployment during working life and mental health of retirees: results of a representative survey. *Aging and Mental Health*, vol. 15(2), p. 178-185.	The term and concept of 'life-satisfaction' is used consistently throughout. It is measured with a validated instrument. Life-satisfaction is the singular construct being used and measured. No explicit link with adjustment or adaptation to retirement
Total Excluded: 55 Total Retained: 22	

Appendix 5

Characteristics of Included Studies

Study ID + reference (brief)	Objectives	Sample charact. (brief)	N	Design	Construct of adjustment or adaptation	Theory/Model Reference	Reference context (Retirement or General)	Measure	Ref. provided in study	Summary of main findings
1 Beck (1982)	To investigate effect of retirement on happiness with life, and determinants of satisfaction with retirement.	USA 'representative probability sample of male population' Age: 55-69 100% Men	3348	Longitud.	Inconsistent constructs used: 'an aspect of psychological well-being (…) happiness with life'; 'satisfaction with life', 'satisfaction with retirement'	Poor conceptual elaboration. Justification of measure/questions only. 1. Life-happiness distinguished from life satisfaction, but high correlation between the two acknowledged. 2. The evaluation of retirement experience in relation to prior expectations is expected to be highly correlated with satisfaction in retirement.	Retirement	Two non-validated, single questions: 1. Happiness 2. Post-retirement assessment of pre-retirement expectations: "All in all, how does your life in retirement compare with what you expected it to be?	NA	1. No significant main effect of retirement 2. Negative bivariate relationship between retirement and happiness with life 3. Poor health, lower income, and early retirement main determinants of negative evaluations of retirement.
2 Braithwaite et al. (1986)	Exploratory study of retirees reporting poor adjustment to retirement, 'of the relationship of factors leading up to retirement, problems encountered upon retirement, and adjustment after retirement'.	Australia. Random sample of community dwelling individuals (census sampling frame) Age: >60 56% Men	487	Cross-sectional	Term adjustment to retirement used throughout. Conceptualised as a 'multidimensional construct'. No further elaboration.	Reference to Continuity Theory and 'individual differences' in adjustment styles. Hypotheses: 1. retirement adjustment is 'better conceptualized as a multi-dimensional construct' 2. 'there are distinct styles of poor adjustment that have different consequences'	Retirement	Non-validated single questions on: 1. Reaction to retirement 2. Problems encountered with retirement 3. Quality of life 'indices' on health, income, involvement, activities, life satisfaction and mental health.	NA	1. Identified four poor adjustment styles (poor health, negativism, change adaptation and retirement reluctance). 2. Latter two responses to retirement tend to report problems in the short term; the former two may have long term consequences.

3 Crawford (1972)	Exploratory study of ageing and experience of retirement, in relation to predictions from two retirement adjustment theories: disengagement and activity theories	UK Non-random, convenience sample of married couples Median age: 62-64 Unclear gender %	53	Longitud.	Adjustment to retirement used throughout. No further elaboration.	Reference to Disengagement Theory and Activity Theory. Hypotheses derived from each one: 1. (activity theory) 'the individual who is prepared to drop the role of worker and to take up others should be well-adjusted and healthy' 2. (disengagement theory) 'the individual who is willing to lose the role of worker and to reduce his total role count should be happy and healthy'.	Retirement	Non-validated questions (unclear how many): 'Anticipation of and adjustment to retirement are represented by two global measures (…) collapsed for convenience into two categories – positive and negative'. Measurement construct very inconsistent in report, positive and negative refer to 'attitudes', 'aspects of retirement', etc	NA	1. 'After retirement, 52% of men and 38% women were positive' – unclear in relation to what. 2. Non-significant changes in pre-post evaluations of retirement experience or attitudes towards retirement (very unclear).
4 Donaldson et al. (2010)	1. to determine if individual characteristics (demographic and health), psychosocial (mastery and planning) and organizational factors (conditions of workforce exit) predict retirement adjustment; 2. To investigate the relationship between planning, mastery and retirement adjustment.	Australia Semi-retired and retired men and women Age: >45 50% men	570	Cross-sectional	Very clearly stated: 'retirement adjustment is conceptualised as a person's positive retirement experiences (Atchley, 1999)'.	Clearly located. Study makes direct reference to Wong and Earl (2009) and Role theory. Hypotheses: both contextual and psychosocial factors influence adjustment".	Retirement	Operationalized using the 13-item retirement adjustment measure. (Cronbach's alpha =.81; Wells et al., 2006). In current study, Cronbach's alpha = .88.'	Wells, DeVaus, Kendig, Quine, Petralia, (2006).	1. Higher income, better psychological and physical health predict better retirement adjustment. 2. Controlling for demographics and health, higher personal sense of mastery and better exit conditions predict adjustment to retirement. 3. Pre-retirement planning not related to retirement adjustment. 4. Post-retirement planning effect on adjustment was mediated by mastery.

5 **Dulin et al. (2010)**	Overview of the New Zealand Health, Work and Retirement Study on cultural-contextual factors which influence on physical and mental health of older individuals.	New Zealand National representative sample. Age: 55-70 45.6% men	6662	Longitud.	Adjustment to retirement used throughout, but not clearly defined. The measure used seems to be about 'anticipation of retirement adjustment', but this is not clear.	Reference to positive ageing models (Hill, 2010) .	General	A four-item measure on respondents' beliefs about their ability to make the retirement transition successfully, assessing: levels of confidence, anxiety and depression, associated with thoughts of retirement. But, this study used the measure with retired participants. 'Taylor and Shore (1995) report Cronbach's alpha for the scale of 0.86; alpha for the present sample = 0.89'.	Taylor and Shore (1995)	1. Socioec. status, social support and retirement status associated with optimal ageing. 2. Maori scored lower on markers of physical and mental health, partially explained by restrictive factors (reduced financial resources, access to health care and physical activity). 3. After controlling for multiple variables, ethnicity predicted health disparities.
6 **Floyd et al. (1992)**	To develop a measure of 'retirement satisfaction and perceptions of retirement-related experiences'. Two studies: development and validation (psychometric properties) of the scale.	USA Non-random, convenience sample, geographically diverse. Men Age: 67 and 69 (men) Approx. 50% men	Study 1: 50 Study 2: 402	Cross-sectional	Clearly defined. 'Retirement as a life transition imposed a temporal perspective on the measure: a. to assess past experiences and feelings about the transition, b. present satisfaction in retirement c. prospect for future adjustment.	Reference to life-span development theory and life-span transition theory (Baltes, 1987; Fiske and Chiriboga, 1990)	Retirement + General	Retirement Satisfaction Inventory Measure assessed 6 areas: pre-retirement work functioning, adjustment and change, reasons for retirement, satisfaction with life in retirement, current sources of enjoyment, leisure and physical activities.	Floyd et al. (1992)	1. Factor analyses produced internally consistent subscales. 2. Moderate but acceptable test-retest reliability. 3. Satisfaction scores correlated with concurrent measures. 4. Discriminated 4 groups of voluntary and involuntary retirees.
7 **Fonseca (2007)**	Provides descriptive information on adjustment to retirement, highlighting different patterns of adjustment to this process	Portugal Heterogeneous, non-random sample, geographically diverse. Mean age: 67 47% men	502	Cross-sectional	Clearly defined. 'The degree of success in the adjustment to retirement is measured by the impact of events and influences on the life of the retiree'.	Reference to life-span development theory. Reference to past research evidence.	General Retirement	Retirement Satisfaction Inventory	Floyd et al. (1992) Fouquereau, Fernandez, Mullet, (1999) Fouquereau, Fernandez,	1. Retirement is not, in itself related to 'psychological damage', in the short term; 2. As retirement progresses, as well as age, satisfaction with life in retirement decreases; 3. subjects retired for >

									Fonseca et al. 2005	9 years have less 'capacity for enjoyment' 3. 3 patterns of adjustment to retirement (openness-gains, vulnerability-risk, losses-withdrawal)
8 **Gall et al. (1997)**	1. Evaluate the impact of retirement, 2. Monitor the change in adjustment across time, 3. Identify resources predictive of short- and long-term adjustment in retirement'.	Canada Non-random, convenience sample Mean age: 69 100% men	117	Longitud.	Unclear concept of adjustment to retirement: a. As change (positive or negative) in all or each of the measured domains (below). Paper's lit. review suggests that adjustment = good outcomes in measures. b. Domains measured: physical health, mental health, locus of control, retirement satisfaction, life-satisfaction. c. Very unclear in the analysis what variables are used as predictors and as dependent variables.	Explicit reference to two: Hypotheses from Atchley's model: 1. the impact of retirement will be positive in the short term (1 year post) with increases in physical and psychological health and satisfaction; 2. Initial increases in adjustment will stabilize or decrease in long-term (6-7 years). Hypotheses from Resource Theory: 1. resources for successful adjustment will differ from short- to long-term retirement; 2. internal locus of control will predict short-term adjustment, physical health will predict long-term adjustment.	Retirement	Multiple measures: 1. Summary Health Measure 2. Kaizer Illness Index 3. SCL-90 4. Retirement Descriptive Index (RDI): measures satisfaction with 4 areas in retirement: activities and work, financial situation, health, interpersonal 5. Life satisfaction: Single-question: "In general, how satisfying do you find the way you're spending your life today? (completely satisfying, pretty satisfying, or not very satisfying?)"	RDI Smith, Kendall and Hulin (1969)	1. Positive impact of retirement in well-being during first year. 2. Evidence of a retirement adjustment process, well-being (interpreted as psychological health) changed from short- to long-term retirement. 3. Physical health, income, and voluntary retirement predicted short-term adjustment; 4. Changes in resources over time also differentially predicted short- and long-term adjustment.
9 **George and Maddox (1977)**	Study on the adaptation to 'loss of central life-roles'.	USA Non-random, purposive sample. Age: unclear 100% Men	58	Longitud.	Retirement conceptualised as loss of central life role. Adaptation is operationalised as	Reference to lit. on 'life events, stress, and their correlates': 'This perspective emphasizes both the coping behaviour	General Retirement	Modified 5-item version of the 7-item Kutner Morale Scale 'to maximize the reliability of the measure'.	Kutner, Fanshel, Togo and Langner (1956)	1. High levels of adaptation in the sample, as reflected in stability of morale over time. 2. Social resources

					changes in morale: 'morale is an indicator of adaptation'. Morale considered an indicator of 'subjective aspects of wellbeing' and 'a measure of subjective adaptation'	occasioned by life events rather than events per se, and the multivariate factors which mediate the adaptive or maladaptive outcomes which follow life events'.		'Reliabilities were acceptably High': .89 at Time 1 and .95 at Time 2.		(marital status and socioeconomic status) moderate the relationship between adaptation and retirement.
10 Herve et al. (2012)	To compare the quality of adaptation and the satisfaction with life of two groups of adults retiring at different ages (before the age of 60, and after 60).	France Convenience sample Mean age: 71 and 72 (those retiring after 60) Unclear gender %	579	Cross-sectional	Adaptation to retired life conceptualised as changes in SWB and 'satisfaction with life'.	Reference to research evidence only.	Retirement	Measures of changes in SWB and satisfaction with life measured by: 1. Satisfaction With Life Scale 2. Ryff's scales of psychological well-being 3. French adaptation of the Boredom Proneness scale 4. 26 items from the OARS methodology re. physical health 5. 2 single-items measured satisfaction with retiring conditions: a. time of retirement (early, normal or late) b. satisfaction with his/her age at retirement'	Diener, Such, Lucas, Smith, (1999) Ryff and Keyes, (1995), Gana and Akremi (1998). Duke University (1978)	Two samples differed only in their own perception and appreciation of the age at which they retired.
11 Isaksson and Johansson (2000)	Comparison of early retirees and persons continuing to work, following work-place downsizing, in relation to satisfaction, well-being, health, and work centrality.	Sweden Workers from a large insurance company downsizing its workforce. 2 groups compared: early retirees and 'stayers'	370: 224 (early retirees) 146 (stayers)	Longitud.	Variable throughout: a. Initial interchangeable use of evidence on adaptation, wellbeing and 'distress'. b. Hypotheses operationalise adjustment in terms	Reference to Atchley's model of adjustment to retirement. Hypotheses not directly related to this: 1. Voluntary choice will lead to better adjustment in terms of general satisfaction, health and well-being, both among retirees and	Retirement	Multiple measures: 1. Satisfaction with outcome of downsizing, single item: 'How satisfied are you at the present time with the outcome?' 2. Psychological distress/well-being. (GHQ–12). Reliability (C. alpha) was .86 at Time 1	Goldberg and Williams (1988)	1. Voluntary (as opposed to forced) choice directly and positively associated with satisfaction, psychological well-being and health for both groups. 2. Females showed lower values of work centrality, were more

		Age: unclear 58% Men.			of: 'general satisfaction, health and wellbeing' and 'adaptation in terms of distress and health complaints' c. 'Adjustment was defined in terms of general satisfaction with downsizing outcome, psychological well-being, general health complaints, and work centrality'	stayers. 2. Stayers will generally report more distress and health complaints than retirees following downsizing.		and .81 at Time 2. 3. Health Symptom Checklist (Andersson, 1986). C. alpha = .66 on both times. 4. Work centrality measured using one item from the Meaning of Working (1987) studies: 'How important is work at the present time'.		inclined to apply for retirement, and were generally more satisfied with outcomes.
12 Mattila et al. (1989)	Preliminary findings of a prospective study (from 1982-1986, 1st wave)	Finland Randomly selected from pop. register community dwelling individuals Age: >61 at recruitment Unclear gender %	200	Longitud.	No clear conceptualisation. But stated 'success in adjustment is measured in terms of mental and physical health, general functional capacity and activity in interpersonal and leisure time, and general life satisfaction.' Thus, it can be inferred	Tangential reference to ageing models; retirement is seen as one of the adjustments inherent to the process of ageing. Not specific.	General	Multiple measures: 1. GHQ 36 2. Life Satisfaction Index A 3. Lawton's Philadelphia Geriatric Centre Morale Scale 4. Physical health measures not described.	Goldberg (1972) Neugarten et al. (1961) Lawton (1975)	1. 'Participants' biopsychosocial situation' remained constant, retirement is not a stressor of crucial importance or a life-crisis.
13 Nuttman-Shwartz (2004)	Phenomenological description of pre- and postretirement experience (qualitatively), and examination of adjustment to retirement (quantitatively) during two periods.	Israel Non-random sample Age: 65 (at recruitment) 100% Men	56	Longitud.	Retirement conceptualised as a 'life transition'. Construct of adjustment to retirement is unclear.	Location of the study as 'based on premises of': 1. life-span developmental theories (Baltes); 2. Atchley's model of adjustment to retirement. Not used to generate hypotheses.	Retirement	Multiple measures: 1. Modified version of the Multidimensional Health scale on subjective perceptions of physical health. Internal consistency for the 6 questions in the present study (alpha = .71). 2. Mental Health Inventory (MHI) to measure subjective well-	Antonovsky (1985) Drory, Florian, Kravetz (1991 Veit and Ware (1983)	1. In pre-retirement 4 main perceptions of retirement: crisis/uncertainty, hope for change, continuation, transition. 2. Dominant perception as period of uncertainty/crisis. 3. Comparison of pre- and post-retirement scores on standardized

								being, separate scores calculated for wellbeing and distress (alpha .95 for distress; alpha .94 for well-being).		tests shows reduced distress.
14 Palmore et al. (1979)	To examine effects of five major life-events (retirement, spouse's retirement, major medical events, widowhood, and departure of last child from home) and of three types of resources (physical, psychological, and social), on the physical and social-psychological adaptation.	USA Probability sample, stratified by age and sex, drawn from the a list of the major health insurance provider Age: 45-70 Unclear gender % 'Black people excluded'	375	Longitud.	Retirement conceptualised as a stressful event; and adaptation 'as the outcome of attempts to use various resources to cope with the stresses of life events'.	Reference to 'various models of adaptation in late life (…) emphasis on the multiple determinants of adaptation to stress and the importance of physical, psychological and social resources as mediating variables which affect the level of adaptation to the stressful event'. 'Crisis orientation' in psychosocial adaptation is specifically outlined.	General	Multiple measures: 1. single-item/question on life-satisfaction, using a 10 point Cantrill (self-anchoring) scale. 2. Affect Balance Scale 3. Frequency of nine psychosomatic symptoms 4. Self-esteem, measured by two sets of adjectives with semantic differential: Respect and Useful. 5. Activity during typical week.	Cantril, (1965) Bradburn and Caplovitz, (1965) Bradburn (1969) Palmore, (1974).	1. Medical events impacted on physical adaptation, but not on social-psychological adaptation. 2. Retirement had the most negative social-psychological effects. 3. Multiple events accumulate in impact. 4. Better physical resources helped physical adaptation, better psychological and social resources helped satisfaction. 5. 'Most potential stressors have less long-term outcomes than crisis orientation would suggest'.
15 Potocnik et al. (2010)	To explore the influence of early retirement on adjustment, measuring individual characteristics and social context. Proposes an integrated model of factors	Spain Sample of early retirees, non-random selected from attendants of University (of 3rd Age type) cohort. Mean age: 62.5 71% Men	213	Cross sectional	Clearly defined: 'we examine satisfaction with early retirement as a specific indicator of adjustment, and psychological well-being as a more general indicator of adjustment'.	Hypothesis related to Role theory: 'Early retirement is a work role exit process during which individuals are susceptible to different sources of role expectations (individual, organizational and group levels)'	Retirement	Two measures: 1. Satisfaction with early retirement, assessed by 5 items, adapted from Bussing et al. (1999); C. alpha = .92. 2. Psychological well-being, GHQ 12 (C. alpha = .88)	Bruggerman (1976) cited in Bussing, Bissels, Fuchs, and Perrar (1999) Goldberg, (1979)	1. Organizational pressures related to lower retirement age, low perceived ability to continue working and group norms favourable to early retirement 2. Group norms favourable to early retirement and low perceived ability to continue working predicted satisfaction with early retirement. 3. Satisfaction with

										early retirement was related to psychological well-being, moderated by voluntariness and organisational pressures for early retirement.
16 Potocnik et al. (2013)	1. To identify different retirement satisfaction forms and antecedents of observed retirement satisfaction forms, 2. To explore their impact on psychological well-being.	Spain Non-random sample, as in Potocnik et al. (2010) Mean age: 64 68% Men.	270	Cross-sectional	Adjustment to retirement is used consistently: 'Satisfaction with retirement' '(is) one of the most studied retirement Adjustment indicators'.	Reference to 'dynamic model of job satisfaction' (Büssing, 1992; Büssing and Bissels, 1998; Büssing et al., 1999). Parallel is made between retirement and work models, 'an individual develops a certain degree of retirement satisfaction or dissatisfaction based on the match between expectations about retirement and the actual retirement situation'.	Retirement General	As in Potocnik et al. (2010) an adaptation to retirement of the 'Job Satisfaction Questionnaire, shortened also. Tool measures 'the degree, intensity, and dynamics of job satisfaction': a. Total job satisfaction b. Psychological well-being at work c. Changes in levels of aspirations d. Forms of (dis)satisfaction. 'Questionnaire is a collection of single items rather than a coherent scale, reliability coefficients only partially apply to this scale'. No further information provided.	Bruggeman n (1976), in German only	1. 4 retirement satisfaction forms identified: stabilized-progressive, resigned-stabilized, resigned retirement satisfaction, constructive-fixated retirement dissatisfaction. 2. Gender, retirement intentions, and voluntariness of retirement transition predicted retirement satisfaction forms. 3. Participants in the constructive-fixated retirement dissatisfaction reported lower psychological well-being.

17 Reitzes and Mutran (2004)	To explore prospectively stages in retirement adjustment, and the impact of social psychological, social background, and gender factors.	USA Random sample, based on public records. Age: 58-64 (at recruitment) 60% Men	826	Longitud.	Clear conceptualisation: 'we used positive retirement attitudes (Atchley & Robinson, 1982) as an indicator of retirement adjustment'	Direct reference and testing of Atchley's model of retirement adjustment. Hypothesis derived: 'There will be an initial increase in positive attitudes toward retirement at six months post-retirement, followed by a decline in positive attitudes at the one year post-retirement, and an increase in positive attitudes toward retirement by two years post-retirement'. Further hypotheses derived from Role theory.	Retirement General	Measure 'positive attitudes toward retirement': "I think retirement means being . . . ," followed by 14 adjective pairs (e.g., sad-happy, idle-busy, full-empty). 'The 14 adjective pairs were organized into a 5-point semantic differential format and coded, with responses closest to the positive adjectives receiving scores of 5). Alpha reliability coefficients ranged from .91 for pre-retirement to .92, .94, and .93 for the 3 postretirement measures'.	Atchley and Robinson (1982)	1. General support for Atchley's model of retirement adjustment. 2. Multiple factors influence retirement adjustment: a. pre-retirement self-esteem, pension eligibility, increased positive attitudes post-retirement; b. retirement planning and voluntary retirement increased positive attitudes toward retirement in the short, but not long term; c. poor health decreased positive attitudes, in the long term, but not short term.
18 Reitzes and Mutran (2006)	To explore effects of preretirement identities on retirement adjustment and post-retirement self-esteem.	As Reitzes and Mutran (2004)	826	Longitud	As above	As Reitzes and Mutran (2004)	Retirement General	Same measure as Reitzes and Mutran (2004). Additional info. on measure used: 1. 'Alpha coefficients ranged from .92, .94, and .93 for the three retirement periods 2. Relationship between retirement adjustment and a single item measuring happiness in retirement at each of the three measurement times, found r = from .488–.607 range.	As Reitzes and Mutran (2004)	1. 'Preretirement identities and social background characteristics influenced initial retirement adjustment and later changes in retirement adjustment'.

19 **van Solinge and Henkens (2008)**	To explore, measure and test differentially the fit of constructs of adjustment to and satisfaction with retirement.	Netherlands Panel study; sampling from large companies, with self-selection Age: >55 50% Men	778	Longitud.	Conceptualisation is rich and clear. 1. Differentiation of adjustment to retirement and satisfaction with retirement on the basis of the developmental tasks required: a. adjustment involves social and psychological detachment from work and dealing with loss. b. satisfaction involves the development of a satisfactory post-retirement lifestyle. 2. The two constructs are assumed to be 'contingent on the context in which the transition is made (access to resources and characteristics of the transition) and psychological factors'.	Reference to life-course perspective on retirement (i.e. contextualisation of the event and experience). Hypotheses in accordance: 1. Psychological factors would be more important for understanding adjustment 2. 'Access to resources (such as health and income) and changes in these resources would be more salient in understanding retirement satisfaction'	Retirement General	Two dependent variables used to 'investigate the subjective experience of retirement': 1. Adjustment represented by three items (adjustment scale) (alpha =.80) 'capturing difficulties that he or she had in adjusting to retirement'. 2. Satisfaction represented by 4 items (satisfaction scale) (alpha .65), 'capturing contentment with retirement'. Correlation between the two variables r =.50, p>.001.	Van Solinge and Henkens (2008)	1. 'Adjustment and satisfaction are related, but not identical' constructs. 2. Adjustment problems arise from preretirement anxiety about the social consequences of retirement, and from lack of control about the decision. 3. 'Retirement satisfaction is primarily related to access to key resources (finances, health, and the marital relationship). 4. Retirement transition is multidimensional, with two developmental challenges: a. adjustment to the loss of the work role and the social ties of work, b. development of a satisfactory postretirement lifestyle.
20 **van Solinge and Henkens (2005)**	To examine adjustment as a contingent process on individual resources and conditions of the transition. To examine couples' adjustment to	As van Solinge and Henkens (2008)	559 (and partners)	Longitud.	The construct of adjustment is used throughout, but not operationalised clearly.	As van Solinge and Henkens (2008), reference to life-course perspective on adjustment to retirement, and its contextualisation. Specifically the relationship context is	Retirement General	Measure of adjustment as van Solinge and Henkens (2008) Additional psychometric info.: Alpha = .82 M = 3.89, SD = 2.28	van Solinge and Henkens (2005)	1. Partners' influence on each other's adjustment is 'limited' 2. Strong 'attachment' to work (full-time jobs, long work histories), lack of control over the transition, negative preretirement

	retirement.					focused on here: the interdependency of adjustment processes between partners.				expectations, and low self-efficacy predict difficult adjustment.
21 Wang (2007)	The study predicted and tested retirement adjustment patterns, based on 3 theoretical models.	USA Health and Retirement Study 2 samples Random, representativ e sample.	994 + 1066	Longitud.	Retirement adjustment is conceptualised as 'change patterns of retirees' psychological well-being'. Later operationalised as symptoms of depression, measured by CES-D	Role theory, continuity theory, and life-course perspective are explicitly used to 'form hypotheses regarding different retirement transition and adjustment patterns'. Hypotheses elaborated accordingly.	Retirement	CES-D 8; yes/no response format is used. 'Previous studies using HRS data have suggested good reliability of this 8-item scale (e.g., Siegel et al., 2003).	Radloff (1977)	1. '3 latent growth curve patterns of retirees' psychological well-being identified. 2. Subgroups of retirees corresponded to different growth curve patterns of change in psychological well-being. 3. 'Retirees do not follow a uniform adjustment pattern during the retirement process, which reconciles inconsistent previous findings'.
22 Wong and Earl (2009)	1. To examine 3 predictors of retirement adjustment: 'individual (demographic and health), psychosocial (work centrality), and organizational (conditions of workforce exit). 2. To examine the effect of work centrality on post-retirement activity levels.	Australia Non-random sample Age: 45-93	394	Cross-Sectional	Construct of retirement adjustment used consistently, but not clearly operationalised.	Role theory is used to support hypotheses related to work centrality; other hypotheses related to previous research evidence, but not theoretical elaboration.	Retirement	Retirement adjustment assessed with 13-item measure reported by Wells et al. (2006). 'The scale has been supported with evidence of high internal consistency (Cronbach's a = .81). In the current study, C. alpha = .83'	Wells et al. (2006)	1. 'Better psychological health, higher income, and being married predict better retirement adjustment'. 2. Work centrality was not a predictor of retirement adjustment or activity levels. 3. Exit conditions significantly predicted retirement adjustment.

Appendix 6

Properties and psychometric evaluation of self-report measures of adjustment or adaptation to retirement

Measure	Reference: Source and other (if necessary)	Available?	Brief description: N. of items, subscales, response format, score	Concept clarity	Content validity	Internal consist.	Conver. validity	Constr. validity	Test-R reliab.	Interp.
Retirement Satisfaction Inventory	Source: Floyd et al. (1992)	•	Items: 51 Measures: 'satisfaction in retirement and perceptions of retirement experiences relevant to post-retirement adjustment' Subscales: pre-retirement work functioning, adjustment and change, reasons for retirement, satisfaction with life in retirement, current sources of enjoyment, leisure and physical activities Response: 7-point Likert scale Score: Global and subscales	+	+	+	-	-	- (.68, range .56-.77)	+
Retirement Adjustment Questionnaire [A]	Source: Wells et al. (2006) [A] Other: Donaldson et al. (2010), Wong and Earl (2009)		Items: 13 Measures: adjustment to retirement, perceptions of the experiences of leaving work and retiring Subscales: none Response: 5-point semantic scale Score: Total adjustment score (range 13-65); higher scores = better adjustment	+	?	+	?	?	?	?
Expected adjustment to retirement sub-scale	Source: Taylor and Shore (1995)	•	Items: 4 Measures: 'expected adjustment to or ability to successfully make the retirement transition', for retirement-eligible individuals. Subscales: NA Response: 5-point scale (no other detail) Score: scores on each item compute a continuous variable (range: 4 to 22), with higher scores representing higher levels of anticipated adjustment to retirement (e.g. more confidence in one's ability to make the transition).	+	?	+	?	?	?	?
Retirement Descriptive Index	Source: Smith, Kendall, Hulin (1969) Other:		Items: 63 Measures: affective responses to retirement with 4 areas, captured in subscales	-	?	?	?	?	+	?

	Hooker and Ventis (1984) Dorfman (1989) Hanish and Hulin (1990)		Subscales: activities and work, financial situation, health, and interpersonal relationships. Response: ? Score: ?							
Retirement Adjustment and Satisfaction Scales	Source: van Solinge and Henkens (2007)	•	Items: 7 Measures: differential experiences of adjustment and satisfaction with retirement Subscales: Adjustment (3 items), satisfaction (4 Items) Response: unstandardised/variable length scales Score: scales 'linearly transformed' into 0-10 score; high score on adjustment scale indicated few difficulties in adjusting to retirement; high score on the satisfaction scale indicated high satisfaction with retirement.	+	-	+ re. adjust. - re. satisfact.	?	?	?	?
Positive Attitudes Towards Retirement	Source: Atchley and Robinson (1982) Other: Reitzes and Mutran (2004, 2006)		Items: 14 Measures: attitude/meanings of retirement; lead phrase, "I think retirement means being . . . ," followed by 14 adjective pairs (e.g., sad-happy, idle-busy, full-empty) Subscales: ? Response: 5-point semantic differential format Score: ?	+	?	?	?	?	?	?
Job-Satisfaction Questionnaire Short Form [B]	Source:Bruggemann (1976) German Other: Bussing, Bissels, Fuchs, Perrar (1999)	•	NA	NA	NA	NA	NA	NA	NA	NA
Boredom Proneness Scale (French version)	Source: Gana and Akremi (1998) Other: Farmer and Sundberg (1986)	•	Items: 28 Measures: boredom proneness Subscales: none Response mode: true/false Score: range 1-24	+	+	-	?	?	+	?
CES-D 8	Source: Radloff (1977) Other: Kohut et al. (1993) Turvey et al. (1999) Gallo et al. (2000)	•	Items: 8 (item selection from original 20: 6 indicated presence of depression; 2 items indicated absence of depression); Measures: depressive symptoms; shorter version designed for use with older populations (> 65) Subscales: none Response mode: binary (1/0) Score: sum of scores (range 0-8), higher values indicate greater distress	+	+	+	+	+	+	+
GHQ 12 and 36	Source: Goldberg (1972) Goldberg (1978); Other:		Items: 12 and 36 (multiple versions) Measures: a psychiatric disorders screen in community and non-psychiatric clinical settings; assessment of	+	+	+	+	+	+	+

	Goldberg, Garter, et al. (1997) Goldberg, et al. (1998) Werneke, et al. (2000)	current state vs 'usual' state. Subscales: variable; 12 item version, depression and social function; no stable structure for 36 item found In literature Response mode: 4-point scale Scoring: 4 methods – binary (0/1), split binary for negative and positive items, likert (0-3), modified likert (0-2)							
SCL-90	Source: Derogatis, Lipman, Covi (1973) Other: Derogatis and Cleary (1977) Cyr, et al. (1985) Holi (2003)	Items: 90 Measures: symptom intensity on nine different subscales; it has been widely used in different ways: as assessment of psychiatric case-ness; or as descriptive measure of psychopathology in different patient populations; the dimensional/factorial structure of the instrument is highly contested. Subscales: Somatization, Obsessive-compulsive, Interpersonal sensitivity, Depression, Anxiety, Hostility, Phobic anxiety, Paranoid ideation, Psychoticism Response mode: 5-point scale (rate of occurrence of the symptom). Score: Global index of distress/Global Severity Index (GSI), which is the mean value of all items; Positive Symptom Distress Index (PSDI), Positive Symptom Total (PST); separate scale scores also computed	+	+	+	+	?	+	+
MHI	Source: Veit and Ware (1983) Other: Davies, et al. (1998)	Items: 38 Measures: psychological distress and wellbeing, designed for general populations Subscales: (hierarchical factor model) Anxiety, Depression, Loss of Behavioral • Emotional Control, General Positive Affect, Emotional Ties, Psychological Distress, Psychological Well-Being, Mental Health Index Response mode: 6-point scale Score: Complex scoring method; scale scores computed by summing scores over all items in the scale; high scores are consistent with each scale name. Six subscales – Anxiety, Depression, Loss of Behavioural / Emotional Control, General Positive Affect, Emotional Ties and Life Satisfaction; Two global scales - Psychological Distress and Psychological Well-being; and A global Mental Health Index score. (variable ranges)	+	+	+	+	+	+	+

Kutner Morale Scale	Source: Kutner, Fanshel, Togo and Langner (1956) Other: Dick and Friedsam (1964) Lohmann (1977) Gilhooly (1984) Helmes (2010)		Items: 7 Measures: morale as a uni-dimensional construct, in older populations. This scale has been used much less extensively than the PGCMS. Psychometric info. very unclear. Subscales: none Response mode: dichotomous scale (0/1) Score: sum of scores in each item (range 0-7)	-	?	?	-	?	?	?
Lawton's Philadelphia Geriatric Centre Morale Scale (PGCMS)	Source: Lawton (2003), Lawton (1975) Other: Morris and Sherwood (1975) Larson (1978) Ryff (1989) McDowell (2010)	•	Items: 22 (original), 17 (revised version) Measures: morale, designed to measure morale among the very old (70-90), and was based primarily on an institutional population. High morale is defined as a basic sense of wellbeing, satisfaction with past life achievements/attainments, feeling useful now, and thoughts of being a "an adequate person" (Lawton, 1975). Subscales: Agitation, Attitude Toward Own Aging, Lonely Dissatisfaction. Response mode: dichotomous scale (0/1) Score: global sum of scores (range 0-17)	+	+	-	-	?	+	+
Affect Balance Scale	Source: Bradburn (1969) Other: Moriaki (1974, OA only); Larson (1978) McDowell and Praught (1982); Ryff (1989); Schiaffino (2003) McDowell (2010)		Items: 10 Measures: positive and negative affect, in reaction to day-to-day life events or stressors; indicates life satisfaction and/or psychological well-being, designed for general population. Psychometric info. is unclear and contradictory in relation to different populations. Subscales: Positive affect, Negative Affect, and Affect Balance (difference between Positive and Negative Affect) used as indicator of overall happiness or wellbeing. Response mode: scale, or binary (1/0) Score: negative and positive affect scores calculated separately (range 0–5) for positive higher scores indicate higher positive and higher negative affect.	-	?	-	?	?	+	+
Satisfaction With Life Scale	Source: Diener, Emmons, Larsen, Griffin (1985) Other: Schiatffino (2003), Kim-Prieto et al. (2005) McDowell (2010)	•	Items: 5 Measures: global judgment of life satisfaction, rather than satisfaction with specific domains, designed for general population. Subscales: none Response mode: 7-point scale Score: sum of scores in each item (range 0-35); Absolute	+	+	+	-	+	+	+

			or relative life satisfaction: 20 = neutral point on the scale; 31-35 extremely satisfied; 26–30 satisfied; 21–25 slightly satisfied; 15–19 slightly dissatisfied; 10–14 dissatisfied; 5–9 extremely dissatisfied.							
The Ryff's Scales of Psychological Well-Being	Source: Ryff and Keyes (1995) Other: McDowell (2010)	•	Items: 18 (3 items per scale) Measures: focused on successful ageing, a six-component measure, corresponding to a model of personal growth and psychological well-being (part of the eudaimonistic approach to well-being) Subscales: self-acceptance; positive relations with others; autonomy; environmental mastery; purpose in life; personal growth. Response mode: 6-point scale Score: sum of scores for the three subscales (range 0-18, each subscale).	-	+	-	+	?	+	?
Life-Satisfaction Index A	Source: Neugarten, Havinghurst, Tobin (1961) Other: Schiatffino (2003) McDowell (2010)	•	Items: 20 Measures: life satisfaction in older adults, in 5 domains: zest (vs. apathy), resolution and fortitude, congruence between desired and achieved goals, positive self-concept and mood tone. Designed for general population, it aims to capture successful ageing Subscales: none Response mode: scale, with 3 levels Score: global score (range 0–40), higher scores indicate greater life satisfaction.	-	+	+	-	?	+	+

Key: + Present, - Absent, ? Cannot Tell (If the information is not available in the paper or additional sources), NA (Not Applicable)
Notes:
Source reference is 'Available' (marked with a dot) if source/original reference could be obtained through reasonable methods (including, journals, databases, wider www sources, or contact with main author).
[A] Author contacted two times, awaiting response at the time of conclusion/writing of the present review.
[B] This scale was highly adapted from another domain of study, and very shortened (from 12 items to a "selection" of 5) by Potocnik et al. (2010) and Potocnik et al. (2013); therefore it was excluded from detailed evaluation.

Exploring what we know about retirement: A meta-analysis of the relationship between retirement and depression in later life

Authors: Claudia Coelho[ab], David Huxtable[c], Emily Newman[a]

[a] Section of Clinical and Health Psychology, School of Health in Social Science, University of Edinburgh

[b] NHS Lothian

[c] NHS Grampian

Address correspondence to:

Claudia Coelho,

Royal Edinburgh Hospital,

Morningside Terrace

Edinburgh

Word Count: 18493

Abstract

Objectives: To systematically evaluate the evidence of the relationship between retirement and self-reported symptoms of depression in later life.

Methods: A comprehensive list of electronic databases, and additional sources, were searched from December 2012 to March 2013. A meta-analysis and systematic assessment of risk of bias were carried out.

Results: 8 non-randomised studies were included in the meta-analysis, 5 cohort and 3 cross-sectional studies. Studies were grouped and analysed according to the two design-type subgroups There was evidence of high dispersion of effect sizes, variable risk of bias and methodological and statistical heterogeneity between studies in both sub-groups – cohort (Q=640.728, df =4, p<0.001), cross-sectional (Q=76.611, df=2, p<0.001). Summary effects were therefore not meaningful. Sensitivity and sub-group analyses did not account for high heterogeneity of effect sizes.

Conclusions: The relationship between retirement and self-reported depressive symptoms seems to be complex and variable. Effect-sizes of individual studies were small, non-significant and highly dispersed, and heterogeneity of true effects was high. These results may be limited by confounding factors in primary studies. This is discussed and contextualised in relation to the use of non-randomised studies in meta-analysis.

Word count (200 max.): 184

Key words (6 maximum):

RETIREMENT, DEPRESSION, OLDER ADULTS, META-ANALYSIS, NON-RANDOMISED STUDIES

1. Introduction

1.1 Defining retirement

In most industrialised and developing societies, retirement from work constitutes a crucial transition, typically defined as a normative withdrawal from the work force in late adult life (e.g. Atchley,1982; Moen, 1996; Szinovacz and DeViney, 1999; Hardy, 2002; van Solinge and Henkens, 2007; Jokela, Ferrie, et al., 2010; Wang, Henkens, van Solinge, 2011; Kubicek, Korunka, Raymo, Hoonakker, 2011). Retirement is, in itself, a recent (19^{th}-20^{th} century) notion, event or social institution (Atchley, 1982; Marshall and Taylor, 2005; Higo and Williamson, 2008; Warner, Hayward, Hardy, 2010). It has also evolved radically over the last 60 years, from a more or less predictable, uniform transition around a specific age, to a variable and more individually determined event (Guillemard and Rein, 1993; Han and Moen 1999; Vickerstaff and Cox, 2005; van Solinge and Henkens, 2009). Contemporary work-life has become much more diverse. The cessation or withdrawal from active labour has also become much more heterogeneous in timing, path, permanency and form (Clark and Quinn, 2002; Cahill, Giandrea, Quinn, 2006; Ekerdt, 2010). As an institution, retirement is both an instrument and a consequence of social and economic policy. While 30 years ago the early transition into retirement was actively encouraged (e.g. Angelini, Brugiavini, Webber, 2009), current economic and political strategies focus on the exact opposite, encouraging the extension of working lives, economic contribution of older workers, and delay of the retirement transition (Shills, 2008; Brugiavini, Pasini, Peracchi, 2008a; Behncke, 2012). Despite this diversity and looming radical changes to its timing and process[6], retirement is still, as stated above, a normative (i.e. expected) transition in late adulthood (Kohli and Künemund, 2002). There is still an enduring expectation that, around the age of 60 or 65, workers are able end their working lives (Organisation for Economic Co-operation and Development, OECD, 2011).

1.2 Retirement as topic of theory and research

1.2.1 Theoretical context

As described in Coelho and Newman (2014), theories of retirement generally dovetail with theories of ageing and life-span development, or present retirement as a specific instance of ageing. Retirement is conceptualised as a meaningful transition in developmental terms, which carries specific demands and consequences for the individual and social structures around him or her (e.g. Elder, Johnson, Crosnoe, 2003; Higo and Williamson, 2008; Calvo and Sarkisian, 2011; Löckenhoff, 2012). Theories

[6] See Coelho and Newman (2014) for further detail.

in this area include: disengagement theory (Cumming and Henry, 1961, cited in Higo and Williamson, 2008); activity theory (Havinghurst, 1961); continuity theory (Atchley, 1971, 1989); role theory (e.g. Moen, 1996; Kim and Moen, 2002); life-course or life-span development perspectives on retirement (e.g. Moen, 1996; Elder et al., 2003); and resource theory (e.g. Wang, 2007; Wang et al., 2011). Overall, these use psychological and sociological constructs to model the impact that retirement may have on individuals, or their post-retirement adjustment (for a recent review see Coelho and Newman, 2014).

1.2.2 Empirical research context

Research and modelling of retirement has cut across various fields of social, political and medical sciences. It has been obviously dependent on, and contextualised by, the way retirement in itself is defined and operationalised. The current heterogeneity addressed above, regarding its timing, permanency, pathway and form, carries with it difficulty in defining, operationalising and measuring retirement (Denton and Spencer, 2009) in a way that is valid, reliable or useful. Age alone does not provide this, despite the meaning that it still carries for individuals (Elder, 1994; Moen, 1996; van Solinge and Henkens, 2009; Schalk, van Veldhoven, et al., 2010). Nonetheless, retirement has been a focus in large epidemiological studies of ageing in which the challenge to use appropriate measures of the construct is critical and evolving (Ekerdt and DeViney, 1990; Stang, 2008; Houser and Weir, 2010). Such operationalisations have been questioned in their validity, as they are used across genders (e.g. Price, 2000, Price and Nesteruk, 2010), ethnic/cultural settings (e.g. Zsembik and Singer, 1990; Luborski and LeBlanc, 2003), or professional groups (e.g. Thelin and Holmberg, 2010). Generally, most large studies on retirement have adopted multifaceted criteria and measurement types, including objective indicators, routine records and self-reports of current work status (Ekerdt and deViney, 1999). However, these large scale studies have also been criticised for their reductionist ('one size fits all') views of retirement (e.g. Szinovacz and DeViney, 1999; McVittie, McKinlay, Widdicombe, 2008; Sargent, Bataille, Vough, Lee, 2011; McVittie and Goodall, 2012).

As a topic of research, retirement has tended to be treated as a temporal phenomenon, i.e. as a process, not a state. Retirement has also become increasingly understood as an individual and internalised psychological phenomenon (e.g. van Solinge and Henkens, 2007, Shultz and Wang, 2011, Sargent, Bataille, Vough, Lee, 2011), i.e. as a late-life developmental transition. As a simplifying overview, research on retirement has tended to treated it, both, as a dependent and independent variable (Ekerdt, 2010; Shultz and Wang, 2011). As the former, retirement has been operationalised as behaviour or as a decision-making process. These studies have focused on the interaction of push-pull factors between work life and anticipated retirement life (van Solinge and Henkens, 2009), as predictors or moderators of the retirement decision or process (economic, labour market, political, demographic, cultural, social factors; or person-specific financial circumstances, work, age, health, individual differences, etc.) (Atchley, 1979; Topa, Moriano, Depolo, Alcover and Morales, 2009; Ekerdt, 2010;

Sargent-Cox, Anstey, Kendig, Skladzien, 2012). As an independent variable, research has sought to characterize and understand *'the course of the retirement and post-retirement experience' [focusing on] an array of pre-retirement variables and enclosing structural contexts to account for variations in the way that retirement goes'* (Ekerdt, 2010:71). Research in this sense has focused: a) on the physical and mental health effects or correlates of retirement for individuals; b) on the impact of retirement on social structures; and b) on individuals' patterns of transition, adjustment or adaptation to retirement.

1.3 Research and evidence on the effects of retirement

The present meta-analysis in interested specifically on the mental health, effects or correlates of retirement for individuals. Therefore, it is located in that specific segment of this literature.

Retirement as a process carries wide-ranging consequences for individuals and those around them. Among others, Gill, Butterworth et al. (2006), Christ, Lee et al. (2007) and Alavinia and Burdorf (2008) have argued that, in younger populations, the relationship between loss of work participation and physical and psychological health is relatively well understood. Although this is complex, bi-directional, and some inconsistency is found in the data, there is evidence to suggest a detrimental relationship between unemployment and physical and mental health (Wilson and Walker, 1993; McKee-Ryan, Song, Wanberg, Kinicki, 2005; Paul and Moser, 2009; Jeffris and Nazareth et al., 2011). Comparatively, less is understood about this relationship in later life (Moen, 1996). This is not due to lack of interest in this area. On the contrary, the impact of retirement has been widely researched since the 1950s (van Solinge and Henkens, 2008; Shultz and Wang, 2011). It has included the examination of the impact of retirement on men and women (e.g. George, Fillenbaum, Palmore, 1984; Calasanti, 1996), or the manipulation of the retirement process (e.g. early vs late retirement) or work-life characteristics and the study of their impact after the transition (e.g. Buxton, Singleton, Meltzer, 2005) – e.g. voluntary vs involuntary retirement (e.g. Villamil, Huppert, Melzer, 2006; van Solinge and Henkens, 2007); occupational grade or type (e.g. Hyde and Ferrie et al., 2004). However, the relationship between retirement and physical and psychological health has been difficult to characterise, as trends in evidence of this relationship are very inconsistent (e.g. Midanik, Soguikian, Ranson, Tekawa, 1995; Gall, Evans and Howard, 1997; Pinquart and Schindler, 2007; Jokela et al., 2010; Oksannen, Vahtera, et al., 2011). In what specifically concerns this meta-analysis, knowledge about the mental health effects or correlates of retirement is fragmentary.

The impact of retirement has been studied, predominantly, through observational studies. The lack of randomised controlled trials in this area may be understood as a consequence of the topic itself (Reeves, Deeks, Higgins, Wells, 2011), as a naturally-occurring, observable 'exposure' rather than an intervention. The predominance of these kinds of studies brings with it challenges related to the

compatibility and comparability of measures, data, analyses and outcomes, and their amenability for review and reduction (Hofer and Piccinin, 2009). This area is characterised by studies of large cohorts of individuals, over sequential waves of data collection, focusing on multiple aspects of ageing, of which retirement is one. These include past and on-going studies, among others: the Health and Retirement Study (HRS, United States of America), English Longitudinal Study of Ageing (ELSA), Survey of Health, Ageing and Retirement in Europe (SHARE), GAZEL cohort studies (France), China Health and Retirement Longitudinal Study (CHARLS), The Melbourne Longitudinal Studies on Healthy Ageing Program (MELSHA). In addition to these, this area also includes a large number of cross-sectional studies.

Overall, the recurrent upshot in numerous qualitative and partial (i.e. not systematic) reviews of the literature is the characterisation of the impact of retirement as a negative, positive, both or neutral event for individuals' psychological wellbeing or mental health (e.g. Bosse, Aldwin, Levenson, Ekerdt, 1987; Reitzes, Mutran, Fernandez, 1996; Drentea, 2002; Mein, Martikainen, Hemingway, Stansfels, Marmot, 2003; Wang et al.. 2011; Oksannen, et al. 2011, Luhmann, Hoffmann, Eid, Lucas, 2012). This variability in outcome is typically attributed to variations in the methodologies used, selection of samples, operationalization of retirement, selection of dependent variables, selection of moderator variables and their measurement, and the uncontrolled impact of (unknown) confounders (e.g. Wang, 2007; Topa et al., 2009; Behncke, 2012). The following methodological issues, observed recurrently in primary studies, have been of particular concern. First, the variability in the choice of mental health related outcome constructs and measures found across studies (e.g. Ross and Drentea, 1998), leads to questions about what is actually being researched. Second, the existence of data emerging from cross-sectional studies (typically, comparing samples of retirees to samples of workers), raises questions over the level of confounding in these studies – i.e. the presence of systematic differences between samples, which are not controlled for (Luhmann, et al., 2012). Third, coexisting with data generated in cross-sectional studies, data emerging from large cohort studies (e.g. Westerlund et al., 2010) raises concerns over the commensurability of outcomes in a review process (Lipsey and Wilson, 2001). Fourth, the so called 'endogeneity bias' in the study of relationships between health and retirement (e.g. Coe and Lindeboom, 2008; Coe and Zamarro, 2011; Calvo, Sarkisian, Tamborini, 2013). Establishing or making inferences about causal or sequential relationships between retirement and health (physical or mental) is difficult, as '*the retirement decision is not exogenous to health'* (Behncke, 2012:282). The observation of poor health after retirement, cannot be un-problematically causally attributed to retirement itself (Calvo et al. 2013). Fifth and final, the unavoidable co-occurrence of ageing and retirement makes distinguishing between health patterns associated with retirement from those associated with ageing very challenging (Westerlund et al. 2010).

1.4 The current meta-analysis – rationale

An integrative, systematic review and data synthesis in this area has not been done to date, and this meta-analysis proposes to contribute to this area in this way. As retirement is a topic that cuts across boundaries of different disciplines, it aims to synthesise data from diverse areas (Luhmann et al., 2012). In examining the literature, it is noticeable that, with rare exceptions (e.g. Calvo and Sarkisian, 2011), authors' use of theoretical concepts and empirical findings is compartmentalised in the discipline they are broadly located in (e.g. Behncke, 2012; Fe and Hollingsworth, 2012). This lack of integration and, crucially, consensus in size and direction of effects across primary studies, makes a meta-analysis a useful methodological choice (Sharpe, 1997; Petticrew and Roberts, 2006, Elwood, 2007). However, the inclusion of observational/Non-Randomised Studies presents specific challenges. This meta-analysis seeks to both a) characterise the relationship between retirement and mental health in later life; and b) explore the variability in this relationship found within this body of literature (Elwood, 2007).

1.4.1 Why the focus on retirement

Currently, retirement is a key issue, both to individuals' lives and to social and economic decision-making. Gill et al.. (2006), Ekerdt (2010) or Weller (2010), among others, emphasise that this is a point in time when two macro-factors collude to make the understanding of the impact of retirement in later life a pressing concern. The first, the coming of age of a large cohort of older individuals (so called 'baby boomers'), who have begun to reach the age of statutory retirement (e.g. Villanueva, 2000; Kuerbis and Sacco, 2012). The second, current social and economic policies, which encourage the extension of work participation and the delay of retirement altogether (e.g. Solinge and Henkens, 2009, Maimaris, Hogan, Lock, 2010, Loreto, 2010). Holding this wider context in mind, this meta-analysis is specifically concerned with the significance of retirement to individuals' mental health, namely, to its relationship with symptoms of depression or depression in later life. To ignore the potential impact of retirement (in all its current complexities) on older individuals' mental health, would be as surprising as ignoring the mental health impact of unemployment in younger populations (McKee-Ryan et al., 2005; Paul and Moser, 2009).

1.4.2 Why the focus on the relationship between retirement and depression

It is widely quoted that, by 2020, depression is expected to be the highest ranking disease in the developed world (WHO, 2001). Clinically, depression or depressive symptoms are the most common mental ill-health presentations in older individuals (e.g. Blazer, 2003; Smit, Ederveen, Cuijpers, Deeg, Beekman, 2006; Buber and Engelhardt, 2011; Luppa, and Sikorski, et. al, 2011). Although adults over 65 present relatively low prevalence rates of major depressive disorder (Karel and Hinrichsen, 2000), with one-year rates between 1% and 5% (Fiske, Wetherell, Gatz, 2009), depression is more likely to

present in sub-clinical forms (according to diagnostic criteria like DSM-IV) and be under-recognised in this population (Karel and Hirichsen, 2000; Blazer, 2003; Luppa, and Sikorski, et. al, 2012). Blazer (2003) locates the prevalence of clinically significant depressive symptoms between 15% and 25% of community-dwelling older individuals, while Steffens, Fisher, Langa, Potter, Plassman (2009) locate it around 11%. Depression or depressive symptoms in later life are multi-determined and diverse in presentation, however, these are associated with significant cognitive and functional impairment (e.g. Jajodia and Borders, 2011; Bunce, Batterhamb, Mackinnon, Christensen, 2012), social isolation (e.g. Schwarzbach and Luppa, et al., 2013), excess disability related to physical co-morbidities and increased use of health services (e.g. Peytremann-Bridevaux, Voellinger, Santos-Eggimann, 2008; Ladin, 2012), poor adherence to treatment (Gallagher, Savva, Kenny, Lawlor, 2013), and mortality (e.g. Katz, 1996; Zhang, Kahana, Kahana, Hu, Pozuelo, 2009). Older men present three times higher risk of completed suicides or passive death wishes compared any other age group, which are often related to major depression (Conwell, Y., Duberstein, P. R., Caine, 2002; Conwell, 2009, Ayalom, 2011).

The choice of depression as the outcome variable focused on this meta-analysis was motivated by two concerns. Firstly, as described above, depression is a common, debilitating and risk-laden presentation in older people. The personal, social and economic cost of depression is very significant (e.g. Kleine-Budde and Muller, et al., 2013), and the possibility of its prevention or reduction of its burden for the individual is, therefore, also of clinical significance (Campion, Bhui, Bhugra, 2012). Past research has characterised retirement as an increasingly less linear, potentially delayed and more variable process for individuals. Contemporary retirement may be, potentially, a disruptive transition in later life, as a prolonged period of loss, burden or stress (Fonseca, 2007). Alternatively, the move to or anticipation of retirement may be a protective factor for older individuals' mental health, as period of renewal and investment in meaningful aspects of life. The current political and social trend towards the extension of working lives has again raised interest and questions around these issues (e.g. Westerlund et al., 2010; Behncke, 2012; Sahlgren, 2013). A clearer uunderstanding of the relationship between retirement and the experience of depression or depressive symptoms, may point us in the direction of an important clinical pathway or strategy for early identification or intervention in late life depression (Brugiavini, Croda, Dewey, 2008b). Secondly, methodological concerns also motivated the choice of depression as the outcome focused on this meta-analysis. As addressed above, this large body of literature characterized by noticeable variability in the outcome constructs that are used (van Solinge and Henkens, 2008; Coelho and Newman, 2014). Depression as a construct offered the most consistent operationalization and measurement across primary studies. In addition to this, standard clinical measures of symptoms of psychological distress, like depression, are considered to be valid measures in this population (Kolodziej, 2011).

Finally, the focus on a relationship between retirement and depression or symptoms of depression, without a statement on the causal direction of this potential relationship, is deliberate. This meta-

analysis, therefore, aims to conservatively analyse a *relationship* between retirement and depression or depressive symptoms.

2. Objectives

2.1 Primary objective

The main objective of this meta-analysis was to assess the relationship between retirement from the work force and the experience of depression or depressive symptoms in retired older individuals.

2.2 Secondary objectives

The secondary objective of this meta-analysis was to explore how studies' methodological features and sample characteristics may account for some of the variability found in this relationship (Lipsey and Wilson, 2001; Petticrew and Roberts, 2006; Elwood, 2007; Borenstein, Hedges, Higgins and Rothstein, 2009). Based on existing theoretical and empirical research, the following seven variables were identified for the exploration of heterogeneity of effects between studies.

2.2.1 Study characteristics

2.2.1.1 Overall study design

Various authors (e.g. Wang, 2007; Oksannen et al., 2011) have emphasised the possible influence of study design on the outcomes pertaining to the relationship between retirement and mental health. Therefore, cross-sectional and cohort studies were analysed separately, following the method used by McKee-Ryan et al. (2005) and Paul and Moser (2009).

2.2.1.2 Type of outcome measure

The consideration of the influence of the type of measure of depression on the relationship of interest was two-fold. Firstly, this meta-analysis was interested in exploring contrasts between the use of diagnostic tools and measures of symptom severity in depression. This follows empirical data related to the challenges in the measurement and, consequently, calculation of prevalence of depression in older individuals (Karel and Hirichsen, 2000; Watson, Lewis, Kistler, Amick, Boustani, 2004). Secondly, contrasts between the use of self-report and "objective" indicators (i.e. either clinician-led assessment or use of proxy-indicators, like medication intake) were also of interest.

2.2.1.3 Date of study

Given the evolving nature of retirement, retirement policy and research concerns throughout the last six decades or so (Shultz and Wang, 2011), this meta-analysis was interested in exploring how the outcomes may relate to the dominant social and political context at the time of the study. In an attempt to model a potential impact of the political and social shift from the desirability of retirement (and early retirement) to the reversal of such policies (towards extension of working lives and delay in retirement) since, approximately, 2000 (Leber and Wagner, 2007; Shultz and Wang, 2011), contrasts were planned between studies pre-2005 and post-2005 (to allow the median publication time-lag of 4-6 years [Ioannidis, 1998] for studies carried out in the late 1990s).

2.2.2 Characteristics of the sample

2.2.2.1 Time in retirement

In accordance with the current longitudinal views of retirement, i.e. as a process of transition (e.g. Szinovacz and Davey, 2004; Shultz and Wang, 2011), this meta-analysis considered how the number of years spent in retirement may moderate the relationship between retirement and the experience of depression or depressive symptoms. Indeed, phase models of retirement adaptation, such as Atchley's (1971), include the idea of a 'honeymoon' phase, which if followed by a 'disenchantment' phase and a 'reorientation' phase. This has informed research hypotheses regarding the relationship between time in retirement and the experience of depressive symptoms or psychological wellbeing (e.g. Ekerdt, Bossé, Levkoff, 1985; Gall, Evans, Howard, 1997; Kim and Moen, 2002; Szinovacz and Davey, 2004), with variable results.

2.2.2.2 Mean age

This meta-analysis also considered the moderating effect of age. Age has been consistently related to differences in retirement antecedents and consequences in this research field (Gall and Evans, 2000; Pinquart and Shindler, 2007; Topa et al., 2009). As a predictor of life-satisfaction or adjustment to retirement, age is considered in relation to changes in physical health, social function and mental health (Gall and Evans, 2000). However, despite the consistency in the use of age as a correlate or moderator, its relationship with retirement wellbeing or adjustment is variable and contradictory (Topa et al., 2009).

2.2.2.3 Gender

Retirement, it is argued, has been typically researched as a predominantly male process (Quick and Moen, 1998; Kim and Moen, 2001; Price and Nesteruk, 2010). This is clearly at odds with the gender structure of the workforce in most societies. Kim and Moen (2001) consider that '*gender is a key source of heterogeneity in the nature and impact of retirement (…) women tend to have more negative attitudes toward retirement, plan for it less, adjust to it more poorly, and are more likely to experience*

depression following retirement' (p. 85). Gender was, therefore, focused as a moderator in the relationship of interest in this meta-analysis.

2.2.2.4 Geographical origin of the sample

The relationship between retirement and depressive symptoms may also be moderated by the geographical origin of the research samples. The experience of retirement is contextualised in social, normative and political contexts. Differences in how retirement is experienced, for instance, in North America, Europe and Asia will be, to some extent, related to the particular social institutions (e.g. the availability social security benefits) and local political decisions or pressures (e.g. the encouragement of early retirement). According to Hershey, Henkens, van Dalen (2007) and Topa et al. (2009) the origin of the sample '*may be why empirical research has detected notable discrepancies among the results of studies from diverse countries, so the origin of the participants can be expected to be a potential moderator of variability among the studies*' (Topa et al., 2009:42).

3. Methods

3.1 Criteria for considering studies for this review

3.1.1 Types of studies

This meta-analysis considered all quantitative, empirical primary studies (i.e. not theoretical or review studies) which included the examination of the relationship between retirement and at least one valid measure of depression in this population (see Table 3.1). Retirement was defined as the state that follows the normative, permanent or temporary exit from the work force in late adulthood, thus distinguished from job loss or unemployment (Moen, 1996; Marshall and Taylor, 2005). A valid measure of depression was defined as an instrument (e.g. scale or diagnostic tool), which allows the characterisation or quantification of a) mental states such as depression or depressed mood, or b) severity of symptoms of depression. The definition of these terms was based on a preliminary examination of a sample of studies.

This meta-analysis did not use inclusion/exclusion criteria related to type of research design in the initial search and selection of studies. The advantages (breadth and representativeness of the resulting sample) and disadvantages (increased risk of bias) of this decision have been extensively discussed

(e.g. Lipsey and Wilson, 2001; Higgins, Ramsay, et al., 2013). As recommended by Higgins et al. (2013), this meta-analysis it did not establish, at the outset, absolute thresholds of quality eligibility for studies, and did not rely on pre-set hierarchical paradigms related to quality of study designs (e.g. Oxford CEBM Levels of Evidence, Howick, Chalmers, Glasziou, et al., 2011). Inclusion was decided on the study's fitness for the purpose of the meta-analysis (Tugwell, Petticrew, Kristjansson, et al., 2010).

The preliminary literature review highlighted that there were no available randomised controlled trials (henceforth RCTs) in this area. In the absence of RCTs, the present review focused on the available studies in this area, which fell into the broad category of non-randomized study (henceforth NRS), and more specifically *observational studies* (Reeves, Higgins, Ramsay, Shea, Tugwell, Wells, 2013). Stroup, Berlin, et al. (2000) define this category as: '*an etiologic or effectiveness study using data from an existing database, a cross-sectional study, a case series, a case-control design, a design with historical controls, or a cohort design'* (p. 2008).

The validity of meta-analytic procedures in relation to these studies is a controversial issue (Valentine and Thompson, 2013). It is also the case that most data review and synthesis protocols and internal validity (risk of bias) assessment tools tend to be designed with RCTs in mind (Wells, Shea, et a., 2013; Valentine and Thomson, 2013). On the one hand, some authors consider NRSs to be valuable in the study of areas of human experience (exposures) that are very unlikely to be addressed competently in RCTs (Stroup et al., 2000; Elwood, 2007; Shünemann, Tugwel, et al., 2013). Retirement is one of such exposures. On the other hand, other authors consider these studies to be methodologically prone to various biases, challenging the valid development of review or meta-analytic procedures and conclusions (Reeves et al., 2011; Reeves et al., 2013). Valentine and Thomson (2013) caution that meta-analyses of NRSs '*cannot adjust for the effects of unmeasured or unmeasurable factors (...) and the degree of resulting bias from residual confounding may still be large and unpredictable in direction*' (p. 30). Nonetheless, the number of meta-analyses of NRSs has increased substantially and feature heavily in Cochrane Reviews (Stroup et al., 2000).

An aspect of NRSs that presents specific challenges to a meta-analytic process is the inherent variability of study designs in this category (Higgins et al., 2013). In what concerns the identification and selection of studies, Von Elm, Altman, et al., (2007), Elwood (2007), Reeves et al. (2011), Hartling, Bond, Santaguida, Viswanathan, Dryden (2011), Higgins et al. (2013), amongst others, advise that the assignment of study design labels in primary studies is too misleading to be relied upon. As such, this meta-analysis chose to not use explicit study design labels as inclusion/exclusion criteria. It included all studies, which, overall, fell into two commonly used labels (used here just as an available short-hand), defined by Reeves et al. (2011) and Higgins et al. (2013:15) as cohort studies, (prospective or retrospective) and cross-sectional studies. In relation to both types of study, one inclusion design feature was later explicitly applied. Only cohort studies that included a comparison

between pre-retirement and post-retirement measurements of participant's in relation to the index variable were retained; in the same way, only cross-sectional studies that featured comparisons between retired and non-retired groups of participants were retained.

The period of inclusion of studies was set from 1950 onwards, a date from which there seems to be more significant research on retirement (Shultz and Wang, 2011). Selective language limits were also imposed. The following languages were considered: English, Spanish, French, Portuguese and Italian (based on the author's available language resources), which allowed access to evidence beyond that written in English (Moher, Pham, Lawson, Klassen, 2003). Published and unpublished studies were eligible.

3.1.2 Types of participants

Participants were defined as individuals of either gender, who exited or who were about to exit the labour force, i.e. retiring from working for profit or pay in formal labour (Joyce, Pabayo, Critchley, Bambra, 2010). Studies were considered if the age of participants was over 40 years, which represents the lower limit used for the estimation of the effective age of retirement, as defined by the OECD's latest global statistics (OECD, 2011)[7]. There are therefore '*differences in the age at which people were estimated to actually retire, as opposed to the age at which they become pension eligible*' (OECD, 2011:60). Because of the weak relationship between chronological age and retired status, studies that inferred/assumed the labour participation status of participants based on their age, were considered too unsafe to be included (Warr, Butcher, Robertson, Callinan, 2004).

This meta-analysis focused on studies of retirement as a normative experience (as much as possible), in the general population. Therefore, it did not focus on studies of specific populations which a) inherently introduce particular characteristics or needs to the process of retirement, and b) if included in the sample of studies, would influence the outcomes of the meta-analysis significantly (Vahtera, Westerlund, et al., 2009; Westerlund, Kivimaki, et al., 2009). These specific populations were: individuals with learning disabilities; individuals with pre-retirement existing chronic or acute physical conditions (e.g. heart disease, stroke, diabetes, arthritis, cancer, COPD); individuals with pre-existing chronic or acute mental illnesses/conditions (e.g. psychosis, personality disorder, cognitive impairment); individuals who retire early to take up illness/disability-related retirement; individuals who retire very early as part of group, or specifically incentivized employer schemes; individuals who served professionally in the armed forces or police services; individuals who had been professional or elite sports men and women.

[7] '*The average effective age of retirement is derived from observed participation rate changes over a five-year period for successive cohorts of workers (by five-year age groups) aged 40 and over*' (OECD, 2011:60).

3.1.3 Types of outcome measures

This meta-analysis included studies that provided at least one valid and specific outcome measure of depression, and reported on the magnitude and direction of the effect of retirement on this measure *or* on the association between retirement and this measure. Only studies from which an effect size could be computed were eligible (Lipsey and Wilson, 2001). The types of measures initially considered for inclusion were: self-report tools (e.g. Centre for Epidemiological Studies of Depression Scale [CES-D, Radloff, 1977].), clinician-rated tools (e.g. Hamilton Depression Rating Scale [Hamilton, 1967]), structured diagnostic tools (e.g. the Composite International Diagnostic Interview [CIDI, World Health Organisation, 1990]), pre-exiting diagnoses or proxy-indicators (e.g. use of anti-depressant medication). Constructs such as depression, anxiety, mood, psychological wellbeing, mental health were initially considered. However, depression was specifically chosen as the outcome construct here because: a) it is the less ambiguously defined construct present in this body of literature (compared to constructs such as adjustment, mental health or psychological wellbeing)[8]; b) the measures that are used can be readily identified; and c) these measures have direct clinical relevance and familiarity, but are also routinely used in non-clinical settings and populations.

Table 3.1 – Summary of inclusion and exclusion criteria for studies

Inclusion criteria
3. Study reports on quantitative data (no research design limits).
4. Study reports on the quantification of the relationship between the state of retirement and at least one measure of depression.
5. In the case of cohort studies, study reports on comparisons between pre-retirement and post-retirement measurements of the index variable. In the case of cross-sectional studies, study reports on comparisons between retired and non-retired participants.
6. Study includes an appropriate definition of retirement or retired status, and is thus distinguished from late-life job loss or unemployment.
7. Study includes an appropriate and specific definition of depression as a dependent variable or outcome variable.
8. Study provides at least one valid measure of depression in this population, including: self-report, clinician-rated or structured diagnostic tools; pre-exiting diagnoses; or proxy-indicators of mental health state.
9. Study includes an expression of the magnitude and direction of estimates of the effect of retirement on this measure or association of retirement and this measure.
10. Study's participants are defined as: individuals of either gender, who exited or who were about to exit the labour force, and were 40 years old or over.
11. Retired status is explicitly assessed and determined in the study's methodology, i.e. not assumed based on age only.
12. Study's outcomes allow an effect size to be computed.
13. Study was carried out form 1950 to the present.
14. Study is written in English, Spanish, French, Portuguese or Italian.
15. Study is published or unpublished.
Exclusion Criteria
1. Study reports on qualitative data, review data, or data of unidentifiable origin.
2. All the study's participants are selected from the specific populations (in relation to their retirement process and experiences) listed above.

3.2 Search methods for identification of studies

The use of NRSs in meta-analyses poses added difficulties to the identification of studies, compared

[8] See Coelho and Newman (2014) for a recent review of this issue.

to RCTs. These difficulties are related to the fact that these studies are more variable in type, design and methodologies. There is also the added challenge of the absence of a central register of non-randomised studies (Higgins et al., 2013; Reeves et al., 2011). Overall, NRSs are considered to be exposed to increased risk of various types of publication bias, namely selective reporting of outcomes and analysis (Sterne, Egger, Moher, 2011; Norris, Moher, et al. 2013). The implications of these issues for meta-analyses are: a) greater difficulty in estimating the extent of the population of studies about a specific review question, i.e. estimate the 'file-drawer' (Rosenthal, 1979); b) lesser known magnitude of the influence and determinants of publication bias for these studies; c) greater potential for misleading or invalid conclusions (Sterne, et al., 2011; Higgins, et al., 2013).

These challenges were taken into account in the design of the search strategy. It prompted the compilation of an inclusive list of electronic databases, specific publication sources and grey literature. Designed to complement the specific inclusion/exclusion criteria above, the search strategy therefore emphasized sensitivity over precision (Elwood, 2007; Golder and Locke, 2009; Lefebvre, Manheimer, Glanville, 2011). This trade-off was significant as, while prioritising sensitivity increased confidence in the outcomes of the search, the loss of precision demanded a significant investment of time at the search stage, and the development of a pragmatic approach for dealing with large numbers of records (Golder and Locke, 2009). A funnelling approach in the selection of relevant studies was therefore used (Higgins and Deeks, 2011). All search terms were used in English and, when relevant, translated into the other languages listed above. The terms included: retirement, work-cessation, bridge-employment, employment; mental health, mental illness, mental ill-health, distress, wellbeing, psychological, affective, depression, anxiety; self-esteem; retired, retirees, old/older, elder. These were, where possible, combined as: exposure, outcome, population. The literature search was conducted between December 2012 and March 2013 (full search strategy in Appendix 1).

3.2.1 Electronic searches

Data was searched in electronic databases from various disciplines, including: CENTRAL, The Campbell Collaboration Library, MEDLINE, EMBASE, CINAHL, PsycINFO, ASSIA, Social Sciences Citation Index, Sociological Abstracts, Global Health, CAB, IBSS, ABI/INFORM Complete, Business Source Complete, Bibliography of Asian Studies Online, The European Library, US Library of Congress Online Catalogue (see Appendix 1).

3.2.2 Searching other resources

3.2.2.1 Grey literature

Although the category of grey literature is somewhat ambiguous (Higgins and Deeks, 2011), it is the source of around 10% of studies referenced in Cochrane reviews (Mallett 2002, cited in Higgins and Deeks, 2011). The grey literature search in this meta-analysis was wide-ranging, including: databases

of conference proceedings, databases of theses and dissertations, labour organisations or relevant institutions (governmental, academic, charitable bodies) (see Appendix 1).

3.2.2.2 Hand-searching

Individual journals (digital format), from various research areas, which regularly publish on retirement were also consulted, including: The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, Gerontologist, Aging and Mental Health, British Journal of Psychiatry, Epidemiology, Journal of Epidemiology and Community Health, The Lancet, Journal of Vocational Behaviour, Journal of Occupational Health Psychology. Reference lists of relevant publications (primary studies, reviews, institutional, reports, guidance or policy documents) were also searched.

In addition to the traditional literature sources above, this meta-analysis also searched for individual past or ongoing studies, host institutions of studies on ageing and/or retirement, national and international. The full list of these studies and institutions is presented in Appendix 1.

3.2.2.3 Correspondence

Finally, individual authors were contacted for unpublished studies and additional information on published studies via email, as recommended by Young and Hopewell (2011) (see Appendix 2).

3.3 Data collection and analysis

3.3.1 Selection of studies and unit of analysis

Studies were selected by the main author only. This process was developed in two steps, following the methodology used, for instance, in Siegfried, Muller, Deeks, et al. (2005), Topa et al. (2009) and Luhmann, et al. (2012). In the first step, the thirteen inclusion criteria and two exclusion criteria were applied systematically to the results of the data search, using a funnelling approach over successive sifts of data (Figure 3.1). Records were appraised for inclusion, first, by screening of title and abstracts (Sifts 1 to 4) and, second, by reading the full-text (Sift 5 and 6). After this, a second step was used, taking into account a characteristic of this research area, i.e. the existence of large epidemiological studies of ageing or social surveys (e.g. GAZEL cohort in France, or RAND Survey Meta Data Repository), which generate multiple waves of data, widely accessible to researchers for use in secondary analyses. These are, in turn, reported in multiple individual articles, reports or abstracts. For this meta-analysis the unit of interest was the study, not the report (Lipsey and Wilson, 2001; Higgins and Deeks, 2011). As it was expected that the data search would yield large numbers of reports, this required that multiple reports of the same study/data were linked together (Higgins and Deeks, 2011). Therefore, this second step was taken, in which reports were checked for duplicate analyses of the same data, and linked to a "parent study". At this stage, only one report per "parent study" was

retained. Priority was given to publications reporting: a) more time points (in the case of cohort studies), b) larger sample sizes, c) clearer descriptive and/or inferential statistics (Luhmann et al., 2012), and d) peer-reviewed reports, over non-peer-reviewed published reports, over non-published reports. Reports or studies that did not meet criteria but were, nevertheless, relevant, were used as background information (Fonner, Denison, Kennedy, O'Reilly, Sweat, 2012). The sequence of data selection process, sifts and outcomes is presented in Figure 3.1.

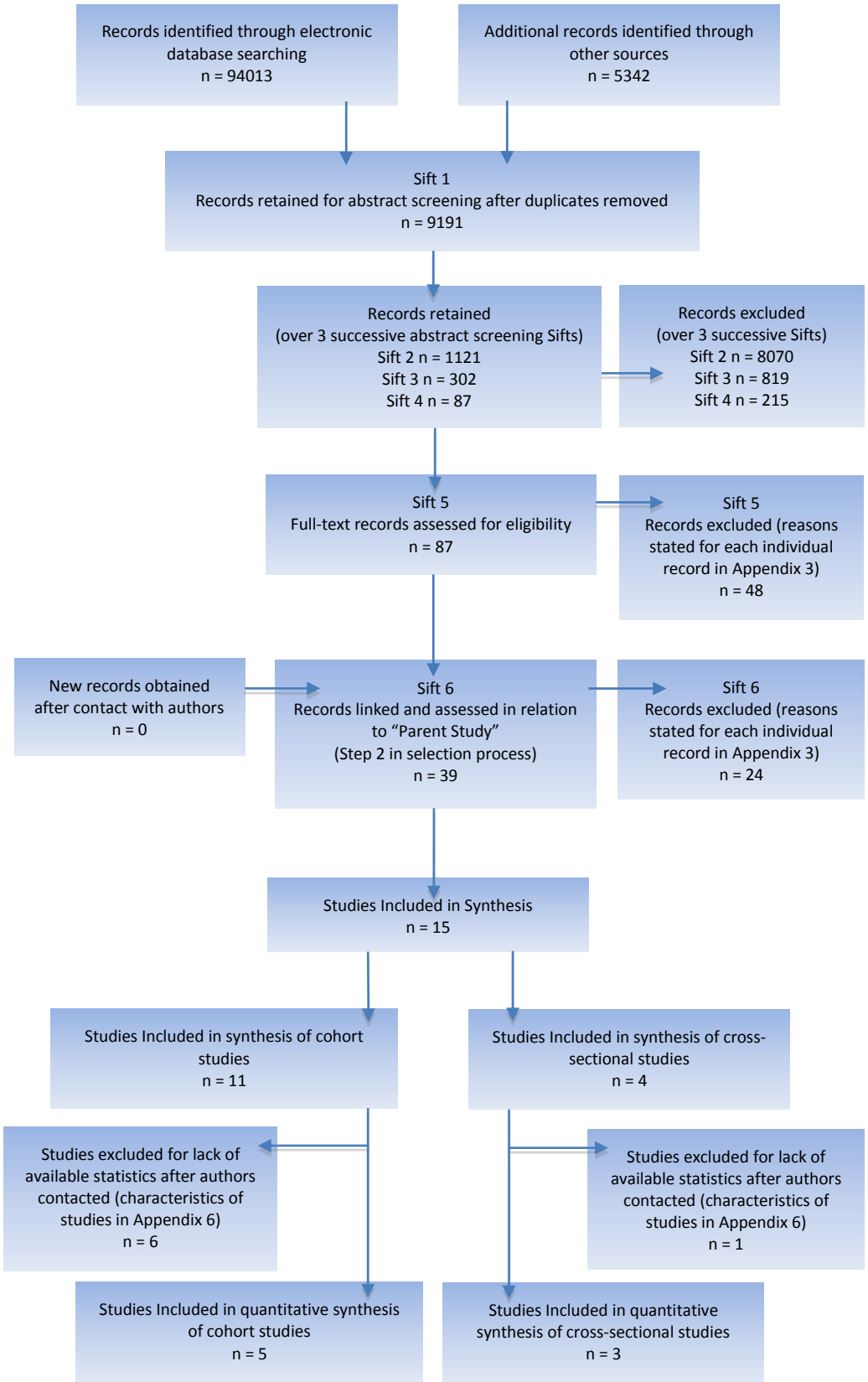
3.3.2 Data extraction and management

The search and selection strategies yielded primary studies designs of two kinds: cohort (related samples/repeated measures) and cross-sectional (unrelated samples/independent groups). From this point onwards, these were treated separately (Lipsey and Wilson, 2001; Morris and DeShon, 2002). The decision to separate the two types of data also follows the methodology used by Siegfried, Muller, Volmink, et al. (2003), McKee-Ryan, Song, Wanberg, Kinicki (2005) and Paul and Moser (2009).

Data extraction and coding was initially done by the first author, using two purpose-built coding manuals (in Appendix 4a/4b). Before the final exclusion of studies (due to lack of usable statistical data) 3 cross-sectional studies and 5 cohort studies were randomly selected[9] and coded by the second author. The average percentage of agreement across all of the coded categories was 82% for the cross-sectional sample and 83% for the cohort sample. Disagreements were resolved by discussion, and resulted in the review of the coding manual and data. Data were then collated and managed in Excel.

[9] Obtained with http://www.random.org/

Figure 3.1 – Flow-chart of study identification, screening, selection and inclusion (Moher, Liberati, Tetzlaff, Altman, 2009)



3.3.3 Assessment of risk of bias in included studies

3.3.3.1 Risk of Bias and Study Quality

Risk of bias is understood as *'risk of systematic error'* (Higgins, Altman, Sterne, 2011), which can introduce distortions of various kinds, magnitude and directions, and can mask or create an apparent effect (Higgins et al., 2011). Bias is distinguished here from imprecision (random error) or quality (the extent to which a study was conducted to the highest possible standard [Higgins et al., 2011]). Following Higgins et al. (2011) and Higgins et al. (2013), an assessment of risk of bias was preferred to an assessment of the quality of the study, as: a) it addresses directly the '*extent to which results of included studies should be believed*' (Higgins et al., 2011); and b) disentangles an overall judgement of quality from the assessment of the influence of specific biases (i.e. a study may be performed to the highest possible standard, yet, still have an important risk of bias [Higgins et al., 2011]).

The proliferation of quality assessment tools (scales and checklists) in recent years has been well documented and reviewed (e.g. Sanderson, Tatt, Higgins, 2007; Bai, Shukla, Bak, Wells, 2012). This effort has also been accompanied by the development of guidelines for the reporting of observational studies (e.g. Tooth, Ware, Bain, Purdie, Dobson, 2005; von Elm et al., 2007). These reviews point out two important issues. First, the design of tools fit for use with NRSs has received much less attention than the design of tools for RCTs. This leads to the use of tools designed with RCTs in mind to evaluate NRSs, which then systematically fail to satisfy quality assessment characteristics (like random allocation to groups, or blinding) that are not meaningful in this context. Second, as there is a lack of consensus over which elements are critical in assessing vulnerability to bias in NRSs (Sanderson, et al., 2007), and given the broad range of designs included in this category, one assessment tool '*is unlikely to adequately address concerns across all these designs*' (Higgins, at al 2013:20). Therefore, this meta-analysis opted to use a specifically designed risk of bias checklist, rather than a general ready-made quality tool, adjusted to the specific characteristics of the studies under review.

3.3.3.2 Specific aspects of risk of bias in NRSs

NRSs raise specific issues in relation to the assessment of risk of bias. These differ from RCTs crucially in relation to their exposure to risk of selection bias, i.e. '*differences in the baseline characteristics of individuals in different [exposure] groups*' (Higgins et al., 2013:13). In NRSs the distortions introduced in the allocation of individuals to groups are unknown. Confounding occurs when characteristics are unevenly distributed between groups and create differences in the outcome under investigation. According to Higgins et al. (2013) confounding produces two possible effects for a meta-analysis: '*If confounding produces biases in one direction, then the overall estimate of the intervention effect will be shifted (systematic bias). If biases vary across studies, then this will lead to increased variability of the observed effects, introducing excessive heterogeneity among the studies, and the potential for true effects to be missed*' (p. 13). Wells, et al. (2013) and Norris et al. (2013)

consider that, at this point, the assessment of risk of bias in NRSs should focus on the following domains: a) strengths and weaknesses of the design; b) details of the study's execution; c) the potential for selection bias and confounding; and d) the potential for reporting biases, namely, selective reporting.

3.3.3.3 Risk of Bias assessment

The Cochrane Collaboration Risk of Bias Tool (CCROB Tool, Higgins et al., 2011) was developed for the assessment of RCTs[10]. The use of this tool with NRSs is a subject of disagreement. The CCROB Tool was tested on a sample of 3 studies included in this meta-analysis (across cohort and cross-sectional studies) by the third author (experienced in the use of this tool). The conclusion was that this tool did not fit the needs of the studies in this meta-analysis. It was therefore necessary to develop a tailored risk of bias assessment tool, following, as indicated by Higgins et al. (2013), the approach and framework used by Siegfried et al. (2003), supplemented by the reporting bias checklists suggested by Wells et al. (2013) (in Appendix 5). The assessment of risk of bias was done by the main author, with a randomly selected sample of 6 studies (across all studies) independently assessed by the third author. The reviewers were not blinded to the names of the authors, institutions, or journal of publication of the studies.

Formal statistical methods for combining results of studies at different risk of bias are not yet currently available (Higgins et al., 2011). Therefore, the following procedure was used. First, a narrative description of risk of bias was produced for each study, in relation to each of the domains listed above (included in Table of Characteristics of Included Studies, Appendix 6). Second, in case risk of bias estimates varied critically between studies, this information was included in further sensitivity analyses of effect estimates (Higgins et al., 2011).

3.3.4 Measures of the effect

Morris and DeShon (2009) advise that, when meta-analyses include studies based on both independent-groups and repeated-measures, combining outcomes across design-types should be avoided or done with extreme care. Borenstein, et al. (2009), are less prescriptive in this sense. They advise that combining data should be considered in relation to the studies under review and the review question itself. For this meta-analysis, it was assumed that studies with different design-types were functionally different. The two effect size metrics for each design-type are, therefore, described below (effect size calculation formulas and further effect-size estimation procedures in Appendix 7).

[10] A project to extend the existing risk of bias tool to NRSs has recently been funded by The Cochrane Collaboration, and is still in its early stages of development (Higgins et al., 2013).

3.3.4.1 Cross-sectional studies

These studies typically focus on the difference between two independent groups (here defined by work participation status), at a single point of measurement in time. For cross-sectional studies, the standardized mean difference (SMD) was chosen as the effect size metric (commonly known as *d*), in order to allow the meaningful comparison between the various measures of depression/depressive symptoms used across studies (Lipsey and Wilson, 2001). In addition, the raw mean difference effect size (*D*) was also calculated for each cross-sectional study.

3.3.4.2 Cohort studies

The standardized mean gain was initially considered as the effect size statistic for this sub-group of included studies (Lipsey and Wilson, 2001). However, Morris and DeShon (2002) and Borenstein et al. (2009) suggest the use of the standardised mean difference (SMD) in the case of repeated measures, such as these. In comparison to the standardised mean gain, the standardized mean difference tends to yield a more conservative estimate of effect size (Morris and DeShon, 2002). Also, in situations where *r* (the correlation between pairs of observations) cannot be assumed to be homogenous between studies, the standardised mean difference is also preferable (Morris and DeShon, 2002). Methodological concerns regarding this set of studies, guided the choice of the standardised mean difference as a more conservative metric for cohort studies. Finally, in addition to this, as above, the raw mean difference effect size was also calculated for each cohort study.

3.3.5 Dealing with missing data

For all studies, at the coding stage, additional information was sought from other sources (e.g. parent-study websites; papers related to the same study or data), if the information was ambiguous or absent from the paper included in the analysis. The level of confidence in the information used was also coded (see Coding Manuals Appendix 4a/4b). In relation to missing effect-size level data, authors were contacted in order to obtain missing information. This was successful in only a small number of cases – only 3 out of 11 contacted authors replied with requested additional data/clarification (Appendix 2). If this information could not be obtained in this way, missing terms necessary for the calculation of the relevant effect-size statistic were, if possible, estimated according to procedures described in Lipsey and Wilson (2001) and Borenstein et al. (2009). The level of confidence and estimation of the effect size was also coded (see Coding Manuals and Forms) (Lipsey and Wilson, 2001). Finally, if information could not be obtained by any of the strategies above, it was coded as missing.

3.3.6 Data synthesis

Given the characteristics of this research area, a random-effects model was chosen as the statistical model in this meta-analysis. This is based on the expectation that true effects[11] will vary from study to study, reflecting systematic differences in how studies from different research areas explored the relationship between retirement and depression. Because of these differences, the possibility that there may be different effect sizes underlying different studies is considered in this model (Borenstein, et al., 2009). Overall effect size data was interpreted according to Cohen's (1992) criteria (small [0.2], medium [0.5] and large [0.8]), and all analyses were performed using the Comprehensive Meta-Analysis (CMA) software[12].

3.3.7 Assessment of heterogeneity

Heterogeneity is understood as a concept close to variance, in the context of meta-analyses, and used in reference to true effects. One of the planned aims of this meta-analysis was to characterise the variation in observed effect sizes. However, this observed variation incorporates both true heterogeneity and random error/chance (Borenstein, et al., 2009). The analysis of heterogeneity seeks to tease apart these two components. Borenstein, et al. (2009) advise that a meaningful presentation of heterogeneity indices should include a) a measure of the magnitude of heterogeneity and, b) *'a measure of uncertainty over whether the apparent heterogeneity is genuine'* (p. 122). In this meta-analysis, the I^2 statistic and the *Q* statistic were chosen, respectively, to fulfil these functions.

I^2 is defined as an index of inconsistency, that is, the *'ratio of true heterogeneity to the total variation in observed effects, akin to a signal to noise ratio'* (Borenstein, et al., 2009:120). This is not sensitive to either effect size metric or number of studies included, and its value falls between 0 (all observed variance is due to chance) and 100% (all variance is real, and reasons for this should be investigated). Higgins and Thompson (2002) and Higgins, Thompson, Deeks, Altman (2003) suggest guidance benchmarks for I^2 heterogeneity values: low (25%), moderate (50%) and high (75%). The *Q* statistic is a standardised measure of total dispersion. In addition to *Q*, the calculation of *Q*-*df* reflects excess dispersion, where *df* (degrees of freedom, *K*-1, *K* is the number of studies) is the expected value of *Q* on the assumption that all studies share a common effect size. The *Q* statistic (and its *p*-value) was then used to test the assumption of homogeneity of true effect sizes (in a regular test of the null hypothesis). Finally, Tau (T) was also used to reflect an estimation of the standard deviation of true effect sizes, and Tau-squared (T^2) an estimation of the variance of the true effect sizes (Borenstein, et al., 2009).

---

[11] '*A true effect size is the effect size in the underlying population, and the effect size we would observe if the study had an infinitely large sample size (and therefore no sampling error). A study's observed effect size is the effect size that is actually observed'* (Borenstein et al., 2009).

[12] Borenstein et al. (www.Meta-Analysis.com)

Valentine and Thompson (2013) argue that greater heterogeneity is more likely when a meta-analysis includes NRSs. Thus, in addition to the use of a random-effects statistical model, it will also focus on the exploration of study and sample characteristics outlined in the objectives as potential sources of heterogeneity (Petticrew and Roberts, 2006), in sub-group and sensitivity analyses.

3.3.8 Planned sub-group and sensitivity analyses

These include the exploration of the effect of the following (theoretically-based) methodological and sample characteristics on the variability of effect sizes: overall study design (planned comparisons between cohort studies and cross-sectional studies sub-groups); type and quality of instrument used to measure depression/symptoms of depression (planned comparisons between a) diagnostic and symptom-severity instruments, and, b) self-report and 'objective/independent' measures); date of study (planned comparisons between pre-2005 and post-2005 studies); samples' mean time in retirement; gender distribution of samples; samples' mean chronological age; geographical origin of the sample.

3.3.9 Assessment of publication biases

Publication biases are defined as the outcome of the '*dissemination of research findings being influenced by the nature and direction of results*' (Sterne, et al., 2011). Reporting biases were assessed here using both funnel-plots and '*fail-safe N'* (Rosenthal, 1979). This is a widely used measure of reporting bias, although its limitations (reliance on significance/*p*-value and tendency towards over-estimation of the threshold) have been outlined, among others, by Borenstein, et al. (2009) and Sterne, et al. (2011).

4. Results

4.1 Description of studies

4.1.1 Results of the search and selection process

Figure 3.1 outlines the search process and outcomes in detail. As already mentioned, the search strategy adopted in this meta-analysis (in Appendix 1) prioritized sensitivity over precision for specific methodological reasons, yielding initially a very large number of records. This was managed by a pragmatic sequential approach of selection 'sifts', which included screening of titles and abstracts, and, if necessary, full-text screening. The final stage of selection put forward 87 records for full-text reading. At this point (Sift 5 in Figure 3.1), 48 records were excluded (with reasons for each study stated in Appendix 3) and 39 records retained. After this, the second step in the selection process described before linked all records to "parent studies", allowing the selection of one record per "parent study" (using specific criteria also outlined above), as the unit of interest in this meta-analysis was the study not the record. At this point, 24 records were excluded (reasons again in Appendix 3) and 15 retained: 11 cohort studies and 4 cross-sectional studies. All 15 studies were data-coded (with a random selection of 8 studies also second-coded). However, as gaps in statistical data for the calculation of effect-sizes were identified, 11 authors were contacted with requests for additional statistical information at this point. At the end of the search and selection process 6 cohort studies and 1 cross-sectional study were excluded because of lack of usable or complete statistical data for the calculation of effect-sizes. The resulting 5 cohort studies and 3 cross-sectional studies were therefore put forward for analysis.

4.1.2 Included studies

A summary of key characteristics of each study included in this meta-analysis is presented in Tables 4.1 and 4.2. An extensive table of Characteristics of Included Studies is found in Appendix 6. All studies, with the exception of Steptoe, et al. (2012), were published in peer-reviewed journals. Also, most studies were secondary analyses of existing data (i.e. data collected for other or multiple purposes, previous to the planning and execution of the analysis), with the exception of Reitzes, et al. (1996) and Steptoe, et al. (2012).

The summary of outcomes of cohort studies emphasises the inconsistency found in relation to the effect of retirement on self-reported depressive symptoms or association between retirement and self-reported depressive symptoms. All studies used considerably large samples, increasing their statistical power and the likelihood of reporting statistically significant outcomes. Kubicek et al. (2011), Clarke

et al. (2011) and Calvo et al. (2013) all reported considerable loss of data between the two measurement time points (45%, 50.6% and 21.3%, respectively). Three studies (Reitzes et al., 1996; Westerlund et al., 2010; Kubicek et al., 2011) reported, an association between retirement and reduction in depressive symptoms or a reducing effect of retirement on depressive symptoms. One study (Clarke et al., 2011) reported no association between retirement and increase in depressive symptoms. One study (Calvo et al., 2011) reported an increasing effect of retirement on depressive symptoms, although this effect weakened if retirement occurred after the age of 60. All cohort studies used self-report measures, namely, variations of the CES-D scale.

The summary of outcomes of cross-sectional studies on differences between retirees and workers in self-reported symptoms of retirement is equally inconsistent. The same large samples are found, and all 3 studies used self-report measures. With the exception of Bosse et al. (1987) which used the SCL-90-R depression subscale, the remaining two use versions of the CES-D. Two studies (Bosse et al., 1987; Steptoe et al., 2012) reported on, respectively, higher or more prevalent symptoms of depression in retirees. One study, (Drentea, 2002) reported inconsistent outcomes, suggesting both no association between retirement and depressive symptoms and a positive effect or association of retirement on/with self-reported depressive symptoms.

Table 4.1 – Summary characteristics of included cohort studies

	Design						*Sample/Sampling*						*Measurement*		*Outcomes*
Study ID	*Study design*	*Time points*	*Interval*	*2^{dary} analysis*	*Analysis design*	*Discipline*	*Origin*	*Method*	*Total N T1*	*Data loss %*	*Male % (T1)*	*Age (M) T1*	*Measure*	*Other DVs*	*Summary of outcomes re. depression*
CO4 Reitzes et al. (1996)	Cohort, prospective	2	1992-1994	No	Multiple Regression	Sociology	USA	Random Non-general population	826	11	48	No data	CES-D, 20 full scale	. Self-esteem	1. Retirement is associated with decrease in depressive symptoms
CO6 Kubicek et al. (2011)	Cohort, prospective	2	1993-2004	Yes	Multiple regression	Psychology	USA	Stratified sample (in 2^{ary} analysis) Non-general population	2899	44.5	No data	53	CES-D, 20, full scale	. Psychological well-being	1. Retirement has a positive effect (reducing) on depressive symptoms
CO8 Westerlund et al. (2010)	Cohort, prospective	3	1996-2005	Yes	Multiple regression	Psychology	France	Convenient, self-selected Non-general population	14104	0	80	No data	CES-D, 20 items, full scale	. Mental fatigue . Physical fatigue . Chronic conditions	1. Significant effect of retirement in reduction of depressive symptoms
CO9 Clarke et al. (2011)	Cohort, prospective	2	1989-2001	Yes	Growth curve modeling	Sociology	USA	Random multistage, stratified General population	3617	50.6	37.5	47.05	CES-D, 7 items	. Physical health . Functional health	1. No association between retirement and increase in depressive symptoms
CO10 Calvo et al. (2013)	Cohort, prospective	2	1992-2010	Yes	Multiple Regression	Sociology	USA	Random multistage, stratified General population	6275	21.31	53.69	55.67	CES-D, 8 items	. Self-reported health	1. Retirement increases depressive symptoms at age 60 2. Retirement has less of a negative impact if it occurs later in life

Key: Time Points = number of measurement time points used in the analysis; Interval = interval between measurement time points ; Secondary Analysis = is the record a secondary analysis of previously collected data?; Analysis design = the design of the statistical analysis performed on data; Discipline = scientific area of the study's first author or affiliated department/institution; Origin = geographical origin of the sample; Method = sampling method and target population (general or non-general); T1 = Measurement time point 1; Data loss = percentage of data lost between T1 and T2 (measurement time point 2); Measure = Measure of depression used; Other DVs = other outcome variables included in the analysis.

Table 4.2 – Summary characteristics of included cross-sectional studies

	Design					*Samples/Sampling*						*Measure*		*Outcomes*
Study ID	*Study design*	*2^{dary} Analysis*	*Analysis design*	*Compare.*	*Discipline*	*Origin*	*Method*	*N*	*Equivalent ?*	*Male %*	*Age (M)*	*Measure*	*Other DVs*	*Summary of outcomes re. depression*
CS1 Bosse et al. (1987)	Cross-sectional	Yes	Cross-sectional; between group differences and multiple regression	Retired vs Workers	Sociology	USA	Convenient Non-general population	N:1513 R: 673 C.:840	Yes Differences found Samples from same population	100	61	SCL-90-R, depress. subscale	. Physical health symptoms . SCL-90 subscales	1. Significantly higher depressive symptoms in retirees; 2. Effect maintained after controlling for severity of health symptoms
CS2 Drentea (2002)	Cross-sectional	Yes	Cross-sectional; multiple regression	Retired vs Non-retired	Sociology	USA	Random General population	N:2587 R: 672 C: 1915	No data Sample sizes differ Samples from same population	43	47.6	CES-D, 7 items	. Distress . Anxiety . Sense of control . Positive affect	1. Difference in means suggests positive effect of retirement on depressive symptoms 2. Regression models show no consistent (positive or negative) association between retirement and depressive symptoms
CS4 Steptoe et al. (2012)	Cohort prospective	No	Cross-sectional; between group differences	Retired vs Workers	Psychology	UK	Random, multistage. General population	N:8414 R:5680 C:2734	No data Samples from same population	44.7	67.76	CES-D, 8 items	. Wealth . Physical health . Health behaviours . Cognitive function . Social particip. . Psych. well-being: affective, eudemonic	1. Higher prevalence of depressive symptoms in retirees, compared to workers.

Key: 2^{adry} Analysis = is the record a secondary analysis of previously collected data?; Analysis design = the design of the statistical analysis performed on data; Compare. = comparison between what groups; Discipline = scientific area of the study's first author or affiliated department/institution; Origin = geographical origin of the sample; Method = sampling method and target population (general or non-general); R = retired sample; C = comparison sample (non-retired or working); Equivalent? = were equivalence tests performed on the two samples; Measure = Measure of depression used; Other DVs = other outcome variables included in the study

4.1.3 Excluded studies

Given the significant number of studies excluded from analysis due to lack of available data, the same summary of key characteristics of each study is presented in Appendix 6 (Tables 3 and 4). These included 6 cohort studies and 1 cross-sectional study. All studies in this set were published in peer-reviewed journals, and all were secondary analyses of data. Tables 3 and 4 (Appendix 6) highlight how, in contrast to the set of included studies, two of the excluded studies (Laaksonen et al., 2012; Oksannen et al., 2011) used the purchase of antidepressant or psychotropic medication as an operationalisation of the effect of retirement on depressive symptoms. Also, in relation to the area of scientific research, medical and public health related studies appear in this set, in contrast to the set of included studies. Finally, European studies dominate this set of excluded studies. With the exception Westerlund et al. (2010) and Steptoe et al. (2012), the majority of included studies used data collected in the USA. However, crucially, the inconsistency in outcomes in the reported relationship or effect of interest in this meta-analysis is again present in these excluded studies. Across all studies: one study (Oksannen et al., 2011) reported a positive effect of retirement on purchased antidepressant medication; three studies (Midanik et al., 1995; Brugiavini et al., 2008; Laaksonen et al., 2012) reported no association between retirement and adopted measure of depression; two studies (Kim and Moen, 2002; Jang et al., 2009) reported an increasing effect or association between retirement and depressive symptoms; finally, one study (Lindeboom et al., 2002), reported very unclear outcomes which do not allow inferences in relation to the direction or existence of an effect without additional information.

4.2 Risk of bias in included studies

Risk of bias was measured by two independent raters (the first and the third authors). The percentage of agreement between the two raters was high (88.45%), suggesting that the custom-made risk of bias assessment tools yielded reliable ratings. Disagreements were resolved in discussion. Tables 4.3 and 4.4 outline the assessment of risk of bias for each study, in relation to each bias domain (further details of studies' characteristics and support for bias assessment judgment in Appendix 6).

In relation to the sub-group of cohort studies, overall, these do not vary so widely that it would compromise the performance of meta-analysis on these data (Higgins et al., 2013). Within this sub-group, Westerlund et al. (2010) and Calvo et al. (2013) stand out as these seem to be at low risk of bias, across all domains. However, while Westerlund et al. (2010) used a convenient sample obtained from a non-general population (individuals recruited from a workplace setting), Calvo et al. (2013) used a sample from the Health and Retirement Study (USA), which is a random, stratified, multistage sample. This implies different levels of risk of confounding, however, their analysis design (multiple regression models) provides some evidence of control of confounding, namely, with both attempting

to control for endogeneity bias (i.e. the impact of existing poor health on the decision to retire and, therefore, on the outcomes of retirement). Conversely, Clarke et al. (2011) stands out, because it suggests high risk of bias in the performance, detection and attrition domains, and high or unknown risk of reporting bias. Particularly, it used a heavily transformed version of the CES-D (reduced to 7-items, as an 'index'). This raises important concerns regarding the overall validity of this measure. Additionally, the paper's account of what data is included in analyses is unclear, raising questions around missing data. Reitzes et al. (1996) and Kubicek et al. (2011) both seem to present a mix of high and low risk of bias, across domains. However, both share the same issue regarding the representativeness of their samples and participation rates, which place these studies at risk of lack of external validity. Finally, with the exception of Reitzes et al. (1996), there was evidence in all studies of methods for dealing with missing data that were not fully pre-specified or accounted for. These studies presented loss of data in their analysis sections, which were simply stated, but not clearly supported/justified.

In relation to the sub-group of cross-sectional studies, overall, these suggest higher risk of bias, across all domains, when compared to the cohort studies. This, again, supports the decision to perform the subsequent meta-analysis in two design-based subgroups (Higgins et al., 2013). Specifically, all three studies raise concerns regarding reporting biases. It is difficult to assess how much selective reporting there was in Bosse et al. (1987) and Steptoe et al. (2012), but Drentea (2002) clearly stated that results were not fully presented in the report, due to lack of statistical significance. Within this sub-group, the risk of bias assessment of Drentea (2002), suggests that this study might be at a higher or unknown risk of bias across all domains, with the exception of attrition bias. It raises particular questions about unclear control of confounding, also using two samples for comparison which seem unbalanced in size (retired n=672 and non-retired n=1915, respectively). Regarding detection bias, as Clarke et al. (2011) above, Drentea (2002) used an idiosyncratic version of the CES-D (selecting 11 items), potentially compromising the scale's measurement properties (conceptual, content and construct validity, and internal consistency). Finally, the assessment of Bosse et al. (1987) raises concerns about a) the representativeness of this sample, exclusively composed of men; and, b) the potential influence of confounding factors in the study's outcomes, as significant differences were found between the retired and comparison/working sub-samples.

Table 4.3 – Risk of bias assessment of included cohort studies

Study ID	External Validity		Internal Validity										
			Performance Bias	*Detection Bias*	*Attrition Bias*		*Selection Bias/control of confounding*						
	1. Represent?	2. Participation rate	3. Retirement status	4. Depression measure	5. Equal follow-up	6. Final analysis %	7. Clear retire/worker status at start?	8. Age	9. Gender	10. Geo. origin	11. Soc.-econ. status	12. Physical health	13. Marital status
CO4 Reitzes et al. (1996)	-	-	-	+	+	+	-	?	+	+	+	-	+
CO6 Kubicek et al. (2011)	-	-	-	+	+	-	+	+	-	+	+	-	+
CO8 Westerlund et al. (2010)	-	+	+	+	+	+	+	?	+	+	+	+	+
CO9 Clarke et al. (2011)	+	-	-	-	-	-	-	+	+	+	+	+	+
CO10 Calvo et al. (2013)	+	+	+	+	+	+	+	+	+	+	+	+	+

Table 4.3 (continued)

Study ID	Internal Validity					Overall risk of bias assessment in each domain (mainly +, - or ?)	
	Reporting Bias						
	14. Pre-specified outcomes?	15. M-A friendly?	16. Partial/unreported outcomes?	17. Multiple/adjusted analyses?	18. Missing data methods?		
CO4 Reitzes et al. (1996)	+	-	+	+	+	External validity Performance bias Detection bias Attrition bias Selection bias/confounding Reporting bias	- - + + + +
CO6 Kubicek et al. (2011)	+	+	+	-	-	External validity Performance bias Detection bias Attrition bias Selection bias/confounding Reporting bias	- - + +/- + +
CO8 Westerlund et al. (2010)	+	-	+	+	-	External validity Performance bias Detection bias Attrition bias	+/- + + +

						Selection bias/confounding	+
						Reporting bias	+
CO9 **Clarke et al. (2011)**	?	-	-	+	-	External validity	+/-
						Performance bias	-
						Detection bias	-
						Attrition bias	-
						Selection bias/confounding	+
						Reporting bias	-
CO10 **Calvo et al. (2013)**	+	+	+	+	-	External validity	+
						Performance bias	+
						Detection bias	+
						Attrition bias	+
						Selection bias/confounding	+
						Reporting bias	+

Key: Risk of bias judgment: + (criterion satisfied), - (criterion not satisfied), ? (cannot tell); Overall risk of bias judgment: mainly +, -, ? in each domain

1. Representativeness of the sample for the target population: study receives a + if the sample included all eligible workers/retirees over a defined period of time, or in a defined catchment area or context/setting (e.g. employer, school graduates), or if sample is a random sample.

2. What percentage of selected individuals agreed to participate (at baseline): study receives a + if the percentage participation was 80% or more of the initially recruited/selected sample.

Performance bias (systematic difference in measurement of exposure)

3. How was retirement status ascertained?: study receives a + if participants' retired status is ascertained via independent record(s) (e.g. employment, medical or social security records); or, if self-assessed retired status is confirmed by independent records.

Detection bias (systematic differences in outcome assessment)

4. How was depression status ascertained?: study receives a + if the measure of depression was adequate (valid) for the population.

Attrition bias (systematic differences in follow-up)

5. All participants followed up for the same amount of time?: study receives a + if all participants were followed-up for the same amount of time.

6. Percentage of participants included in the final analysis (i.e. adequacy of follow-up of cohorts): study receives a + if the percentage participants in the final analysis was 80% or more (of baseline participants); or, if a full description of those lost-to-follow-up was not suggestive of bias.

Selection bias/control of confounding (systematic differences in comparison groups)

7. Clear worker/retired status at start of the study: study receives a + if no participant was retired at the start of the study; or, if those participants that were retired at the start of the study can be clearly identified throughout the analysis.

List of pre-specified potential confounding factors (i.e. characteristics of samples or groups) – for each of the factors listed (8-13), the item receives a + if the factor was balanced between workers and retired groups (10% or less difference) or adjusted for in analysis; an absence of a + indicates that the factor was either not measured, or, if it was, it did not meet the + requirements: 8. Age (at baseline), 9.Gender, 10. Region/country of origin, 11. Socio-economic status, 12. Physical health status, 13. Marital status

Reporting bias (systematic differences in the way outcomes or analyses are reported)

14. All pre-specified primary outcomes are reported: study receives a + if the pre-specified outcome related to depression is reported.

15. The outcome of interest in the review is reported completely so that it is meta-analysis friendly?: study receives a + if the outcome related to depression is reported completely (including descriptive statistics) so that it can enter in a meta-analysis.

16. There is no evidence that an outcome that was partially reported or not reported was statistically insignificant or low in magnitude: study receives a + if there is no evidence that an outcome that was partially reported or not reported was statistically insignificant or related to an effect low in magnitude.

17. There is no evidence that multiple adjusted analyses were carried out but only one or a subset of one was fully reported: study receives a + if there is no evidence that multiple adjusted analyses were carried out but only one or a subset of one was fully reported.

18. There is no evidence of methods being applied to deal with missing data that were not pre-specified: study receives a + if there is no evidence of methods applied to deal with missing data that were not pre-specified.

Table 4.4 – Risk of bias assessment of included cross-sectional studies

Study ID	External Validity		Internal Validity								
			Performance Bias	*Detection Bias*	*Attrition Bias*	*Selection Bias/control of confounding*					
	1. Represent?	2. Participation rate	3. Retired/worker status	4. Depression measure	5. Final analysis %	6. Age	7. Gender	8. Geo. origin	9. Soc.-econ. status	10. Physical health	11. Marital status
CS1 Bosse et al. (1987)	-	+	-	+	+	+	-	+	-	+	-
CO6 Drentea (2002)	+	-	-	-	+	?	?	?	?	-	?
CS4 Steptoe et al. (2012)	+	+	-	+	?	+	+	+	+	+	+

Table 4.4 (continued)

Study ID	Internal Validity						
	Reporting Bias						
	12. Pre-specified outcomes?	13. M-A friendly?	14. Partial/unreported outcomes?	15. Multiple/adjusted analyses?	16. Missing data methods?	**Overall risk of bias assessment in each domain (mainly +, - or ?)**	
CS1 Bosse et al. (1987)	+	-	-	+	-	External validity Performance bias Detection bias Attrition bias Selection bias/confounding Reporting bias	+/- - + + +/- -
CO6 Drentea (2002)	-	-	-	?	?	External validity Performance bias Detection bias Attrition bias Selection bias/confounding Reporting bias	+/- - - + ? -
CS4 Steptoe et al. (2012)	+	-	-	-	?	External validity Performance bias Detection bias Attrition bias Selection bias/confounding Reporting bias	+ - + ? + -

Key: Risk of bias judgment: + (criterion satisfied), - (criterion not satisfied), ? (cannot tell); Overall risk of bias judgment: mainly +, -, ? in each domain
1. to 4. Criteria defined as above; 5. Percentage of participants included in the final analysis: study receives a + if the percentage participants in the final analysis was 80% or more (of baseline participants); or, if a full description of those lost was not suggestive of bias; 6-16 Criteria defined as 8-18 above.

4.3 Effects of methods

4.3.1 Summary effects by design-type sub-group

The inconsistency of observed effects is well represented in Figure 4.1. The relatively narrow confidence intervals around all observed effects suggest that, while these may be fairly precise (due potentially to the use of large samples), the observed dispersion cannot be comfortably accounted for as sampling error. All effects are significant, with the exception of Clarke, et al. (2011). This section will not comment on the overall summary effect, as this is not a meaningful value (as it combines two distinct design types). Rather, it will focus on the summary effects of the design-type subgroups.

In relation to the sub-group of cohort studies, Figure 4.1 and Table 4.5, show an overall very small, non-significant effect of retirement on self-reported symptoms of depression, favouring the post-retirement period (i.e. fewer symptoms of depression) (SMD= -0.077, 95% CI -0.402 to 0.247). Heterogeneity was high and significant (Q=640.728, df =4, p<0.001), where the observed variation is greater than what would be expected based on within-study error. The I^2 value suggests that 99.37% of the variance of true effect sizes can be attributed to real heterogeneity between studies. The Tau value (T=0.368), suggests that some proportion of true effects may fall in the trivial range, while others in the moderate range. The dispersion of effects is high within this subgroup, with Calvo et al. (2013) contributing with the largest, significant observed effect size (SMD= +0.333, 95% CI 0.304 to 0.362, p<0.001), in the opposite direction of the summary effect (i.e. favouring the pre-retirement period). Finally, among the cohort studies sub-group, only Kubicek et al. (2011) did not require further estimation procedures to obtain the effect-size value (Table 4.5). All other studies required some estimation, which introduces some level of uncertainty around the individual effect sizes.

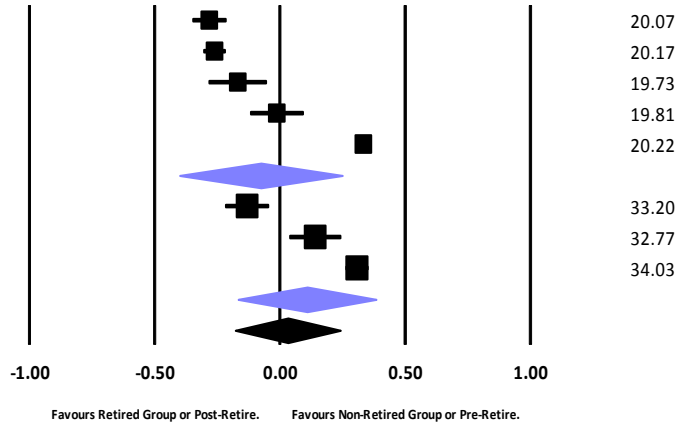
In relation to the sub-group of cross-sectional studies, Figure 4.1 and Table 4.6 also show an overall very small, non-significant effect of retirement on self-reported symptoms of depression, this time favouring non-retired samples (SMD= +0.107, 95% CI -0.169 to 0.383). Again, heterogeneity was high and significant (Q=76.611, df=2, p<0.001), with the I^2 statistic suggesting that 97.39% can be attributed to heterogeneity between cross-sectional studies. Equally, the Tau value (T=0.240) indicates that some proportion of true effects may fall in the trivial range, while others may fall in the moderate range. Again, the dispersion of effects is high within this subgroup. Steptoe et al. (2012) contribute with the largest (moderate range), significant effect size (SMD= +0.307, 95% CI 0.262 to 0.353, p<0.001), in the same direction of the summary effect. In relation to the level of extra estimation procedures for individual effect sizes, Bosse et al. (1987) was the only study in this sub-group to require some estimation, again, raising some uncertainty around this effect size value.

Overall, however, the estimation of excess dispersion (heterogeneity) between the two design-type subgroups using mixed-effects model was less than what would be expected based on within-study

error (Q<df) and not significant (Q=0.723, df=1, p<0.395). This suggests that the studies' effect sizes in the two sub-groups do not differ significantly enough to reject the hypothesis that they share a common effect size (however, this is only a tentative suggestion). The subgroup analysis in relation to design type, therefore, yielded high heterogeneity and inconsistency estimates of true effects within the two design types. This suggests that, while the overall summary effects for the two sub-groups are not meaningful (i.e. can be misleading), the exploration of further sources of heterogeneity is warranted.

Figure 4.1 – Random effects model forest plot and descriptive statistics for cohort and cross-sectional studies, grouped by study design

Group by Study Design	Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Relative weight
Cohort	Westerlund et al (2010)	-0.282	0.035	0.001	-0.350	-0.213	-8.071	0.000	20.07
Cohort	Kubicek et al (2011)	-0.261	0.023	0.001	-0.306	-0.216	-11.359	0.000	20.17
Cohort	Reitzes et al (1996)	-0.168	0.060	0.004	-0.285	-0.052	-2.824	0.005	19.73
Cohort	Clarke et al (2011)	-0.012	0.055	0.003	-0.120	0.095	-0.222	0.824	19.81
Cohort	Calvo et al (2013)	0.333	0.015	0.000	0.304	0.362	22.581	0.000	20.22
Cohort		-0.077	0.166	0.027	-0.402	0.247	-0.467	0.641	
Cross-Sectional	Drentea (2002)	-0.131	0.045	0.002	-0.219	-0.043	-2.914	0.004	33.20
Cross-Sectional	Bosse et al (1987)	0.141	0.053	0.003	0.037	0.244	2.669	0.008	32.77
Cross-Sectional	Steptoe et al (2012)	0.307	0.023	0.001	0.262	0.353	13.137	0.000	34.03
Cross-Sectional		0.107	0.140	0.020	-0.168	0.383	0.764	0.445	
Overall		0.030	0.107	0.011	-0.180	0.240	0.281	0.779	



Fixed-effect Heterogeneity Cohort Studies: Tau = 0.368; Tau² = 0.19; Q =640.728, df = 4 (p<0.001); I² = 99.37%
Fixed-effect Heterogeneity Cross-sectional Studies: Tau = 0.240; Tau² = 0.057; Q =76.611, df = 2 (p<0.001); I² = 97.39%
Mixed-effects Heterogeneity between Cohort and Cross-sectional studies: Q=0.723, df=1 (p<0.395)

Table 4.5 – Summary of effect sizes and descriptive statistics of included cohort studies

Study ID	Pre-Retire. Mean	Pre-Retire. SD	Post-Retire. Mean	Post-Retire. SD	Post-Retire. N	Correlation Pre/Post	Effect Direction	SMD	SE_{SMD}	UMD	SE_{UMD}	ES Estimation
CO4 Reitzes et al. (1996)	**5.705**	6.576	**4.673**	5.557	291	0.49	Favours post-retire.	**- 0.168**	0.060	- 1.032	0.363	3[a]
CO6 Kubicek et al. (2011)	**9.62**	7.75	**7.68**	7.07	1609	0.59	Favours post-retire.	**- 0.261**	0.023	-1.940	0.168	5
CO8 Westerlund et al. (2010)	Odds-Ratio: **0.60 (95% CI, 0.53-0.67)**				14104	-	Favours post-retire.	**- 0.282**	0.035	-	-	3[a]
Clarke (2011)	**0.305**	0.343	**0.301**	0.312	340	0.49	Favours post-retire.	**- 0.012**	0.055	- 0.004	0.018	3[a]
CO10 Calvo et al. (2013)	**0.59**	1.11	**1.14**	1.9	4938	0.49	Favours pre-retire.	**+ 0.333**	0.015	+ 0.550	0.024	3[a]

Key:
Effect Direction: effect favours the post-retirement period (-), i.e. fewer self-reported symptoms of depression
effect favours pre-retirement period (+), i.e. fewer self-reported symptoms of depression
SMD: Standardised Difference in Means; SE_{SMD}: Standard Error of the Standardised Difference in Means; UMD: Unstandardised (raw) Mean Difference; SE_{UMD}: Standard Error of the Unstandardised Mean Difference
ES Estimation level: used here as a proxy of level of confidence in the estimation of effect size (and variance) for each study, using the following scale (adapted from Lipsey and Wilson, 2001):
1 = highly estimated (have N and crude p-value only, such as p<.10 and must reconstruct via rough t-test equivalence)
2 = Moderate estimation (have complex but relatively complete statistics, such as multifactor ANOVA as basis for estimation)
3 = Some estimation (have unconventional statistics and must convert to equivalent t-values; or have conventional statistics but incomplete, such as exact p-level)
4 = Slight estimation (must use significance testing statistics, rather than descriptive statistics, but have complete statistics of the conventional sort)
5 = No estimation (have descriptive data such as means, standard deviations, frequencies, proportions, etc, and can calculate the effect size directly)
[a] Further effect size estimation procedures for Reitzes et al. (1996) and Calvo et al. (2013) in Appendix 7

Table 4.6 – Summary of effect sizes and descriptive statistics of included cross sectional studies

Study ID	Non-Retired Mean	Non-Retired SD	Non-Retired N	Retired Mean	Retired SD	Retired N	Effect Direction	SMD	SE_{SMD}	UMD	SE_{UMD}	ES Estimation
CS1 Bosse et al. (1987)	**0.28**	0.597	817	**0.36**	0.529	644	Favours non-retired group	**+ 0.141**	0.053	+ 0.080	0.030	3[a]
CS3 Drentea (2002)	**0.96**	1.33	1915	**0.79**	1.21	672	Favours retired group	**- 0.131**	0.045	- 0.170	0.058	5
CS4 Steptoe et al. (2012)	**1.16**	1.77	2734	**1.77**	2.08	5680	Favours non-retired group	**+ 0.307**	0.023	+ 0.610	0.046	5

Key:
Effect Direction: - effect favours the retired group (i.e. fewer self-reported symptoms of depression)
+ effect favours non-retired group (i.e. fewer self-reported symptoms of depression)
SMD: Standardised Mean Difference; SE_{SMD}: Standard Error of the Standardised Mean Difference; UMD: Unstandardised (raw) Mean Difference ; SE_{UMD}: Standard Error of the Unstandardised Mean Difference;
ES Estimation level: used here as a proxy of level of confidence in the estimation of effect size (and variance) for each study, same scale as above.
[a] Further effect size estimation procedures for Bosse et al. (1987) in Appendix 7

4.3.2 Further Sub-group and Sensitivity analyses

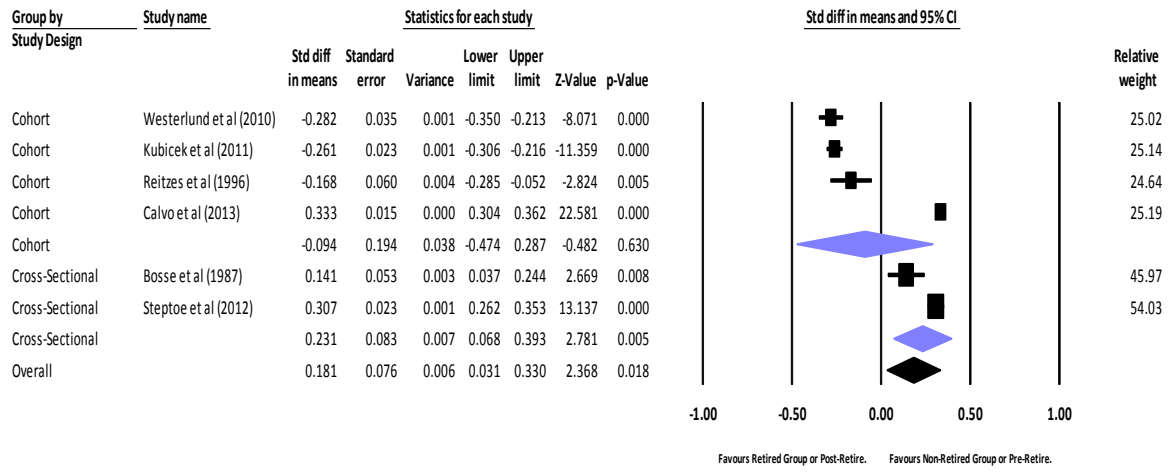
Further sub-group and sensitivity analyses sought to explore sources of heterogeneity between studies. The results are presented in relation to the *a priori* defined potential sources of heterogeneity.

4.3.2.1 Outcome measure

This analysis was initially planned to relate to contrasts between both a) the use of diagnostic and symptom-severity measures of depression, and b) self-report and 'objective/independent' indicators of depression. However, the sample of studies included in the analysis did not support the fulfilment of this objective.

One related aspect that became salient during the process of this meta-analysis was the integrity with which the CES-D (as an outcome measure) was used. With the exception of Bosse et al. (1987), all studies included in this meta-analysis used some form of the CES-D (full scale [20 items], short scale [8 items] and 'idiosyncratic' versions]. As outlined in the Table of Characteristics of Included Studies (Appendix 6), both Clarke et al. (2011) and Drentea (2002) used heavily transformed forms of the scale. This was considered to potentially compromise the measurement properties of the CES-D, and, therefore, considered unsafe measures of symptoms of depression. Furthermore, these two studies were thought to be at relatively high risk or unknown risk of bias, across all domains. Removing these two studies from, respectively, cohort and cross-sectional sub-groups, it is possible to see in Figure 4.2, that the effect of removing Clarke et al. (2011) marginally changed the overall cohort sub-group effect size, which remains non-significant (SMD= -0.094, 95% CI -0.474 to 0.287). Equally, heterogeneity and inconsistency of the effects remain high. The impact of removing Drentea (2002) was more significant within the cross-sectional sub-group. The effect size is larger (though still in the small range), but statistically significant (SMD= +0.231, 95% CI 0.068 to 0.393). Heterogeneity and inconsistency of effects estimates were still high. However, given the small sample of studies in this sub-group, the impact of removing Drentea (2002) does not allow meaningful conclusions to be drawn.

Figure 4.2 – Random effects model forest plot and descriptive statistics grouped by study design - sensitivity analysis of impact of integrity of CES-D measure.



Fixed-effect Heterogeneity Cohort Studies: Tau = 0.387; Tau^2 = 0.15; Q =636.74, df = 3 (p<0.001); I^2 = 99.53%
Fixed-effect Heterogeneity Cross-sectional Studies: Tau = 0.110; Tau^2 = 0.012; Q =8.33, df = 1 (p=0.004); I^2 = 87.99%

4.3.2.2 Date of the study

Contrasts were planned between studies conducted pre-2005 and post-2005. However, the time distribution of studies in the sample did not support the fulfilment of these contrasts. Most studies used data collected before the cut-off or, in case of Calvo et al. (2013), the cohort study measurement points (1992 and 2010) straddled the 2005 cut-off. Only one study, Steptoe at al (2012), includes data collected in 2010-2011.

4.3.2.3 Time in retirement

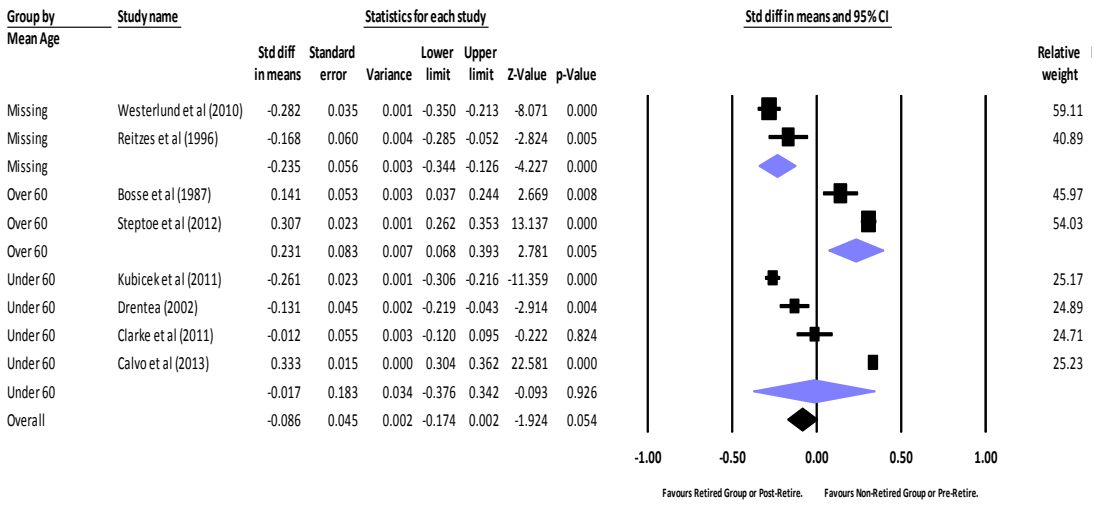
The same issue as above arose for the exploration of heterogeneity in relation to the samples' mean time in retirement. The studies included in this analysis did not regularly report this characteristic of their samples – data recorded as missing for the majority of studies.

4.3.2.4 Mean age of the sample

Although most studies reported the mean age of their samples, this was not consistent. In relation to the cohort subgroup, both Reitzes et al. (1997) and Westerlund et al. (2010) did not include complete data regarding the mean age of their studies' samples. In relation to the cross-sectional subgroup, all included data on the mean age of samples. Looking at Tables 4.1 and 4.2, it is possible to notice that Bosse et al.'s (1987) and Steptoe et al.'s (2012) samples were, on average, older (61 and 67.76) than the other samples (ranging from 47.05 to 55.67, where data was available). A sub-group comparison was made across all studies, including only those studies for which data was available, between samples with a mean age over 60 and a mean age under 60. Figure 4.3 indicates that for older samples, the effect's direction suggests worse outcomes (i.e. more self-reported depressive symptoms) in the post-retirement period or retired samples. However, this effect is still small, though statistically

significant (SMD=+0.231, 95% CI 0.069 to 0.393, p<0.005) and heterogeneity and effect inconsistency are high. Looking in turn at the effect size related to younger samples, heterogeneity and inconsistency values are high (Q =520.027, df =3, p<0.001; I^2 =99.42%), and the effect size in this sub-group is still very small and not statistically significant (SMD= -0.017, 95% CI -0.376 to 0.342, p<0.926). However, it is in the opposite direction to the effect observed in older samples, i.e. suggesting worse outcomes (i.e. more self-reported depressive symptoms) in the pre-retirement period or non-retired samples. The fact that the older samples sub-group only included two studies and combined effect-sizes across design sub-groups, does not allow conclusions to be drawn.

Figure 4.3 – Random effects model forest plot and descriptive statistics grouped by study design - sensitivity analysis of impact of mean age of sample.

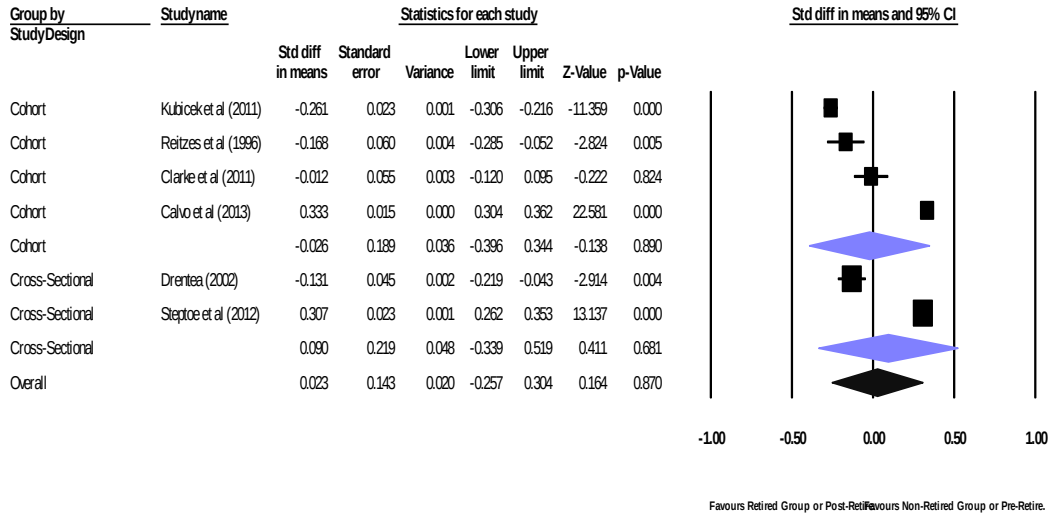


Fixed-effect Heterogeneity Over 60 Studies: Tau =0.11; Tau² =0.012; Q =8.326, df =1 (p<0.004); I² = 87.99%
Fixed-effect Heterogeneity Under 60 Studies: Tau =0.364; Tau² =0.133; Q =520.027, df =3 (p<0.001); I² =99.42%
Mixed effects Heterogeneity Between studies: Q=21.908, df=1 (p<0.001)

4.3.2.5 Gender distribution of the sample

Most studies included samples of both genders, with the exception of Bosse et al. (1987) and Westerlund et al. (2010), which included samples with 100% and 80% of men in their studies. The exclusion of both of these studies from the analysis of cohort and cross-sectional subgroups (Figure 4.4), respectively, did not change meaningfully the magnitude, direction or significance of the overall summary effect for the cohort subgroup (SMD= -0.026, 95% CI -0.396 to 0.344, p<0.890). Heterogeneity and inconsistency remained very high. The same happened in relation to the cross-sectional sub-group of studies (SDM= +0.090, 95% CI -0.339 to 0.519, p<0.681), and, again, heterogeneity and inconsistency between true effects remained high. The estimation of heterogeneity between the two design-type subgroups (using mixed-effects model), was again not significant (Q=0.161, df=1, p<0.688), where the observed variation was less than what would be expected based on within-study error (Q<df).

Figure 4.4 – Random effects model forest plot and descriptive statistics grouped by study design - sensitivity analysis of impact of gender distribution of sample.

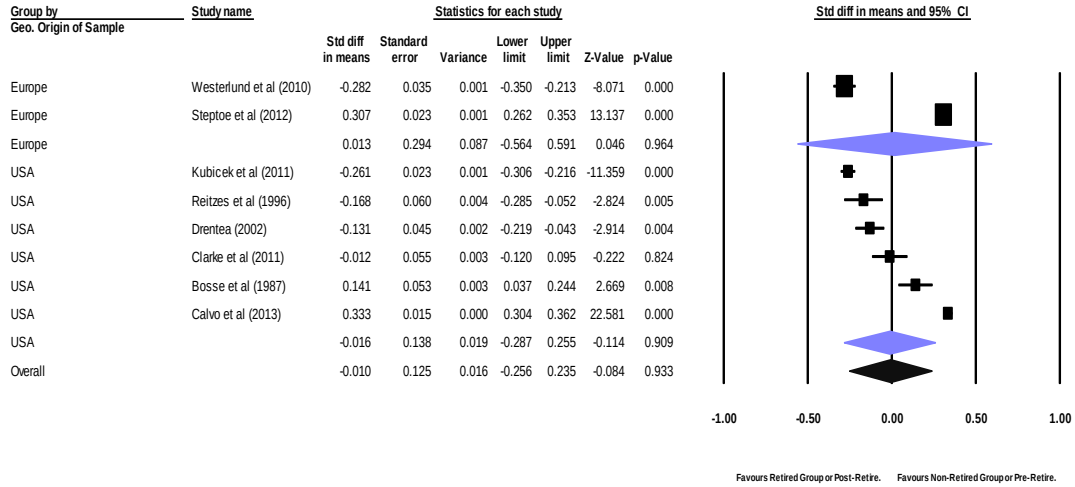


Fixed-effect Heterogeneity Cohort Studies: Tau =376; Tau² =0.141; Q =510.819, df =3 (p<0.001); I² = 99.41%
Fixed-effect Heterogeneity Cross-Sectional Studies: Tau =0.304; Tau² =0.092; Q =74.958, df =1 (p<0.001); I² =98.66%
Mixed effects Heterogeneity Between studies: Q=0.161, df=1 (p<0.688)

4.3.2.6 Geographical origin of the sample

Most studies included in this analysis used samples originating from the USA. Only Westerlund et al. (2010) and Steptoe et al. (2012) used European samples (France and UK). Both Westerlund et al. (2010) and Steptoe et al. (2012) yielded individually relatively larger (though still in the small and low-moderate ranges) effect sizes (respectively, SMD= -0.282, 95% CI -0.35 to 0.213, *N*=14104; SMD= +0.307, 95% CI 0.262 to 0.353, *N*=8414), however, in opposite directions. Analysing the sample comparing these two studies to the remaining included studies (in Figure 4.5), indicates that the heterogeneity and inconsistency between the two geographical origin sub-groups is not significant (*Q*=0.008, *df*=1, *p*<0.929, as the observed variation is less than the expected within-study error), suggesting that these two geographically defined sub-groups may share a common effect size.

Figure 4.5 – Random effects model forest plot and descriptive statistics grouped by study design - sensitivity analysis of impact of geographical location of the sample.

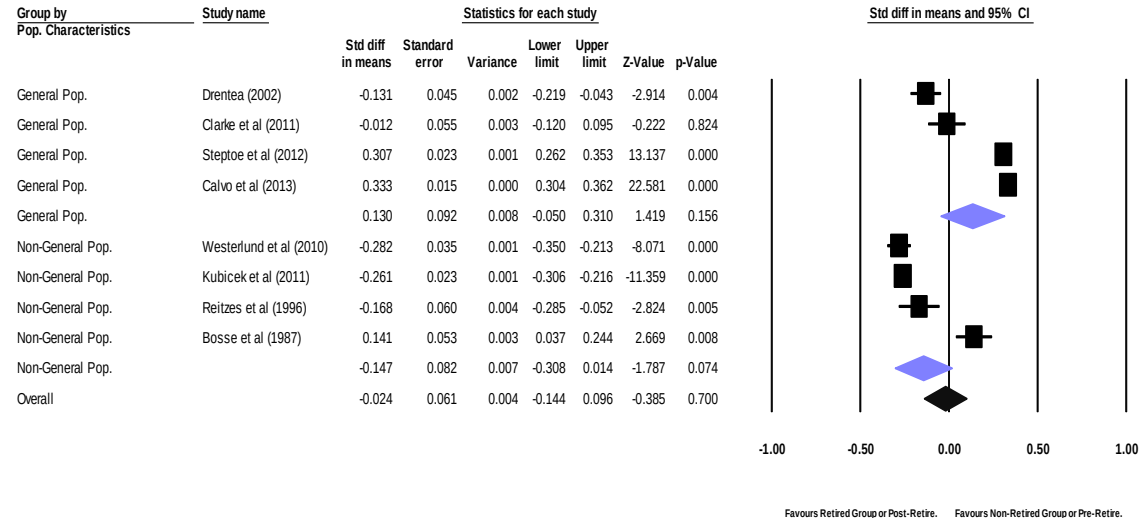


Fixed-effect Heterogeneity European Samples Studies: Tau =0.415; Tau² =0.173; Q =196.541, df =1 (p<0.001); I² = 99.49%
Fixed-effect Heterogeneity USA Samples Studies: Tau =0.336; Tau² =0.113; Q =544.606, df =5 (p<0.001); I² =99.08%
Mixed effects Heterogeneity Between studies: Q=0.008, df=1 (p<0.929)

4.3.2.7 Further sub-group analysis – population (general vs non-general)

A further exploration of heterogeneity focused on an additional characteristic of the samples used in these studies, and corresponding populations. This compared the effects yielded from studies that used samples extracted from general populations (e.g. the HRS or the ELSA) and non-general populations (e.g. the GAZEL cohort or the Carolina Health and Transitions Study). The analysis (Figure 4.6) indicated that the summary effect for each of the sub-groups was still small and non-significant, but in opposite directions (general populations, SMD=+0.130, 95% CI -0.050 to 0.310, p<0.156; non-general populations, SMD= -0.147, 95% CI -0.308 to 0.014, p<0.074). Respectively, the sub-group of samples of non-general populations indicated an effect that favoured the post-retirement period or retired samples, whereas samples from general populations suggested an effect that favoured pre-retirement period or non-retired populations. Again, while heterogeneity and inconsistency between studies in each of the sub-groups remains high, the mixed effects heterogeneity statistic (Q=5.057, df=1, p<0.025) suggests that it may be possible to reject the hypothesis that these two sub-groups share a common effect size, i.e. that the true effects vary.

Figure 4.6 – Random effects model forest plot and descriptive statistics grouped by study design - sensitivity analysis of impact of population characteristics.



Fixed-effect Heterogeneity General Pop. Samples: Tau =0.180; Tau² =0.032; Q =126.71, df =3 (p<0.001); I² = 97.632%
Fixed-effect Heterogeneity Non-Gen. Pop. Samples: Tau =0.158; Tau² =0.0.025; Q =53.717, df =3 (p<0.001); I² =94.42%
Mixed effects Heterogeneity Between studies: Q=5.057, df=1 (p<0.025)

4.3.3 Publication bias

The small number of studies included in this meta-analysis rendered the use of funnel-plot unsafe for the visual inspection and identification of publication bias (Terrin, Schmid, Lau, 2005). Therefore, the *Fail-safe N* (Rosenthal, 1979) was used. This meta-analysis incorporated data from 8 studies, which yielded a z-value of 4.5949 and corresponding 2-tailed p-value of 0.0000. The *fail-safe N* was 36, i.e. 36 'null' studies would have to be located and included in order for the combined 2-tailed p-value to exceed 0.050. That is, there would be need to be 4.5 missing studies for every observed study for the effect to be nullified. This value does not exceed Rosenthal's (1991, cited in Borenstein et al., 2009) robust fail-safe criterion (5*k*+10, where *k* is the number of included studies in this meta-analysis), and therefore suggests the possibility that these results cannot be considered robust to the effects of publication bias.

5. Discussion

5.1 Summary of main results

5.1.1 Main sub-group analysis

This meta-analysis sought to evaluate the evidence and explore the inconsistency found in studies of the relationship between retirement (as a transition) and symptoms of depression in later life. It included 8 studies, 5 cohort studies and 3 cross-sectional studies. The inconsistency of observed effects was clearly presented. The option to analyse this sample of studies in relation to design sub-groups was based on both theoretical and methodological reasons, which assumed that these studies would differ functionally in important ways (Borenstein et al., 2009; Higgins et al., 2013). It was, however, noticeable that even within the two design sub-groups, the observed dispersion of effects seemed excessive to what could be accommodated by sampling error. The summary of effects for cohort studies suggested that these evidenced a small, non-significant effect of retirement on self-reported symptoms of depression, favouring the post-retirement period (i.e. fewer symptoms of depression) (SMD= -0.077, 95% CI -0.402 to 0.247). However, heterogeneity found between these studies was very high. The summary effect size of the cross-sectional sub-group also suggested an overall very small, non-significant effect of retirement on self-reported symptoms of depression, this time favouring non-retired samples (SMD= +0.107, 95% CI -0.169 to 0.383). Again, heterogeneity was very high, strongly indicating that the variance found across these studies is real, rather than spurious. Overall, these effect sizes were very small, non-significant and potentially misleading, given the high dispersion of effects observed and true effect heterogeneity found throughout.

5.1.2 Exploration of heterogeneity

Further subgroup and sensitivity analyses were used to explore potential sources of heterogeneity (including both methodology and sample characteristics of studies). Two important caveats are made here. First, the exploration of sources of dispersion should not be perceived as suggestive of moderating variables in the relationship between retirement and symptoms of depression in later life (Borenstein et al., 2009). Second, given the small number of studies included in this review (and despite their large *Ns*) and the substantial variability found between studies, under the random-effects model, it is likely that these analyses were underpowered to find sub-group differences (Borenstein et al., 2009). The exploration of heterogeneity is therefore seen here as only informative, within these conservative parameters.

Looking first at the variable way with which the CES-D was used (this was the measure of depression in all but one study), the sensitivity analysis removed 2 studies with very idiosyncratic uses of the CES-D (Drentea, 2002; Clarke et al., 2011). These were also studies considered to be at high risk of bias, which, again, was likely to introduce increased variability of observed effects (Deeks, Dinnes, et al., 2003). The resulting summary effect size did not change significantly, and the high heterogeneity found before remained unchanged. This, however, seems to be an important characteristic of this area of research on retirement. Measures seem to be adapted (namely, shortened) to the methodological needs of studies (possibly, to increase response rates in large surveys), forfeiting the quality, measurement properties and, consequently, the validity of the instrument. This is clearly an issue that warrants the attention of authors in this area. A further sensitivity analysis queried the impact of gender distribution of samples (as a source of confounding) on the observed dispersion of effects. Two studies were singled-out as these included male only and mostly male samples (Bosse et al., 1987; Westerlund et al., 2010). However, again, the exclusion of these studies from the analysis of the sub-groups did not change meaningfully the magnitude, direction or significance of the overall summary effects. Crucially, heterogeneity and inconsistency of effect sizes remained very high.

Sub-group analyses explored the differences in age distribution of samples. However, these encountered challenging inconsistency/lack of clarity with which the age distribution of samples was reported. As an obvious potential confounder this was a surprising find. A sub-group comparison was made (including only 6 out of 8 studies, for which this data was available), between samples with mean age >60 and <60. The summary effects for each age-defined sub-group were inconsistent in direction – respectively, suggesting worse outcomes in the post-retirement period/retired samples in older samples, and worse outcomes in the pre-retirement period/non-retired samples in younger samples. However, both effect sizes were small, and, given the high heterogeneity and inconsistency found between observed effect sizes, these are not likely to be meaningful. Nevertheless, this may be suggestive (though not robustly) the complex relationship between age and self-reported depressive symptoms, in non-clinical samples (Lebowitz, Pearson, et al., 1997; Wilson, Mottram, Vassilas, 2008).

Other sub-group analyses focused on the geographical origin of samples (namely, samples from the USA [which dominate this area of research and the included dataset] and samples from Europe), and on whether the sample was extracted from a general population or non-general population. Again, the outcomes continued to reveal very small, non-significant effect sizes and, crucially, high heterogeneity and inconsistency between studies, in both sub-group analyses. However, the between-subgroups heterogeneity statistic (Q=0.008, df=1, p<0.929) suggested that European and USA samples sub-groups could potentially share a common effect size, though Borenstein et al. (2009) caution against the over-interpretation of these statistics.

Finally, two sub-group and sensitivity analyses had been planned in relation to samples' mean time in

retirement and date of the study. The first, was widely supported by the continuity and life-long development models of the impact of retirement (e.g. Atchley, 1989; Moen, 1996). If adjustment to retirement is widely considered to be a temporal process, i.e. a transition, then it is possible that the dispersion of outcomes related to symptoms of depression could be underpinned by samples that are assessed at different stages of this transition, but never systematically analysed in this sense (Horner, 2012). The fact that studies included in this analysis did not consistently report the mean time in retirement of samples or mean age of retirement, did not allow this kind of analysis. This is likely to be a powerful source of variability of effects, one that is theoretically consistent and would therefore warrant more rigorous attention by authors in this area (Butterworth et al., 2006; Calvo and Sarkisian, 2011).

5.2 Quality of the evidence

Overall, the assessment of domain-specific bias of each study suggested variability between studies in this sense. Clarke et al. (2011) and Drentea (2002) presented very specific concerns regarding most bias domains. The decision to include them in the analysis, though problematic (Valentine and Thompson, 2013), was based on an exploratory stance towards these data (Lipsey and Wilson, 2001). The observed effects in this meta-analysis are likely to reflect the impact of the variable methodological quality of studies, and other unknown or unmeasured confounding factors. The introduction of excessive heterogeneity, in this sense, may have increased the potential for true effects to be missed (Higgins et al., 2013). Siegfried et al. (2005) and Valentine and Thompson (2013) argue that NRSs, unlike RCTs, can only adjust for known confounders, and only those that are measured without error. Assessing the potential risk of bias from confounding in these studies, 6 factors (age, gender, etc.) were initially identified. However, studies did not consistently or clearly report or measure the impact of these factors. Another important characteristic of almost all included studies was the unclear way in which these presented information about procedures to deal with missing data. Although studies included very large samples, there was regular unexplained inconsistency between the stated *Ns* in methodology sections, and the *Ns* that were subsequently reported in analyses/outcome tables. Further, with the exception of two studies, most required some level of estimation during the calculation of their individual effect sizes because of missing or unclear data. This necessarily introduced a level of uncertainty in the calculation of overall effect summary and estimation of heterogeneity. Finally, endogeneity bias (e.g. Coe and Lindeboom, 2008), was a concern that only two studies addressed (Westerlund et al., 2010; and Calvo et al., 2013). This is a critical problem in this field of research, as the observation of poor health after retirement, cannot be un-problematically attributed to retirement itself (Calvo et al., 2013).

5.3 Completeness and applicability of evidence

This meta-analysis was based on an extensive, inclusive and systematic data search, balanced with specific inclusion/exclusion criteria, to manage the risk of introducing excessive inconsistency (Sharpe, 1997) into the review. A bi-product of this option, however, was the potential for the final sample of included studies to represent a limited picture of this area of research. Three issues related to the external validity of this review are explored here.

First, the selection strategy sought to apply rigour in relation to the definition of an outcome to be measured. It identified depression as an outcome with clinical relevance and one which offered the possibility of a more consistent measurement of phenomena (i.e. managing risk of detection bias). As described before and elsewhere (e.g. Coelho and Newman, 2014), this is an area of research where constructs related to mental health or wellbeing are used variably and, at times, interchangeably. This methodological option restricted the analysis to studies that explicitly identified and measured depression, excluding broader constructs of 'mental health'. Equally, this review applied stringent criteria to the definition and operationalisation of retired status itself. It was apparent that many studies took a simplistic approach to this, taking the age threshold of 'over 65' to mean 'retired'. This research practice, however, was considered too prone to performance bias. These options necessarily introduced a limitation in the types of outcomes investigated. In that sense, the studies included in this review do not fully represent the population of outcomes in this area of research. However, it is argued that these adequately represent the population of studies that define retirement appropriately and their outcomes clearly as investigations of depression. Second, the analysis of publication bias queried the robustness of the results of this review. Although this could be said in relation to any kind of empirical effort, Sterne et al. (2011) note that cohort studies are particularly vulnerable to the selective publication of large and/or statistically significant results. There are multiple past and ongoing very large cohort studies in this area, mostly using survey designs. Some of these studies make their data electronically accessible, as a data corpus. Among the included studies, 6 out of 8 studies were secondary analyses of data, and all 7 excluded studies were also secondary analyses of this kind. A possible consequence of this type of research landscape (common in fields of epidemiology, economics and medical research) is that it makes it likely for researchers to both a) engage in multiple publications based on the same data, and b) 'try out' analyses, which may not yield "publishable" results. It is therefore possible that this review has been challenged by this issue. Finally, the study of retirement is obviously a multidisciplinary area, and its literature is diverse and spread out. It is then likely, then, that there are studies that are systematically undetected in this area (Reeves et al., 2011; Reeves et al., 2013).

Overall, the area of studies of retirement, given these characteristics and its multi-disciplinary nature, presents challenges to efforts to systematically review and summarise the evidence. Although it is likely that the included studies are not exhaustive of this population of publications, the outcomes of

the review process may be nonetheless representative of a review effort in this area, with what data is available, at this point in time.

5.4 Potential biases in the review process

This review attempted to employ robust and inclusive search and selection strategies, and rigorous analytical methods, following gold-standard methodological and statistical guidance (Cochrane Collaboration and Borenstein et al., 2009). The strengths of this review lie a) in its comprehensive coverage of the literature (Siegfried et al., 2005); b) the systematic assessment of the risk of bias (and therefore quality) of included studies using an assessment tool specifically developed for this review; and c) the rigorous exploration of the influence of *a priori* hypothesised sources of heterogeneity of effects. However, as any review process, it entailed iterative decision-making processes, that will necessary introduce bias (i.e. systematic error), and was also dependent on the characteristics of the available data.

5.4.1 Analytic decisions and process

As argued above, the decision to use stringent inclusion criteria in relation to operational definitions of retirement and depression may have limited the range of effects that were available to this analysis. Widely cited publications and outcomes/effects in this area (e.g. Mein, et al., 2003; Gill, et al., 2006) were not included in this synthesis. Furthermore, the analysis suffered from a lack of available usable statistical data at its final stage, which again will have impacted on its outcome. Namely, data yielding from the SHARE, LASA and CHARLS (large studies of retirement and ageing) were missing from this process, although significant efforts were made to obtain them. Focusing on the sample of included studies, the statistical analysis sought to include as much information as possible, requiring some estimation of terms used to calculate effect sizes. Along with the inclusion of studies with potential high vulnerability to bias, as addressed above, this introduced a level of uncertainty about the validity of the summary effects. Finally, it is very likely that other sources of variability of true effects underlie the dispersion of observed effects that was found. Issues such as co-current health status or financial status may have shed further light onto this relationship, if that data had been available.

5.4.2 Some characteristics of the research area

The volume of existing studies in this area, both published and unpublished, was noticeable. The issue of multiple reporting is understandable as an outcome of academic professional and social processes (e.g. Barnes, Bloor, Henry, 1996). However, it introduces a misleading impression of variety of outcomes and evidence, when it may be, in fact, a case of variety of analysis and reporting. The issue is that this variety of analysis and reporting introduces confounding that are difficult to identify and

explore (Siegfried et al., 2005). The heterogeneity found in this analysis could be, to some extent, related to this characteristic of this area. The concern here is also that the multiple analysis and reporting of these data introduces an impression of new knowledge regarding a specific relationship between these or more variables, when it is, at times, a further secondary unpacking and partitioning of an original effect. This introduces an amount of literature 'noise', which is challenging and introduces bias to a review process.

Furthermore, the variability of terms, concepts and measures of concepts was also noticeable. This challenge is not exclusive to this area, rather, it is a core problem of observation and measurement in psychology (Barnes et al., 1996). However, it poses an essential challenge to the development of reviews and syntheses, as the methodology demands the ability to compare like with like (Lipsey and Wilson, 2001). The observable variability in the measurement of depression, even when studies used the "same" depression scale, was surprising. Studies not only shortened or partitioned that scale (e.g. Clarke et al., 2011), but also changed response formats and scaling to suit the study's aims (Steffick, 2000; Kubicek et al., 2011; Calvo et al. 2013). The various versions of the CES-D used do not guarantee the maintenance of the measurement properties of the scale, and therefore may have introduced additional confounding.

5.5 Implications of the evidence

5.5.1 Research implications

As outlined in the introduction to this meta-analysis, there is growing interest in the question of the relationship between retirement and depression, mental health or health in general. This is assumed to be complex and multi-determined. So far, however, partial and non-systematic reviews of the literature have reiterated the apparent variability found in this relationship, and speculated about its potential sources (e.g. Reitzes, Mutran, Fernandez, 1996; Mein, Martikainen, Hemingway, Stansfels, Marmot, 2003; Wang et al. 2011; Oksannen et al. 2011). This review has added to this area of research: a) a quantitative synthesis of this evidence (with the limits and caveats addressed above in relation to its external and internal validity); and b) a systematic exploration of some of these characteristics in relation to the dispersion of these effects. The upshot is that the characteristics that were explored do not account for the estimated heterogeneity of true effects. This opens up the possibility of generating further questions that may be more relevant to our understanding of the relationship between retirement and depression.

Specifically, it is argued that, in further review and synthesis studies of this relationship, a focus on timing of retirement, time spent in retirement and the control for endogeneity bias are key potential confounders that, if analysed systematically, may shed light on the dispersion of effects in this area.

These are also theoretically meaningful, consistent with the conceptualisation of retirement as a normative late-life transition. Additionally, factors such as co-occurring physical health and financial status would also be meaningful candidates to the moderation of this relationship. However, the possibility of this kind of further analysis is dependent on what data is available in primary studies. The large studies in this area have both advantages and drawbacks. On the one hand, the availability of large cohorts assures power to observe relationships and effects when they truly exist. On the other hand, the design of these large studies needs to satisfy a large number of research aims. The design compromises that are made along the way may hinder the potential to focus on relationships that make specific demands of the data. In a sense, these studies may be useful for wide characterisations of ageing and trends in later-life, but less useful in relation to specific phenomena, such as retirement. Potentially, smaller cohort or matched pairs design studies would allow a more focused look at the phenomenon of retirement.

5.5.2 Clinical implications

Depression or depressive symptoms are widely considered to be the most common mental ill-health presentations in later life (e.g. Blazer, 2003; Luppa, and Sikorski, et. al, 2011). The prevalence of mild and sub-clinical symptoms of depression may be higher and less likely to be detected in this population (Karel and Hirichsen, 2000; Blazer, 2003; Luppa, and Sikorski, et. al, 2012). This meta-analysis was motivated by: a) the complex implications that retirement may bring as a late-life transition (like other developmental transitions), and b) the personal, social and economic costs that depression has on the individual (Kleine-Budde and Muller, et al., 2013). It was anticipated that an integrated understanding of this relationship would allow insight into a significant event in latter part of our lives and, if relevant, point to an important pathway for early (sub-clinical) identification (and intervention) in late life depression (Brugiavini, Croda, Dewey, 2008b). The outcomes of this meta-analysis suggest that the relationship between retirement and self-reported symptoms of depression is not straightforward and multi-determined. While the overall effect sizes were very small (and outwith clinical range), it was hypothesised that these observed effects may have suffered from the impact of confounding in primary studies, and potentially true effects may have been masked.

The acknowledgement and empirical verification of the complexity of this relationship is in itself important. Given the contradictory nature of the conclusions of previous studies, research in this area runs to risk of being selectively used to support (or question) social and economic policy related to retirement. It is possible that, at this point in time, the characterisation of retirement as a detrimental event to mental and physical health (e.g. Sahlgren, 2013) has been given specific emphasis, as it is consistent with current political and economic decisions and outlook in this area. This meta-analysis, by contrast, emphasises that such statements related to the association between retirement and depression are not, in themselves and at this point, robustly evidence-based.

Nonetheless, it is argued, that this question warrants continued interest, especially now, as the reality of retirement is changing critically. It is also clear that the conceptualisation of retirement as a significant transition warrants sustained clinical interest. Specifically, in working clinically with older individuals who present with symptoms of depression, their retirement (circumstances, timing, losses, gains, process, etc) should not be seen as a detail of their later life; rather, it should be queried and explored as a meaningful transition in their life-long development.

6. Conclusions

Retirement is an important late-life transition, and it carries significant implications for every aspect on a person's life, including their mental health. This, however, is also an area where outcomes of research on the significance of retirement to individuals' mental health are inconsistent and lacking in integration. This meta-analysis was a first effort in a systematic, quantitative integration of this evidence. It focused specifically on the relationship between retirement and self-reported depressive symptoms. Though mindful of its limitations, this meta-analysis suggests that this relationship is complex (both methodologically and substantively) and multi-determined. Summary and individual studies' observed effects sizes were mainly small, non-significant, and in different directions, and, the dispersion of effect-sizes and the true-effects heterogeneity was found to be very high. This, it is argued, is underpinned by mostly unknown (or possibly unmeasured) confounding factors in primary studies.

Highlights (see Appendix A):

- Relationship between retirement and self-reported depressive symptoms is complex and highly variable.
- Effect-sizes (individual studies and summary effects) are small, non-significant and dispersed.
- Heterogeneity of true effects is high.
- This is potentially underpinned by mostly unknown or unmeasured confounding factors in primary studies.

References

Key:

Papers included in the Meta-Analysis of Cross-Sectional studies are indicated with one asterisk *

Papers included in the Meta-Analysis of Cohort studies are indicated with **

All papers excluded due to lack of complete data are indicated with °

Aalto, A.M., Elovainio, M., Kivimäki, M., Uutela, A., Pirkola, S. (2012). The Beck Depression Inventory and General Health Questionnaire as measures of depression in the general population: A validation study using the Composite International Diagnostic Interview as the gold standard. *Psychiatry Research*, vol. 197, p. 163-171.

Alavinia, S.M., Burdorf, A. (2008). Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. *International Archives of Occupational and Environmental Health*, vol. 82, p. 39-45

Americans' Changing Lives Survey, http://www.isr.umich.edu/acl/data.htm

Angelini, V., Brugiavini, A., Webber, G. (2009). Aging and unused capacity in Europe: is there an early retirement trap? *Economic Policy*, vol. 24(59), p. 463-504.

Atchley, R.C. (1971). Retirement and leisure participation: continuity or crisis? *The Gerontologist*, vol. 11(1), p. 13-17.

Atchley, R.C. (1979). Issues in retirement research. *The Gerontologist*, vol. 19(1), p. 44-54.

Atchley, R.C. (1982). Retirement as a social institution. *Annual Review of Sociology*, vol. 8, p. 263-287.

Atchley, R.C. (1989). A Continuity theory of normal aging. *The Gerontologist*, vol. 29(2), p. 183-190.

Ayalom, L. (2011). The prevalence and predictors of passive death wishes in Europe: a 2-year follow-up of the survey of health, ageing, and retirement in Europe. *International Journal of Geriatric Psychiatry*, vol. 26, p. 923–929.

Bai, A., Shukla, V.K., Bak, G., Wells, G. (2012). *Quality Assessment Tools Project Report*. Ottawa: Canadian Agency for Drugs and Technologies in Health. Retrieved on the 28th of April 2013 from http://www.cadth.ca

Barnes, B., Bloor, D., Henry, J. (1996). *Scientific knowledge: a sociological analysis*. London: Athlone Press.

Behncke, S. (2012). Does retirement trigger ill health? *Health Economics*, vol. 21, p. 282-300.

Blazer D.G. (2003). Depression in late life: review and commentary. *Journal of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 58(3), p. 249–265.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R. (2009). *Introduction to Meta-Analysis*. Chichester: John Wiley & Sons, Ltd.

* Bosse, R., Aldwin, C.M., Levenson, M.R., Ekerdt, D.J. (1987). Mental health differences among retirees and workers: findings from the Normative Aging Study. *Psychology and Aging*, vol. 2(4), p. 383-389.

Brugiavini, A., Pasini, G., Peracchi, F. (2008a). Exits from the labour force. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 206-214. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

° Brugiavini, A., Croda, E., Dewey, M. (2008b). Retirement and mental health. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 247-254. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

Buber, I., Engelhardt, H. (2011). The association between age and depressive symptoms among older men and women in Europe. findings from SHARE. *Comparative Population Studies*, vol. 36(1), p. 103-126.

Bunce, D., Batterhamb, P.J., Mackinnon, A.J., Christensen, H. (2012). Depression, anxiety and cognition in community-dwelling adults aged 70 years and over. *Journal of Psychiatric Research*, vol. 46, p. 1662-1666.

Butterworth, P., Gill, S.C., Rodgers, B., Anstey, K.J., Villamil, E., Melzer, D. (2006). Retirement and mental health: analysis of the Australian national survey of mental health and wellbeing. *Social Science and Medicine*, vol. 63, p. 1179-1191.

Buxton, J., Singleton, N., Melzer, D. (2005). The mental health of early retirees: national interview survey in Britain. *Social Psychiatry and Psychiatric Epidemiology*, vol. 40, p. 99-105.

Cahill, K. E., Giandrea, M. D., Quinn, J. F. (2006). Retirement patterns from career employment. *The Gerontologist*, 46(2), p. 514–523.

Calasanti, T. (1996). Gender and life satisfaction in retirement: an assessment of the male model. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 51B(1), p. S18-S29.

Calvo, E., Sarkisian, N. (2011). *Retirement and well-being: examining the characteristics of life-course transitions*. Retrieved on the 10th of February 2013 from www.politicaspublicas.udp.cl

** Calvo, E., Sarkisian, N., Tamborini, C.R., (2013). Causal effects of retirement timing on subjective physical and emotional health. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 68(1), p. 73–84

Campion, J., Bhui, K, Bhugra, D. (2012). European Psychiatric Association (EPA) guidance on prevention of mental disorders. *European Psychiatry*, vol. 27, p. 68–80

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., Bula, C., Reisches, F., Wancata, J., Ritchie, K., Tsolaki, M., Mateos, R., Prince, M. (2008). Ascertaining late-life depressive symptoms in Europe: an evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE project. *International Journal of Methods in Psychiatric Research*, vol. 17(1), p. 12-29.

Charles, K.K. (2002). *Is retirement depressing? Labour force inactivity and psychological wellbeing in later life.* Working Paper, National Bureau of Economic Research. Retrieved on the 9th of March 2013 from http://www.nber.org/papers/w9033

Cheshire, H., Hussey, D., Phelps, A., Wood, N. (2012). Chapter 5: Methodology. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), *The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5)*, p. 183-205. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

Christ, S.L., Lee, D.J., Fleming, L.E., LeBlanc, W.G., Arheart, K.L., Chung-Bridges, K., Caban, A.J., McCollister, K.E. (2007). Employment and occupation effects on depressive symptoms in older Americans: does working past the age of 65 protect against depression? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(6), p. S399-S403.

Clark, A., Friedman, M.J. (1983). Factor structure and discriminant validity of the SCL-90 in a veteran psychiatric population. *Journal of Personality Assessment*, vol. 47, p. 396-404.

Clark, R. L., Quinn, J. F. (2002). Patterns of work and retirement for a new century. *Generations*, 26(2), p. 17-24.

Clarke, M., Oxman, A.D., Paulsen, E., Higgins, J.P.T., Green, S. (2011). Appendix A: Guide to the contents of a Cochrane Methodology protocol and review. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 25th of September 2012 from www.cochrane-handbook.org.

** Clarke, P., Marshall, V., House, J., Lantz, P. (2011). The social structuring of mental health over the adult life course: advancing theory in the sociology of aging. *Social Forces*, vol. 89(4), p. 1287-1313.

Coe, N.B., Lindeboom, M. (2008). *Does retirement kill you? Evidence from early retirement windows*. Retrieved on the 8th of February 2013 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295315

Coe, N.B., Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, vol. 30, p. 77-86.

Coelho, C., Newman, E. (2014). *A systematic review of constructs and measures of adjustment and adaptation to retirement*. Unpublished Doctoral Thesis, University of Edinburgh.

Conwell Y. (2009). Suicide prevention in later life: a glass half full, or half empty? *American Journal of Psychiatry*, vol 166, p. 845–848.

Conwell, Y., Duberstein, P. R., Caine, E. D. (2002). Risk factors for suicide in later life. *Biological Psychiatry*, vol. 52(3), p. 193-204.

De Luca, G., Rossetti, C. (2008). Sampling design and weighting strategies in the second wave of SHARE. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 333-338. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

Deeks, J.J., Dinnes, J., D'Amico, R., Sowden, A.J., Sakarovitch, C., Song, F., Petticrew, M., Altman, D.G. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, vol. 7, p. 27-36.

Denton, F.T., Spencer, B.G. (2009). What is retirement?: A review and assessment of alternative concepts and measures. *Canadian Journal on Aging /La Revue Canadienne du Vieillissement*, vol. 28(1), p. 63-76.

Derogatis, L.R. (1983). *SCL-90-R: Administration, scoring and procedures manual-II for the R(evised) version and other instruments of the psychopathology rating scale series (2nd Edition)*. Towson: Clinical Psychometric Research.

* Drentea, P. (2002). Retirement and mental health. *Journal of Aging and Health*, vol. 14, p. 167-194.

Ekerdt, D.J. (2010). Frontiers of research on work and retirement. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 65B(1), p. 69-80.

Ekerdt, D. J., Bossé, R., & Levkoff, S. (1985). An empirical test of phases of retirement: findings from the normative aging study. *Journal of Gerontology*, vol. 40(1), p. 95-101.

Ekerdt, D.J., DeViney, S. (1990). On defining persons as retired. *Journal of Aging Studies*, vol. 4(3), p. 211-229.

Elder, G.H. (1994). Time, human agency, and social change: perspectives on the life course. *Social Psychology Quarterly*, vol. 57(1), p, 4-15.

Elder, G.H., Johnson, M.K., Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer, J.T and Shanahan, M.J. (Eds.), *Handbook of the life course*, p. 3-19. New York: Kluwer.

Elwood, M. (2007). *Critical Appraisal of Epidemiological Studies and Clinical Trials, 3rd Edition*. Oxford: Oxford University Press.

English Longitudinal Study of Ageing (ELSA), http://www.ifs.org.uk/ELSA

Fe, E., Hollingsworth, B. (2012). *Estimating the effect of retirement on mental health via panel discontinuity designs*. Retrieved on the 7th of April 2013 from http://mpra.ub.uni-muenchen.de/38162/

Fiske, A., Wetherell, J.L., Gatz, M. (2009). Depression in older adults. *Annual Review of Clinical Psychology*, vol. 5, p. 363–389.

Fonner V.A., Denison J., Kennedy C.E., O'Reilly K., Sweat M. (2012). Voluntary counselling and testing (VCT) for changing HIV related risk behaviour in developing countries. *Cochrane Database of Systematic Reviews*, Issue 9. Art. No.: CD001224. DOI:10.1002/14651858.CD001224.pub4.

Fonseca, A.M. (2007). Determinants of successful retirement in a Portuguese population. *Reviews in Clinical Gerontology*, vol. 17(3), p. 219-224.

Gall, T.L., Evans, D.R., Howard, J. (1997). The retirement adjustment process: changes in the wellbeing of male retirees across time. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 52B(3), p. 110-117.

Gall, T.L., Evans, D.R. (2000). Preretirement expectations and the quality of life of male retirees in later retirement. *Canadian Journal of Behavioural Science*, vo. 32(3), p. 187-197.

Gallo, W.T., Bradley, E.H., Siegel, M., Kasl, S.V. (2000). Health effects of involuntary job loss among older workers: findings from the health and retirement survey. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 55B(3), p. S131–S140.

Gallagher, D., Savva, G.M., Kenny, R.A., Lawlor, B.A. (2013). What predicts persistent depression in older adults across Europe? Utility of clinical and neuropsychological predictors from the SHARE study. *Journal of Affective Disorders*, vol. 147, p. 192-197.

George, L.K., Fillenbaum, G.G., Palmore, E. (1984). Sex differences in the antecedents and consequences of retirement. *Journal of Gerontology*, vol. 39(3), p. 364-371.

Gill, S.C., Butterworth, P., Rodgers, B., Anstey, K.J. , Villamil, E., Melzer, D. (2006). Mental health and the timing of men's retirement. *Social Psychiatry and Psychiatric Epidemiology*, vol. 41, p. 515-522.

Goldberg, D.P., 1972. *The Detection of Psychiatric Illness by Questionnaire*. Oxford: Oxford University Press.

Goldberg, D.P., Gater, R., Sartorius, N., Ustun, T.B., Piccinelli, M., Gureje, O., Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, vol. 27, p. 191-197.

Golder, S., Loke, Y.K. (2009). Search strategies to identify information on adverse effects: a systematic review. *Journal of the Medical Library Association*, vol. 97(2), p. 84-92.

Guillemard, A.M., Rein, M. (1993). Comparative patterns of retirement: recent trends in developed

societies. *Annual Review of Sociology*, vol. 19, p. 469-503.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology* , vol. 6, p. 278–296.

Han, S.K., Moen, P. (1999). Clocking out: temporal patterning of retirement. *American Journal of Sociology*, vol. 105, p. 191-236.

Hardy, M. A. (2002). The transformation of retirement in twentieth-century America: from discontent to satisfaction. *Generations*, vol. 26(2), p. 9-16.

Hartling, L., Bond, K., Santaguida, P.L., Viswanathan, M., Dryden, D.M. (2011). Testing a tool for the classification of study designs in systematic reviews of interventions and exposures showed moderate reliability and low accuracy. *Journal of Clinical Epidemiology*, vo. 64, p. 861-871.

Hauser, R.M., Weir, D. (2010). Recent developments in longitudinal studies of aging in the United States. *Demography*, vol. 47 (Supplement), p. S111-S130.

Havinghurst, R.J. (1961). Successful aging. *The Gerontologist*, vol. 1, p. 8-13.

Hershey, D., Henkens, K., Van Dalen, H. (2007). Mapping the minds of retirement planners. A cross-cultural perspective. *Journal of Cross-Cultural Psychology*, vol. 38, 361-382.

Higgins, J.P.T., Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, vol. 21, p. 1539-1558.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, vol. 327, p. 557-560.

Higgins, J. P. T., Deeks, J.J. (2011). Chapter 7: Selecting studies and collecting data. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Altman, D.G., Sterne, J.A.C. (2011). Chapter 8: Assessing risk of bias in included studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*. The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Ramsay, C., Reeves, B.C., Deeks, J.J., Shea, B., Valentine, J.C., Tugwell, P., Wells, G. (2013) Issue relating to study design and risk of bias when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 12-25.

Higo, M., Williamson, J.B. (2008). Retirement. In, D. Carr, R. Crosnoe,, M. E. Hughes, A. M. Pienta (Editors), *Encyclopedia of the life course and human development vol. 3 (Social Sciences)*, p. 328-336. Detroit: MacMillan Reference.

Hofer, S.M., Piccinin, A.M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods*, vol. 14(2), p. 150-164.

Horner, E.M. (2012). Subjective well-being and retirement: analysis and policy recommendations. *Journal of Happiness Studies.* Published Online on 1st December 2012, DOI 10.1007/s10902-012-9399-2

Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A., Moschetti, I., Phillips, B., Thornton, H. (2011). *The 2011 Oxford CEBM Evidence Levels of Evidence: Introductory Document.* Oxford Centre for Evidence-Based Medicine. Retrieved on 29th of April 2013 from http://www.cebm.net/index.aspx?o=5653

Hyde, M., Ferrie, J., Higgs, P., Mein, G., Nazroo, J. (2004). The effects of pre-retirement factors and pre-retirement route on circumstances in retirement: findings from the Whitehall II Study. *Aging and Society*, vol. 24, p. 274-296.

Ioannidis, J.P. (1998). Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *Journal of the American Medical Association*, vol. 279, p. 281-286.

Jajodia, A., Borders, A. (2011). Memory predicts changes in depressive symptoms in older adults: a bidirectional longitudinal analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 66(5), p. 571-581.

° Jang, S-N., Cho, S-I., Chang, J., Boo, K., Shin, H-G., Lee, H., Berkman, L.F. (2009). Employment status and depressive symptoms in Koreans: results from a baseline survey of the Korean longitudinal study of aging. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 64(5), p. 677-683.

Jeffris, B.J., Nazareth, I., Marston, L., Moreno-Kustner, B., Bellón, J.A., Svab, I., Rotar, D., Geerlings, M.I., Xavier, M., Goçalves-Pereira, M., Vicente, B., Saldivia, S., Aluoja, A., Kalda, R., King, M. (2011). Associations between unemployment and major depressive disorder: evidence from an international, prospective study (the PREDICT cohort). *Social Science and Medicine*, vol. 73, p. 1627-1634.

Jokela, M., Ferrie, J.E., Gimeno, D., Chandola, T., Shipley, M.J., Head, J., Vahtera, J., Westerlund, H., Marmot, M.G., Kivimaki, M. (2010). From midlife to early old age: health trajectories associated with retirement. *Epidemiology*, vol. 21(3), p. 284-290.

Joyce, K., Pabayo, R., Critchley, J.A., Bambra C. (2010). Flexible working conditions and their effects on employee health and wellbeing. *Cochrane Database of Systematic Reviews*, Issue 2. Art. CD008009. DOI: 10.1002/14651858.CD008009.pub2.

Katz, I. R. (1996). On the inseparability of mental and physical health in aged persons: Lessons from

depression and medical comorbidity. *The American Journal of Geriatric Psychiatry*, vol. 4, p. 1–16.

Keller, S.D, Ware, J.E., Bentler, P.M., Aaronson, N.K., Alonso, J., Apolone, G., Bjomer, J.B., Brazier, J., Bullinger, M., Kaasa, S., Leplège, A., Sullivan, M., Gandek, B. (1998). Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology*, vol. 51(11), p. 1179-1188.

Kim, J. E., Moen, P. (2001). Is retirement good or bad for subjective well-being? *Current Directions in Psychological Science*, vol. 10(3), p. 83-86.

° Kim, J.E., Moen, P. (2002). Retirement transitions, gender and psychological wellbeing: a life course ecological model. *Journals of Gerontology Series B: Psychological Science and Social Sciences*, vol. 57(3), p. 212-222.

Kleine-Budde, K., Muller, R., Kawohl, W., Bramesfeld, A., Moock, J., Rossler, W. (2013). The cost of depression: a cost analysis from a large database. *Journal of Affective Disorders*, vol. 147, p. 137–143.

Kohli, M., Künemund, H. (2002). La fin de carrière et la transition vers la retraite: Les limites d'âge chronologiques sont-elles un anachronism? *Retraite et société*, vol. 36, p. 81-107.

Kolodziej, I. (2011). *The relationship between retirement and mental health: investigating the causal relationship in eleven European countries using SHARE*. Unpublished MSc Thesis, Erasmus University Rotterdam. Retrieved on the 8th of February 2013 from http://arno.uvt.nl/show.cgi?fid=122185

** Kubicek, B., Korunka, C., Raymo, J.M., Hoonakker, P. (2011). Psychological wellbeing in retirement: the effects of personal and gendered contextual resources. *Journal of Occupational Health Psychology*, vol. 16(2), p. 230-246.

Kuerbis, A., Sacco, P. (2012). The impact of retirement on the drinking patterns of older adults: a review. *Addictive Behaviors*, vol. 37, p. 587-595.

° Laaksonen, M., Metsä-Simola, N., Martikainen, P., Pietiläinen, O., Rahkonen, O., Gould, R., Partonen, T., Lahelma, E. (2012). Trajectories of mental health before and after old-age and

disability retirement: a register-based study on purchases of psychotropic drugs. *Scandinavian Journal of Work and Environmental Health*, vol. 38(5), p. 409-417.

Ladin, K. (2012). Decomposing differences in utilization of health services between depressed and non-depressed elders in Europe. *European Journal of Ageing*, vol. 9(1), p. 51-64.

Leber, U., Wagner, A. (2007). *Early and phased retirement in European companies: establishment survey on working time 2004–2005*. Retrieved on the 2nd of March 2013 from http://www.eurofound.europa.eu/publications/htmlfiles/ef0688.htm

Lebowitz, B.D., Pearson, J.L., Schneider, L.S., Reynolds, C.F. III, Alexopoulos, G.S., Bruce ,M.L., Conwell, Y., et al. (1997). Diagnosis and treatment of depression in late life: consensus statement update. *Journal of the American Medical Association*, vol. 278, p. 1186–1190.

Lefebvre, C., Manheimer, E., Glanville, J. (2011). Chapter 6: Searching for studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

° Lindeboom, M., Portrait, F., van den Berg, G.J. (2002). An econometric analysis of the mental-health effects of major events in the life of older individuals. *Health Economics*, vol. 11, p. 505-520.

Lipsey, M.W., Wilson, D.B. (2001). *Practical Meta-Analysis*. Thousand Oaks: Sage Publications, Inc.

Löckenhoff, C.E. (2012). Understanding retirement: the promise of life-span developmental frameworks. *European Journal of Ageing*, vol. 9, p. 227-231.

Longitudinal Aging Study Amsterdam (LASA), http://www.lasa-vu.nl/index.htm

Luborski, M. R., LeBlanc, I.M. (2003). Cross-cultural perspectives on the concept of retirement: an analytic redefinition. *Journal of Cross-Cultural Gerontology*, vol. 18, p. 251-271.

Luhmann, M., Hoffman, W., Eid, M., Lucas, R.E. (2012). Subjective wellbeing and adaptation to life-events: a meta-analysis. *Journal of Personality and Social Psychology*, vol. 102(3), p. 592-615.

Luppa, M., Sikorski, C., Luck, T., Weyerer, S., Villringer, A., König, H-H., Riedel-Heller, S. G. (2011). Epidemiology of depression in old age: results of the Leipzig Longitudinal Study of the Aged (LEILA 75+). *Journal of Epidemiology and Community Health*, vol. 65, Suppl 1, p. 127.

Luppa, M., Sikorski, C., Luck, T., Weyerer, S., Villringer, A., König, H-H., Riedel-Heller, S. G. (2012). Prevalence and risk factors of depressive symptoms in latest life – results of the

Leipzig Longitudinal Study of the Aged (LEILA 75+). *International Journal of Geriatric Psychiatry*, vol. 27, p. 286-295.

Maimaris, W., Hogan, H., Lock, K. (2010). The impact of working beyond traditional retirement ages on mental health: implications for public health and welfare policy. *Public Health Reviews*, vol. 32, p. 532-548.

Marshall, V.R., Taylor, P. (2005). Restructuring the life course: work and retirement. In, Johnson, M.L., Bengston, V.L., Coleman, P.G. and Kirkwood, T.B.L. (Editors), *The Cambridge Handbook of Age and Ageing*, p. 572-582. Cambridge: Cambridge University Press.

McKee-Ryan, F.M., Song, Z., Wanberg, R., Kinicki, A.J. (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of Applied Psychology*, vol. 90(1), p. 53-76.

McVittie, C., McKinlay, A., Widdicombe, S. (2008). Passive and active non-employment: age, employment and the identities of older non-working people. *Journal of Aging Studies*, 22(3), 248-255.

McVittie, C., Goodall, K. (2012). The ever-changing meanings of retirement. *American Psychologist*, vol. 67(1), p. 75-77.

Mein, G., Martikainen, P., Hemingway, H., Stansfels, S., Marmot, M. (2003). Is retirement good or bad for mental and physical health functioning? Whitehall II longitudinal study of civil servants. *Journal of Epidemiology and Community Health*, vol. 57, p. 46-49.

° Midanik, L.T., Soguikian, K., Ranson, L.J., Tekawa, I.S. (1995). The effect of retirement on mental health and health behaviours: the Kaiser Permanente Retirement Study. *Journals of Gerontology Series B, Psychological and Social Sciences*, vol. 50B(1), p. S59-S61.

Moen, P. (1996). A life course perspective on retirement, gender and wellbeing. *Journal of Occupational Health Psychology*, vol. 1(2), p. 131-144.

Moher, D., Pham, B., Lawson, M.L., Klassen, T.P. (2003). The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment*, vol. 7(41).

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6), p. e1000097.

Mojon-Azzi, S., Sousa-Poza, A., Widmer, R.W. (2007). The effect of retirement on health: a panel analysis using data from the Swiss Household Panel. *Swiss Medical Weekly*, vol. 137, p. 581-585.

Morris, S.B., DeShon, R.P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, vol. 7(1), p. 105-125.

Norris, S.L., Moher, D., Reeves, B.C., Shea, B., Loke, Y., Garner, S., Anderson, L., Tugwell, P., Wells, G. (2013). Issues relating to selective reporting when including non-randomized studies in systematic reviews on the effects of healthcare interventions. *Research Synthesis Methods*, vol. 4, p. 36-47.

° Oksannen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjosten, N., Virtanen, M., Kawachi, I., Kivimaki, M. (2011). Is retirement beneficial for mental health? Antidepressant use before and after retirement. *Epidemiology*, vol. 22(4), p.553-559.

Oldfield, Z. (2012). Economics domain tables. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), *The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5)*, p. 214-258. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

Organisation for Economic Co-operation and Development (OECD) (2011). Pensionable years, in *Society at a Glance 2011. OECD Social Indicators*. OECD Publishing. Retrieved on 31st of December 2012 from http://dx.doi.org/10.1787/soc_glance-2011-14-en

Paul, K.I., Moser, K. (2009). Unemployment impairs mental health: meta-analyses. *Journal of Vocational Behaviour*, vol. 79, p. 264-282.

Petticrew, M., Roberts, H. (2006). *Systematic Reviews in the Social Sciences: a practical guide*. Oxford: Blackwell Publishing.

Peytremann-Bridevaux, I., Voellinger, R., Santos-Eggimann, B. (2008). Healthcare and preventive services utilization of elderly Europeans with depressive symptoms. *Journal of Affective Disorders*, vol. 105, p. 247-252.

Pinquart, M., Schindler, I. (2007). Changes of life satisfaction in the transition to retirement: a latent class approach. *Psychology and Aging*, vol. 22(3), p. 442-455.

Price, C.A. (2000). Women and retirement: relinquishing professional identity. *Journal of Aging Studies*, vol. 14(1), p. 81-101.

Price, C.A., Nesteruck, O. (2010). Creating retirement paths: examples from the lives of women. *Journal of Women & Aging*, vol. 22, p. 136-149.

Quick, H. E., Moen, P. (1998). Gender, employment, and retirement quality: a life-course approach to the differential experiences of men and women. *Journal of Occupational Health Psychology*, vol. 3, p. 44-64.

Radloff, L.S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, vol. 1, p. 385-401.

Reeves, B.C., Deeks J.J., Higgins, J.P.T., Wells, G.A. (2011). Chapter 13: Including non-randomized studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Reeves, B.C., Higgins, J.P.T., Ramsay, C., Shea, B., Tugwell, P., Wells. G. (2013). An introduction to methodological issues when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 1-11.

** Reitzes, D.C., Mutran, E.J., Fernandez, M.E. (1996). Does retirement hurt wellbeing? Factors influencing self-esteem and depression among retirees and workers. *The Gerontologist*, vol. 36(5), p. 649-656.

Rosenthal, R. (1979). The file-drawer problem and tolerance for null results. *Psychological Bulletin*, vol. 86, p. 638-641.

Ross, C. E., Drentea, P. (1998). Consequences of retirement activities for distress and sense of personal control. *Journal of Health and Social Behavior*, vol. 39(4). p. 317-334.

Sahlgren, G.H. (2013). *Work longer, live healthier: the relationship between economic activity, health and government policy*. Institute of Economic Affairs (London) Discussion Paper No. 46. Retrieved on the 20th of May 2013 from http://www.iea.org.uk/sites/default/files/publications/files/Work%20Longer,%20Live_Healthier.pdf

Sanderson, S., Tatt, I.D., Higgins, J.P.T. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *International Journal of Epidemiology*, vol. 36, p. 666-676.

Sargent, L.D., Bataille, C.D., Vough, H.C., Lee, M.D. (2011). Metaphors for retirement: unshackled from schedules. *Journal of Vocational Behaviour*, vol. 79, p. 315-324.

Sargent-Cox, K.A., Anstey, K.J., Kendig, H., Skladzien, E. (2012). Determinants of retirement timing expectations in the United States and Australia: A cross-national comparison of the effects of health and retirement benefit policies on retirement timing decisions. *Journal of Aging & Social Policy*, vol. 24(3), p. 291-308.

Schalk, R., van Veldhoven, M., de Lange, A.H., de Witte, H., Kraus, K., Stamov-Roßnagel, C., Tordera, N., van der Heijden, B., Zappalà, S., Bal, M., Bertrand, F., Claes, R., Crego, A., Dorenbosch, L., de Jong, J., Desmette, D., Gellert, F.J., Hansez, I., Iller, C., Kooij, D., Kuipers, B., Linkola, P., van den Broeck, A., van der Schoot, E., Zacher, H. (2010). Moving European research on work and ageing forward: overview and agenda. *European Journal of Work and Organizational Psychology*, vol. 19(1), p. 76-101.

Schwarzbach, M., Luppa, M., Sikorski, C., Fuchs, A., Maier, W., van den Bussche, H., Pentzek, M., Riedel-Heller, S.G. (2013). The relationship between social integration and depression in non-demented primary care patients aged 75 years and older. *Journal of Affective Disorders*, vol. 145, p. 172–178.

Shamliyan, T.A., Kane, R.L., Ansari, M.T. Raman, G. Berkman, N.D., Grant, M., Janes, G., Maglione, M., Moher, D., Nasser, M., Robinson, K.A., Segal, J.B., Tsouros, S. (2011). *Development of Quality Criteria to Evaluate Nontherapeutic Studies of Incidence, Prevalence, or Risk Factors of Chronic Diseases: Pilot Study of New Checklists*. Retrieved on 7th of October 2013 from http://effectivehealthcare.ahrq.gov/ehc/products/281/617/Methods_Research_Report--Shamliyan_final_1-28-11.pdf

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: why validity issues in meta-analyses will not go away. *Clinical Psychology Review*, vol. 17(8), p. 881-901.

Shills, T. (2008). Early retirement in Germany, the Netherlands, and the United Kingdom: a longitudinal analysis of individual factors and institutional regimes. *European Sociological Review*, vol. 24(3), p. 315-329.

Shultz, K.S., Wang, M. (2011). Psychological perspectives on the changing nature of retirement. *American Psychologist*, vol. 66(3), p. 170-179.

Shünemann, H.J., Tugwel, P., Reeves, B.C., Akl, E.A., Santesso, N., Spencer, F.A., Shea, B., Wells, G., Helfand, M. (2013). Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews of on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 49-62.

Siegfried, N., Muller, M., Volmink, J., Deeks, J.J., Egger, M., Low, N., Weiss, H.H., Walker, S.A., Williamson, P.R. (2003). Male circumcision for prevention of heterosexual acquisition of HIV in men. *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD003362. DOI: 10.1002/14651858.CD003362.

Siegfried, N., Muller, M., Deeks, J., Volmink, J., Egger, M., Low, N., Walker, S., Williamson, P. (2005). HIV and male circumcision – a systematic review with assessment of the quality of studies. *Lancet Infectious Diseases*, vol. 5, p. 165-173.

Siegfried, N., Muller, M., Deeks, J.J., Volmink, J. (2009). Male circumcision for prevention of heterosexual acquisition of HIV in men. *Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD003362. DOI: 10.1002/14651858.CD003362.pub2.

Smit, F., Ederveen, A., Cuijpers, P., Deeg, D., Beekman, A. (2006). Opportunities for cost-effective prevention

of late-life depression: an epidemiological approach. *Archives of General Psychiatry*, vol. 63, p. 290-296.

Stang, A. (2008). Appropriate epidemiologic methods as a prerequisite for valid study results. *European Journal of Epidemiology*, vol 23, p. 761-765.

Steffens, D.C., Fisher, G.G., Langa, K.M., Potter, G.G., Plassman, B.L. (2009). Prevalence of depression among older Americans: the Aging, Demographics and Memory Study. *International Psychogeriatrics*, vol. 21(5), p. 879-888.

Steffick, D.E. (2000). *Documentation of affective functioning measures in the Health and Retirement Study*. HRS Health Working Group, Survey Research Center, University of Michigan. Retrieved on the 23rd of December 2013 from http://hrsonline.isr.umich.edu/sitedocs/userg/dr-005.pdf

* Steptoe, A., Demakakos, P., Oliveira, C. (2012). Chapter 4: The psychological well-being, health and functioning of older people in England. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), *The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5)*, p. 98-182. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

Sterne, J.A.C., Egger, M., Moher, D. (2011). Chapter 10: Addressing reporting biases. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A., Thacker, S.B. (2000). Meta-Analysis of Observational Studies in Epidemiology: A Proposal for Reporting. *Journal of the American Medical Association*, vol. 283(15), p. 2008-2015.

Szinovacz, M.E., DeViney, S. (1999). The retiree identity: gender and race differences. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 54B(4), p. S2O7-S218.

Szinovacz, M.E., Davey, A. (2004). Honeymoons and joint lunches: effects of retirement and spouse's employment on depressive symptoms. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 59B(5), p. 233-245.

Terrin, N., Schmid, C. H. Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, vol. 58(9), p. 894-901.

The United Kingdom Government (2013). *HM Treasury: Autumn Statement*. Retrieved on the 5th of December from https://www.gov.uk/government/news/autumn-statement-2013-date-announced-by-chancellor-as2013

Thelin, A., Holmberg, S. (2010). Farmers and retirement: A longitudinal cohort study. *Journal of Agromedicine*, vol. 15(1), p. 38-46.

Tooth, L., Ware, R., Bain, C., Purdie, D.M., Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology*, vol. 161, p. 280-288.

Topa, G., Moriano, J.A., Depolo, M., Alcover, C.M., Morales, J.F. (2009). Antecedents and consequences of retirement planning and consequences: a meta-analysis and model. *Journal of Vocational Behaviour*, vol. 75, p. 38-55.

Tugwell, P., Petticrew, M., Kristjansson, E., Welch, V., Ueffing, E., Waters, E., Bonnefoy, J., Morgan, A., Doohan, E., Kelly, M.P. (2010). Assessing equity in systematic reviews: realising the recommendations of the Commission on Social Determinants of Health. *British Medical Journal*, vol. 341, p. c4739.

Vahtera, J., Westerlund, H., Hall, M., Sjosten, N., Kivimaki, M., Salo, P., Ferie, J.E., Jokela, M., Pentti, J., Singh-Manoux, A., Goldberg, M., Zins, M. (2009). Effect of retirement on sleep disturbances: the GAZEL prospective cohort study. *Sleep*, vol. 32(11), p. 1459-1466.

Valentine, J.C., Thomson, P. (2013). Issue relating to confounding and meta-analysis when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 26-35.

Van de Velde, S., Levecque, K., Bracke, P. (2009). Measurement equivalence of the CES-D 8 in the general population in Belgium: a gender perspective. *Archives of Public Health*, vol. 67, p. 15-29.

van Solinge, H., Henkens, K. (2007). Involuntary retirement: the role of restrictive circumstances, timing and social embeddedness. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(5), p. S295-S303.

van Solinge, H., Henkens, K. (2008). Adjustment to and satisfaction with retirement: two of a kind? *Psychology and Aging*, vol. 23(2), p. 422-434.

van Solinge, H., Henkens, K. (2009). Living longer, working longer? The impact of subjective life expectancy on retirement intentions and behaviour. *European Journal of Public Health*, vol. 20(1), p. 47-51.

Villamil, E., Huppert, F.A., Melzer, D. (2006). Low prevalence of depression and anxiety is linked to statutory retirement ages rather than personal work exit: a national survey. *Psychological Medicine*, vol. 36, p. 999-1009.

Villanueva, A.B. (2000). Review Essay: Population aging and retirement issues in the next millennium. *The Social Science Journal*, vol. 37(2), p. 321-325.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines of reporting observational studies. *Epidemiology*, vol. 18(6), p. 800-804.

Wang, M. (2007). Profiling retirees in the retirement transition and adjustment process: examining the longitudinal change patterns of retirees' psychological well-being. *Journal of Applied Psychology*, vol. 92(2), p. 455-474.

Wang, M., Henkens, K., van Solinge, H. (2011). Retirement adjustment: a review of theoretical and empirical advancements. *American Psychologist*, vol. 66(3), p. 204-213.

Ware, J.E., Sherbourne, C.D. (1992). The MOS 36-Item Short-Form Health Survey (SF-36) I: Conceptual framework and item selection. *Medical Care*, vol. 30(6), p. 473-483.

Warner, D.F., Hayward. M.D., Hardy, M.A. (2010). The retirement life course in America at the dawn of the twenty-first century. *Population Research Policy Review*, vol. 29, p. 893–919.

Warr, P., Butcher, V., Robertson, I., Callinan, M. (2004). Older people's well-being as a function of employment, retirement, environmental characteristics and role preference. *British Journal of Psychology*, vol. 95, p. 297-324.

Watson, L.C., Lewis, C.L., Kistler, C.E., Amick, H.R., Boustani, M. (2004). Can we trust depression screening instruments in healthy 'old-old' adults? *International Journal of Geriatric Psychiatry*, vol. 19, p. 278-285.

Weller, C.E. (2010). Introduction: retirement security in the great recession. *Journal of Aging & Social Policy*, vol. 22, p. 95-98.

Wells, G.A., Shea, B., Higgins, J.P.T., Sterne, J., Tugwell, P., Reeves, B.C. (2013). Checklists of methodological issues for authors to consider when including non-randomized studies in systematic reviews. *Research Synthesis Methods*, vol.4, p. 63-77.

Westerlund, H., Kivimaki, M., Singh-Manoux, A., Melchior, M., Ferrie, J.E., Pentti, J., Jokela, M., Leineweber, C., Goldberg, M., Zins, M., Vahtera, J. (2009). Self-rated health before and after retirement in France (GAZEL): a cohort study. *The Lancet*, vol. 374, p. 1889-1896.

****** Westerlund, H., Vahtera, J., Ferrie, J.E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M., Zins, M., Kivimaki, M. (2010). Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. *British Medical Journal, BMJ 2010:341:c6149.*

Wilson, S.H., Walker, G.M. (1993). Unemployment and health: a review. *Public Health*, vol. 107, p. 153-162.

Wilson, K., Mottram, P.G., Vassilas, C. (2008). Psychotherapeutic treatments for older depressed people. *Cochrane Database of Systematic Reviews*, Art. No: CD004853. doi:10.1002/14651858.CD004853.pub2

World Health Organisation (1990). *Composite International Diagnostic Interview (CIDI).* Geneva: World Health Organisation.

World Health Organization (2001). *World Health Report 2001. Mental Health: New Understanding, New Hope.* Retrieved on the 3rd of September 2013 from http://www.who.int/whr/2001/en/whr01_en.pdf

Young T, Hopewell S. (2011). Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews*, Issue 11, Art. No.: MR000027. DOI: 10.1002/14651858.MR000027.pub2.

Zamarro, G., Meijer, E., Fernandes, M. (2008). Labor force participation and retirement. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 48-55. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

Zhang, J-P., Kahana, B., Kahana, E., Hu, B., Pozuelo, L. (2009). Joint modeling of longitudinal changes in depressive symptoms and mortality in a sample of community-dwelling elderly people. *Psychosomatic Medicine*, vol. 71, p. 704–714.

Zsembik, B.A., Singer, A. (1990). The Problem of defining retirement among minorities: The Mexican Americans. *The Gerontologist*, vol. 30(6), p. 749-757.

Meta-Analysis Appendices

Appendix 1

Meta-Analysis Data Search Strategy

1. Bibliographic databases and other electronic Searches – White and Grey Literature

Topic, Geo. location, white(W) grey(G)	Database	Date	Search Commands (all)
General **W**	The Cochrane Non-Randomised Studies Methods Group Specialised Register	04.01.13 Currently not available	--
W	The Cochrane Central Register of Controlled Trials (CENTRAL)	04.01.13	Retirement Ti, Ab, Kw Retirement FT Retire* Ti, Ab, Kw Work* AND (Old* OR elder* OR senior*) Ti, Ab, Kw Employ* AND (Old*OR elder* OR senior*) Ti, Ab, Kw
W	The Centre for Reviews and Dissemination databases of research in health and social care (all – DARE, NHSEED, HTA)	04.01.13	Retirement FT Work AND (Older OR Elder OR Senior) FT
W	Centre for Reviews and Dissemination publications	04.01.13	"Hand search" (digital equivalent)
G	The Campbell Collaboration Library	04.01.13	Retirement FT Old AND Work FT Employment FT
G	Open Grey	17.12.12	Retire* Retire* AND Health*
W	Science Citation Index *and* Social Sciences Citation Index *and* Conference Proceeding Citation Index	17.12.12	1 Retir* Topic 2 Retir* AND (wellbeing* OR well-being OR health OR illness* OR distress) Refine by subject (gerontology or womens studies or geriatrics gerontology or psychology social or psychiatry or primary health care or psychology multidisciplinary or psychology or nursing or clinical neurology or family studies or sociology or statistics probability or rehabilitation or social sciences biomedical or psychology experimental or psychology applied or history philosophy of science or multidisciplinary sciences or demography or substance abuse or social sciences mathematical methods or psychology clinical or behavioral sciences or social sciences interdisciplinary or social issues or ethics or social work or pharmacology pharmacy or psychology developmental) 3 Retir* AND (anxi* OR depress* OR psych* OR mental* OR "quality of life") Topic 4 S3 Narrow to (gerontology or psychology experimental or substance abuse or psychiatry or geriatrics gerontology or psychology biological or psychology multidisciplinary or social sciences mathematical methods or psychology or social work or health policy services or psychology applied or rehabilitation or psychology clinical or nursing or psychology developmental or psychology psychoanalysis or psychology social or sociology or family studies or social sciences biomedical or social sciences interdisciplinary or psychology educational or social issues or pharmacology pharmacy or demography or women s studies or multidisciplinary sciences or behavioral sciences) 5 (Work* OR Employ*) AND (Old* OR elder* OR senior*) Topic

			6 S5 Narrow by (education scientific disciplines or gerontology or geriatrics gerontology or ergonomics or psychiatry or psychology social or psychology applied or psychology educational or nursing or psychology or multidisciplinary sciences or psychology developmental or psychology experimental or women s studies or social sciences interdisciplinary or psychology multidisciplinary or rehabilitation or social issues or psychology clinical or social sciences biomedical or sociology or family studies or behavioral sciences or pathology) AND (health* OR illness* OR distress* OR wellbeing* OR well-being* OR anxi* OR depress* OR mental* OR psych*)
W	ScienceDirect	04.01.13	1 Retir* Ti, Ab, KW 2 Retir* AND Health* Ti Ab KW 3 S2 LIMIT-TO (topics, "health care,retirement,retirement study,mental health,health,european country,national health,public health,spain,alcohol consumption,early retirement,health status") 4 Wellbeing* Ti Ab KW 5 S4 LIMIT TO (topics, "mental health,social support,woman,depressive symptom,emotional wellbeing,psychological wellbeing,psycho social,subjective wellbeing,health care,life satisfaction,health") 5 Retir* AND (depress* OR Anxi*OR Illness* OR Psych* OR Mental*) FT FT 6 S5 LIMIT TO LIMIT-TO (topics, "health care,public health,lancet,hospital,patient,woman,mental health,china") 7 Retir* AND Quality of life FT FT 8 S7 AND LIMIT-TO (topics, "health care,china,patient,woman,public health,qol,mental health,lancet,hospital,india,qol score") 9 Work* AND (Old* OR elder* OR senior*) AND Health*Ti Ab KW 10 Employ* AND Health* TI Ab KW 11 Work* AND (Old* OR elder* OR senior*) Ti Ab KW 12 Work* AND (Depress*OR Anxi*) Ti Ab KW 18 Employ* AND (Depress*OR Anxi*) TI Ab KW
W	BioMed Central	14.01.13	Retire* Ti+Ab Work* AND (Old* OR elder* OR senior*) Ti+Ab Work* AND (Old* OR elder* OR senior*) AND (Health* OR wellbeing* OR well-being* OR “quality of life” OR illness* OR disease* OR distress* OR anxi* OR depress* OR mental*) Ti+Ab
G	Scopus	14.01.13	Retire* Ti+Ab+KW Work* AND (Old* OR elder* OR senior*) Ti+Ab
W	China Academic Journals	14.01.13	Search function not accessible in English
W	Chinese Science Citation Index	14.01.13	Search function not accessible in English
W	DUXIU (China)	14.01.13	Search function not accessible in English
W	Bibliography of Asian Studies Online	14.01.13	Retirement FT Retirement AND Health FT FT Retirement AND (wellbeing OR depression OR anxiety OR distress OR illness) FT FT (Work OR Worker) AND (Old OR Older OR elder OR senior) FT FT (Employ OR employment) AND (wellbeing OR depression OR anxiety OR distress OR illness OR health OR disease) FT FT
W	British Library for Development Studies	20.12.12	Retirement KW Retirement AND Health KW
Health and Psychology **W**	MEDLINE (US Library of Medicine)	27.12.12	1 Retir$ Ti 2 Retir$ NOT Hous$ NOT Dent$ NOT Communit$ Ti Ab Ab Ab 3 Retir$ AND Health$ Ab Ab 4 Retir$ AND wellbeing$ Ab Ab

			5 Retir$ AND depressi$ Ab Ab 6 Retir$ AND anxi$ Ab Ab 7 Retir$ AND Psych$ Ab Ab 8 Retir$ AND Mental$ Ab Ab 9 Map Term Retirement/ exp Retirement/[Psychology] 10 Tree [Retirement] AND [Mental Disorders] 11 Tree [Retirement] AND ['quality of life'] 12 Work$ AND (Old$ OR elder$ OR senior$) Ti Ti 13 Map Term Employment AND Employment exp 14 Tree [Employment] and [Mental disorders] 15 Tree [Employment] AND [Mental health] 16 Old$ OR elder$ OR senior$ Ab 17 S14 AND S16 18 Employ$ AND (Old$ OR elder$ OR senior$) AND Psych$ Ab Ab Ab 19 Employ$ AND (Old$ OR elder$ OR senior$) AND Health$ Ti Ab Ab
W	EMBASE	20.12.12	1 Retir$ Ti 2 Retir$ NOT Hous$ NOT Dent$ NOT Communit$ NOT Villag$ Ti Ab Ab Ab Ab 3 Retir$ AND Health$ NOT Dent$ NOT Communit$ NOT Villag$ Ab Ab Ab Ab Ab 4 Retir$ AND wellbeing$ Ab Ab 5 Retir$ AND depressi$ Ab Ab 6 Retir$ AND anxi$ Ab Ab 7 Retir$ AND Psych$ NOT Village$ Ab Ab Ab 8 Retir$ AND Mental$ Ab Ab 9 Retirement Map Term / Exp Retirement 10 psychological Map Term OR psychological aspect/ OR psychological theory/ OR "psychological and psychiatric procedures"/ OR psychological rating scale/ OR psychological model/ OR psychological wellbeing/ 11 S9 AND S10 12 Mental Health exp 13 S9 AND S13 14 "Quality of life" exp 15 S9 AND S14 16 Work$ AND Old$ Ti Ti 17 exp work/ OR exp work disability/ OR exp work resumption/ 18 old$ OR elder$ OR senior$ Ab 19 S9 AND S17 AND S18 20 Employment exp 21 Health exp OR Mental Health exp 22 S18 AND S20 AND S 21 23 exp mental disease/di, ep, et, pc, rh, si, th [Diagnosis, Epidemiology, Etiology, Prevention, Rehabilitation, Side Effect, Therapy] 24 S9 AND S23
W	CINAHL Plus	20.12.12	1 retir* exp 2. retir* exp Narrow by Subject Age (80 and over, adult: 19-44 years, middle aged: 45-64 years, aged: 65+ years, all adult) 3 Retir* Ab 4 Retir* AND (wellbeing* OR depress* OR anxi* OR Health* OR Illness* OR disease* OR mental* OR psych* OR distress* OR quality) FT 5 S4 Narrow by Subject Age (adult: 19-44 years, aged, 80 and over, middle aged: 45-64 years, aged: 65+ years, all adult) AND Narrow by Subject Major (life style, ethnic groups, sex factors, decision making, marriage, cognition, support, psychosocial, cognition disorders, mortality, attitude to health, spouses, personal satisfaction, adaptation, psychological, mental health, gerontologic care, women, psychological well-being, health behaviour, activities of daily living, physical activity, chronic disease, socioeconomic factors, functional status, Employment, quality of life, aging, depression, health

			status, retirement, 6 (Work* OR Employ*) AND (Old* OR elder* OR senior*) Ab Ab 7 S6 Narrow By Subject Major (stress, occupational, women, retirement, support, psychosocial, family, attitude to aging, memory, age factors, instrument validation, elder abuse, men, health status, depression, aging, gerontologic care) AND Subject Age(adult: 19-44 years, aged, 80 and over, middle aged: 45-64 years, aged: 65+ years, all adult)
W	PsycINFO *and* PsycArticles	17.12.12	1 exp retirement/ or exp employment status/ or job security/ or exp personnel termination/ or exp reemployment/ or exp unemployment/ 2 Wellbeing* OR depress* OR anxi* OR Health* OR Illness* OR "quality of life" OR distress OR mental* OR disease* OR psych* KW 3 S1 AND S2 4 (Old* OR elder* OR senior*) 5 S3 AND S4
W	Psychology and Behavioural Sciences Collection	17.12.12	1 Retir* FT 2 Narrow by subject (mental illness, mental health services, older people – employment, psychiatry, analysis of variance, surveys, gerontology, older people -- medical care, disability retirement, descriptive statistics, data analysis, older people -- mental health, activities of daily living, families, retirement – planning, public health, mental health, old age, geriatric psychiatry, pensions, older people – health, retirees, social services, great Britain, quality of life, mental depression, aging, caregivers, geriatrics, older people, medical care, retirement, united states, older people -- care 2 Wellbeing* OR depress* OR anxi* OR Health* OR Illness* OR "quality of life" OR distress OR mental* OR disease* OR psych* KW 3 S1 AND S2
W	Global Health *and* CAB	17.12.12	1 exp retirement/ or exp employment status/ or job security/ or exp personnel termination/ or exp reemployment/ or exp unemployment/ 2 Old* OR elder* OR senior* 3 S1 AND S2 retir* AND (Wellbeing* OR depress* OR anxi* OR Health* OR Illness* OR "quality of life" OR distress* OR mental* OR disease* OR psych*) KW
W	Global Health Library (WHO) Geographical sub-indices: Africa (AFRO) Americas (AMRO/PAHO) Eastern Mediterranean (EMRO) Europe (EURO) South-East Asia (SEARO) Western Pacific (WPRO) Latin America and Caribbean (LILACS)	17.12.12	Retirement
W	WHOLIS (World Health Organization Library and Information Networks for Knowledge Database) Global Index	17.12.12	Retirement
G	NHS Evidence – NICE	28.01.13	Retirement
G	SIGN	28.01.13	Retirement
G	King's Fund Publications	04.01.13	Retirement Retire Work

Social Sciences and Social Care **W**	ASSIA	04.01.13	1 retir* Ab 2 retir*exclude (retirement communities; retirement pensions; pensions; state retirement pensions; health insurance; savings) 3. (Work* OR Employ*) AND (Old* OR elder* OR senior*) Ab Ab 4 S3 exclude (nurses; children; nursing; social workers; hospitals; adolescents; young people; dementia; residential care; students; young children; doctors; teachers; universities; psychotherapy; preschool children; cancer; hiv) 5. S4 AND (health* OR illness* OR distress* OR wellbeing* OR distress* OR anxi* OR mental* OR psych* OR 'quality of life') 6. S5 exclude (nurses; children; nursing; hospitals; social workers; social work; adolescents; dementia; nursing homes; health professionals; quality of care; doctors; cancer; psychotherapy; young people; residential care; hiv; health insurance; students; young children; home care)
W	Sociological Abstracts *and* Social Services Abstracts	04.01.13	1 retir* Ab 2 retir* AND (health* OR illness* OR distress* OR wellbeing* OR depress* OR anxi* OR mental* OR psych* OR 'quality of life') 3. (Work* OR Employ*) AND (Old* OR elder* OR senior*) AND (health* OR illness* OR distress* OR wellbeing* OR depress* OR anxi* OR mental* OR psych* OR 'quality')
W	Social Care Online	04.01.13	Retirement OR retiree OR retired OR retiring Topic
W	IBSS (/international Bibliography of Social Sciences)		1 retir* Ab 2 S1 Narrow by Subject: studies OR retirement OR older people OR humans OR retirement planning OR male OR female OR aging OR aged OR middle aged OR elderly OR statistical analysis OR older workers OR health OR employment OR united states OR retirees OR early retirement OR personal profiles OR adult OR deaths OR demographics OR elderly people OR women OR gerontology OR age OR income OR socioeconomic factors OR risk factors OR quality of life OR baby boomers OR polls & surveys OR decision making OR usa OR aged, 80 & over OR health care OR regression analysis OR models OR statistical data OR social research OR mortality OR uk OR effects OR questionnaires OR health care expenditures OR changes OR impact analysis OR research OR comparative analysis OR public health OR mental health OR factors OR men OR families & family life 3 S2 AND (health* OR illness* OR distress* OR wellbeing* OR depress* OR anxi* OR mental* OR psych* OR 'quality of life')
G	CAIRN (French and Belgium journals)	04.01.13	Retraite Ab
G	Joseph Rowntree Foundation	04.01.13	Retirement
G	ESRC research catalogue	04.01.13	Retirement Retirement AND health
G	Social Science Research Network		Retirement Ti Ab KW
Work and Business **W**	ABI/INFORM Complete	31.12.12	1 retir* Ab 2 retir* Narrow to subject (studies OR retirement OR retirement planning OR retirement plans OR statistical analysis OR older people OR older workers OR demographics OR statistical data OR retirees OR early retirement OR employment OR trends OR humans OR aging OR baby boomers OR polls & surveys OR effects OR regression analysis OR public policy OR health insurance OR impact analysis OR impacts OR decision making OR changes OR health care expenditures OR comparative analysis OR advantages OR problems OR women OR united states OR risk management OR aged OR public sector OR male OR international OR correlation analysis OR female

			OR middle aged OR guidelines OR social conditions & trends 3. Retir* AND (wellbeing* OR well-being OR health*) Ab Ab 4. Retir* AND (illness* OR distress) Ab Ab 5. Retir* AND (anxi* OR depress* OR psych* OR mental*) Ab Ab 6 (Work* OR Employ*) AND (Old* OR elder* OR senior*) Ab Ab 7 S6 Narrow by Subject (studies OR older workers OR older people OR statistical analysis OR employment OR humans OR demographics OR female OR retirement OR male OR women OR social security OR aging OR regression analysis OR aged OR models OR unemployment OR middle aged OR trends OR statistical data OR adult OR social policy OR history OR effects OR united states OR polls & surveys OR age discrimination OR elder care OR age OR socioeconomic factors OR impact analysis OR job satisfaction OR social research OR public policy OR comparative analysis OR families & family life OR correlation analysis OR international OR employee attitude OR social conditions & trends OR public health OR changes OR perceptions OR comparative studies OR research OR motivation OR decision making OR health care OR problems OR workers OR behavior OR employees OR guidelines OR quality of life) 8. (work* OR employ*) AND (health*OR illness* OR distress*) Ab Ab 9 Old* OR elder* OR senior* 10 S8 AND S9
W	Business Source Complete	31.12.12	1 retir* Ab 2 Retir* AND (wellbeing* OR well-being* OR depress* OR anxi* OR health OR illness* OR distress* OR psych* OR mental*) FT 3 (Work* OR Employ*) AND (Old* OR elder* OR senior*) FT FT 4 S3 Narrow by Subject (regression analysis, age & employment, medical care, older workers, methodology, longitudinal) 5 (work* OR employ*) AND (Depress*OR Anxi*) FT 6 (work* OR employ*) AND (health*OR illness* OR distress*) FT FT 7 S6 Narrow by Subject (longitudinal method, performance, psychology, applied, affect (psychology), mental illness, social interaction, stress management, cognition, judgments, medicine, evaluation, well-being, personality assessment, pathological psychology, attitude (psychology), motivation (psychology), social aspects, anxiety, human behaviour, quality of life, developing countries, health, surveys, adjustment (psychology), personality, methodology, psychological aspects, public health, mental depression, counselling, distress (psychology), social psychology, emotions (psychology), mental health, stress (psychology), psychology 8 Old* OR elder* OR senior* 9 S7 AND S8
G	International Labour Organisation	31.12.12	Use for general statistics
G	OECD iLibrary	31.12.12	Use for general statistics including 'Society at a Glance'
G	TUC (UK)	31.12.12	Retirement Retire
G	Change to Win Federation	31.12.12	Retirement Retire
G	AFL-CIO	31.12.12	Retirement NARROW (Health Care) KW Retire NARROW (Health Care) KW
G	American Association of	31.12.12	'Hand search'

	Retired Persons		
G	United States Department of Labor	31.12.12	Retirement AND Health Ti Retirement AND Mental Health Ti
G	Australian Council of Trade Unions	31.12.12	Retirement Retire
G	European Union European Commission Libraries Catalogue Europeana The European Library	31.12.12	Retirement Ab Retire Ab Retirement AND Health FT Retirement FT Retirement AND Health Sb
G	US Library of Congress Online Catalogue	31.12.12	Retirement AND Health FT
G	US Office of Public Health and Science Publications	31.12.12	Retirement AND Health FT Retirement AND Health –benefits FT
G	US Centres for Disease Control and Prevention	31.12.12	Retirement AND Health FT Retirement AND Mental Health FT
G	European Centre for Disease Prevention and Control	31.12.12	Retirement FT
G	UKOP online (Official online UK publications)	31.12.12	Retirement AND Health FT
G	Department of Work and Pensions (UK)	31.12.12	Retirement Ti
G	Institute for Fiscal Studies (UK) Search under ELSA	-	-
G	Department of Health Publications and Statistics	31.12.12	Retirement Ti
G	Office for National Statistics UK	31.12.12	Retirement Ti
Theses and Dissertations	ETHOS (UK and Ireland)	18.12.12	Retirement KW Employment AND Health KW Work AND Health AND Old KW Work AND Health AND (old AND elder OR senior) KW Employment AND Illness KW
	Dissertations and Theses (Worldwide)	18.12.12	Retir* Ab
	Index to Theses (UK and Ireland)	18.12.12	Retirement Ti Employment AND Mental AND Health Ti Ti Ti Employment AND well being AND (old OR elder OR senior) Ti Ti TI Employment AND illness Ti Ti Work AND Health AND (Old OR elder OR senior) KW
	China Doctor/Master Dissertations Full-Text Database	18.12.12 Not available nine	-
	Edinburgh Research Archive	18.12.12	Retirement KW Employment KW Work AND Health KW KW Work AND Health Ab Ab (Work OR employment) AND Illness Ab AB
	DART Europe	18.12.12	Retir* KW Employ* AND Health KW KW Work AND Health KW KW KW Work AND Illness KW
	Networked Digital Library of Theses and Dissertations Note: PDFs available	18.12.12	Retirement AND Health Ab Ab Retirement AND Illness Ab Ab Employment AND Illness Ab Ab Employment AND Health AB Ab Worker AND Health Ab Ab Worker AND Illness Ab Ab

2. Published, unpublished and on-going studies (searches between 21.04.13 and 04.05.13) all Grey Literature

Location	Institution	Access	Date
Global	The World Mental Health Survey Initiative Coordinated by: WHO Harvard Medical School	Website	28.01.13
UK	English Longitudinal Study of Ageing (ELSA) Wave 6 (current) (2002-current) By: Department of Epidemiology and Public Health, UCL The Institute for Fiscal Studies The National Centre for Social Research School of Social Sciences, University of Manchester.	Website	28.01.13
	Whitehall Study I and II(UK) Coordinated by: Michael Marmott UCL Research Department of Epidemiology and Public Health	Website	28.01.13
	British Household Panel Survey Access via: Economic and Social Data Service ISER Institute for Socio Economic Reseach	Via UoE	28.01.13
	United Kingdom Household Longitudinal Study (UKHLS) or *Understanding Society* By: ESRC UK Longitudinal Studies Centre Access via: Economic and Social Data Service	Via UoE	28.01.13
	Health Survey for England		28.01.13
	1958 National Child Development Study By: Centre for Longitudinal Studies	Website	28.01.13
	Scotland and European Health for All Database (2006)?	Website (via KN)	28.01.13
Australia and NZ	New Zealand Health Work and Retirement Study Coordinated by: Massey University	Website	28.01.13
	New Zealand Longitudinal Study of Ageing (NZLSA) Coordinated by: Massey University	Website	28.01.13
	Australian National Survey of Mental Health and Wellbeing 2007 Coordinated by: Australian Bureau of Statistics	Website	28.01.13
	The Household, Income and Labour Dynamics in Australia Survey (HILDA)	Website	28.01.13
	The Melbourne Longitudinal Studies on Healthy Ageing Program (MELSHA) 1994-2010 Coordinated by: Monash University and University of Sidney	Website	28.01.13
	The Florey Adelaide Male Ageing Study (FAMAS) 2002-2005 Coordinated by: Florey Medical Research Foundation University of Adelaide Martin (Main Author) at Freemason's Foundation for Men's Health	Website	28.01.13
	Canberra Longitudinal Study 1990-2002 Coordinated by: Centre of Mental Health Research Australian National University	Website	28.01.13
Asia	Korean Longitudinal Study of Ageing (KLOSA)	Website	28.01.13

	Sister study of HRS, ESLA, SHARE, CHARLS		
	China Health and Retirement Longitudinal Study (CHARLS) Sister study of HRS, ESLA, SHARE, CHARLS	Website	28.01.13
	Beiging Longitudinal Study of Health Ageing	Find papers/ no website	28.01.13
	Longitudinal Ageing Study in India (LASI) Coordinated by: Harvard School of Public Health	Website	28.01.13
	Singapore Longitudinal Ageing Studies Coordinated by: Gerontology Research Programme at National University of Singapore	Website	28.01.13
Europe	Survey of Health Ageing and Retirement in Europe (SHARE) 4 Waves (2004-2012) Coordinated by: Axel Börsch-Supan, Ph.D. at Munich Centre for the Economics of Ageing (Max Plank Institute for Social Law and Social Policy)	Website	28.01.13
	European Community Household Panel (1994-2001) Access via: epp.eurostat.ec.europa.eu/portal/page/portal/microdata/echp	Website	28.01.13
	European Prospective Investigation into Cancer and Nutrition (EPIC) Coordinated by: Dr Elio Riboli, Head of the Division of Epidemiology, Public Health and Primary Care at the Imperial College London	Website	28.01.13
Ireland	TILDA – Irish Longitudinal Study on Ageing Coordinated by: Trinity College, University of Dublin	Website	28.01.13
France	Veillissement Sante Travail (VISAT) 1996-?? Coordinated by: L'étude VISAT a été conçue pour permettre diverses opérations de recherche, à la fois autonomes et coordonnées. Les principales équipes institutionnelles locales qui travaillent sur les données VISAT: CLLE-LTC Cognition, Langues, Langage, ErgonomieToulouse Laboratoire de Médecine du Travail, Toulouse Département d'Epidémiologie, Laboratoire d'Epidémiologie et d'Analyses en Santé Publique : Risques, Maladies Chroniques et Handicaps, INSERM, Faculté de Médecine, Toulouse Unité de Pharmacoépidémiologie, Faculté de Médecine, Toulouse	Website	28.01.13
	Enquête santé, travail et vieillissement (ESTEV) 1990-1995 Coordinated by: Société de Médecine du Travail et d'Ergonomie de l'Ouest. Groupe de médecins du travail, épidémiologistes (Inserm) et ergonomes du Centre de Recherche et d'Etude sur l'Age et la population au travail.	Find papers/ no website	28.01.13
	GAZEL Cohort (open cohort) 1989 – on-going Coordinated by (initialy): Electricité de France (EDF), Gaz de France (GDF) et l'Institut National de la Santé et de la Recherche Médicale (INSERM)	Website	28.01.13
	The Constances Cohort – an open epidemiological lab	Find papers/ no website	28.01.13
	AMI cohort – Health and Aging in Elderly Farmers	Find papers/ no website	28.01.13
Germany	German Socio-Economic Panel (GSOEP) 1984-2011 (on-going) Coordinated by:	Website	28.01.13

	German Instituted for Economic Research (DIW Berlin)		
	Berlin Ageing Study (BASE) Coordinated by: P. Baltes group, Max Plank Institute Look here for ageing related papers – successful ageing, affect in old age etc	Website	28.01.13
	Leipzig Longitudinal Study of the Aged (LEILA 75+) Epidemiological studies on depression incidence, prev., risk fators	no website	28.01.13
Italy	Italian Longitudinal Study on Aging (ILSA)	no website	28.01.13
	Italian National Research Council Targeted Project on Ageing	no website	28.01.13
Holland	Longitudinal Aging Study Amsterdam (LASA) 1992 – on-going Coordinated by: VU Amsterdam (Free University Amsterdam)	Website	28.01.13
	GLOBE Study – health inequalities	no website	28.01.13
Sweden	The Swedish National Study of Aging and Research in Kungsholmen (SNACK)	Website	28.01.13
	The Stockholm Birth Cohort	Website	28.01.13
	The Stockholm Public Health Cohort	no website	28.01.13
Finland	The Helsinki Ageing Study	no website	28.01.13
	TURVA project – adjustment to retirement	no website	28.01.13
USA			28.01.13
Institutes an Research Centres	US National Institute on Aging (US)	Website	28.01.13
	14 NIA Demography of Aging Centres (PI, name, affiliation): 1. Agree, Emily M. Hopkins Center for Population Aging and Health Johns Hopkins University 2. Bloom, David E. Center for Global Demography of Aging Harvard University 3. Bound, John Center on the Demography of Aging University of Michigan Population Studies Center - ISR 4. Crimmins, Eileen Center on Biodemography and Population Health University of Southern California and University of California at Los Angeles Andrus Gerontology Center 5. Garber, Alan M. Center on the Economics and Demography for Health and Aging Stanford University Department of Medicine/PCOR 6. Hauser, Robert M. Center for Demography of Health and Aging University of Wisconsin - Madison 7. Hurd, Michael D. Center for the Study of Aging RAND Corporation 8. Lee, Ronald D. Center on the Economics and Demography of Aging (CEDA) University of California at Berkeley 9. Paxson, Christina H. Center for the Demography of Aging Princeton University 10. Soldo, Beth J. Center on the Demography of Aging University of Pennsylvania Populations Studies Center 11. Vaupel, James W. Duke Center on the Demography of Aging Duke University		28.01.13

	Public Policy- DuPRI SSRI 12. Waite, Linda J. Center on the Demography and Economics of Aging University of Chicago & National Opinion Research Center (NORC) 13. Wise, David A. Center for Aging and Health Research National Bureau of Economic Research 14. Wolf, Douglas A. Center for Aging and Policy Studies Syracuse University Center for Policy Research		
	Sloan Centre on Aging and Work Boston College	Website	28.01.13
	The Job Stress Network	Website	28.01.13
Studies	US Health and Retirement Study	Website	28.01.13
	Baltimore Longitudinal Study of Aging (1958-current) Coordinated By: National Institute on Aging	Website	28.01.13
	Georgia Centenarian Study Coordinated by: Institute of Gerontology, University of Georgia	Website	28.01.13
	The Seattle Longitudinal Study 1956-2005 Coordinated by: University of Washington	Website	28.01.13
	Cornell Retirement and Wellbeing Study 1952 – ?? Coordinated by: Cornell University	no website	28.01.13
	Detroit Area Studies 1951 – ??	no website	28.01.13
	The Kaiser-Permanente Retirement Study	no website	28.01.13
	Normative Aging Study Based at VA Outpatient Clinic in Boston, Mass Coordinated by the United States Department of Veteran Affairs	no website	28.01.13
Canada	Canadian Study of Health and Aging	Website	28.01.13
	Fredericton 80+ Study	Website	28.01.13
	The Victoria Longitudinal Study Coordinated by: The University of Alberta	Website	28.01.13
Central and South America	Mexican Health and Aging Study (MHAS) 2001-2005 (planned surveys 2012 and 2014) Coordinated by: Universities of Pennsylvania, Maryland, and Wisconsin and the Instituto Nacional de Estadística, Geografia e Informática (INEGI), Mexico	Website	28.01.13
	Puerto Rican Elderly: Health Conditions (PREHCO) Project Coordinated by: Universities of Puerto Rico, Wisconsin and North-eastern University Affiliated to National Institute on Aging US	Website	28.01.13
	Health, Wellbeing and Ageing in Latin America and the Caribbean (SABE) Coordinated by: Pan American Health Organisation and National Institute on Aging US	Website	28.01.13

Appendix 2

Email Correspondence with Authors

1. Template Emails

1.1 Template email A

Dear [ ]

I am writing to you as the corresponding author for the flowing paper: [Insert title]

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfillment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I am attempting to find any studies or data (published or unpublished), which may be related to the focus of my meta-analysis. In the literature search, I came across the published abstract named above. I am wondering if it would be possible for you to clarify if this abstract corresponds to empirical research that you have now finalized. If this is the case, I am also wondering if it would be at all possible for me to access a full report of this study, with a view to including it in my meta-analysis.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Yours Sincerely,

Claudia Coelho
Specialist Clinical Psychology Practitioner,

Edinburgh Clinical Psychology Services for Older People
National Health Service (NHS) Lothian/University of Edinburgh

1.2. Template email B

Dear [ ]

I am writing to you as the first author for the flowing report: [Insert title]

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfillment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I am attempting to find any studies or data (published or unpublished), which may be related to the focus of my meta-analysis. In the literature search, I came across the title of the report named above. I am wondering if it would be possible for you to clarify if this report corresponds to empirical research that you have now finalized. If this is the case, I am also wondering if it would be at all possible for me to access the report, with a view to including it in my meta-analysis.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Yours Sincerely,

Claudia Coelho
Specialist Clinical Psychology Practitioner,
Edinburgh Clinical Psychology Services for Older People
National Health Service (NHS) Lothian/University of Edinburgh

1.3 Template email C

Dear [ ]

I am writing to you as the main/corresponding author for the flowing paper: [Insert title]

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible for you to provide, in relation to the sample data used in your paper. I outline here the additional information that is needed.

[insert necessary data]

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Yours Sincerely,

Claudia Coelho
Specialist Clinical Psychology Practitioner,
Edinburgh Clinical Psychology Services for Older People
National Health Service (NHS) Lothian/University of Edinburgh

2. Communication with authors regarding other data sources

Table 1 – Other Data Sources – email communication with authors and outcomes

Email communication	Date sent	Response?	Outcome
Author: Silver, M. Title: How Different Definitions of Retirement have Different Implications for Depression Periodical: Gerontologist, 2008, vol. 48, p. 352 Type of document: Meeting Abstract Email: mpsilver@utsc.utoronto.ca	02.09.13	Yes	No new data
Authors: Sugisawa,H.; Sugihara,Y.; Harada,K.; Shibata,H.; Hougham,G. Title: Effects of mandatory retirement on the mental health and social well-being of Japanese men Periodical: Gerontologist, 2003, vol. 43, p. 344 Type of document: Meeting Abstract Email: sugihara@tmig.or.jp	02.09.13	No	-
Authors: Lei, Xiaoyan, Li Tan, Yaohui Zhao Title: The impact of retirement on health, evidence from China, 2011 Type of Document: Unpublished Manuscript. China Center for Economic Research, Peking University, Peking. Email: xylei@ccer.pku.edu.cn	02.09.13	Yes	New data only in Chinese

Dear Dr Silver,

I am writing to you as the author of the flowing publication:
Title: How Different Definitions of Retirement have Different Implications for Depression
Journal: Gerontologist, 2008, vol. 48, p. 352
Type of document: Meeting Abstract

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfillment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I am attempting to find any studies or data (published or unpublished), which may be related to the focus of my meta-analysis. In the literature search, I came across the published abstract named above. I am wondering if it would be possible for you to clarify if this abstract corresponds to empirical research that you have now finalized. If this is the case, I am also wondering if it would be at all possible for me to access a full report of this study, with a view to including it in my meta-analysis.

Thank you so much for taking the time to read this request. I appreciate any help you may be able to provide.

Dear Dr Sugihara,

I am writing to you as the corresponding author for the flowing publication:
Authors: Sugisawa, H.; Sugihara,Y.; Harada,K.; Shibata,H.; Hougham,G.
Title: Effects of mandatory retirement on the mental health and social well-being of Japanese men

Journal: Gerontologist, 2003, vol. 43, p. 344
Type of document: Meeting Abstract

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfillment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I am attempting to find any studies or data (published or unpublished), which may be related to the focus of my meta-analysis. In the literature search, I came across the published abstract named above. I am wondering if it would be possible for you to clarify if this abstract corresponds to empirical research that you have now finalized. If this is the case, I am also wondering if it would be at all possible for me to access a full report of this study, with a view to including it in my meta-analysis.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Dr Lei,

I am writing to you as the first author for the flowing report:
Authors: Lei, Xiaoyan, Li Tan, Yaohui Zhao
Title: The impact of retirement on health, evidence from China, 2011
Type of Document: Unpublished Manuscript. China Center for Economic Research, Peking University, Peking.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfillment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I am attempting to find any studies or data (published or unpublished), which may be related to the focus of my meta-analysis. In the literature search, I came across the title of the report named above. I am wondering if it would be possible for you to clarify if this report corresponds to empirical research that you have now finalized. If this is the case, I am also wondering if it would be at all possible for me to access the report, with a view to include it in my meta-analysis.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

3. Communication with authors regarding additional statistical data

Table 2 – Additional statistical data – email communication with authors and outcomes

Email communication	Date sent	Reply?	Outcome
Cross-sectional studies			
Drentea, P. (2002). Retirement and mental health. Journal of Aging and Health, vol. 14, p. 167-194. Email: pdrentea@uab.edu	19.11.13	22.11.13	Requested data provided
Jang, S-N., Cho, S-I., Chang, J., Boo, K., Shin, H-G., Lee, H., Berkman, L.F. (2009). Employment status and depressive symptoms in Koreans: results from a baseline survey of the Korean longitudinal study of aging. Journal of Gerontology Series B: Psychological Sciences and Social Sciences, vol. 64(5), p. 677-683. Email: sjang@cau.ac.kr	19.11.13	No	-
Steptoe, A., Demakakos, P., Oliveira, C. (2012). Chapter 4: The psychological well-being, health and functioning of older people in England. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5), p. 98-182. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5 Email: a.steptoe@ucl.ac.uk	19.11.13	05.12.13	Requested data provided
Cohort studies			
Brugiavini, A., Croda, E., Dewey, M. (2008). Retirement and mental health. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension, p. 247-254. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf Email: brugiavi@unive.it	19.11.13	No	-
Westerlund, H., Vahtera, J., Ferrie, J.E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M., Zins, M., Kivimaki, M. (2010). Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. British Medical Journal, BMJ 2010:341:c6149. Email: hugo.westerlund@stressforskning.su.se	19.11.13	No	-
Midanik, L.T., Soguikian, K., Ranson, L.J., Tekawa, I.S. (1995). The effect of retirement on mental health and health behaviours: the Kaiser Permanente Retirement Study. Journals of Gerontology Series B, Psychological and Social Sciences, vol. 50B(1), p. S59-S61. Email: lmidanik@berkeley.edu	19.11.13	21.11.13	No further data – author retired and dataset not available
Kim, J.E., Moen, P. (2002). Retirement transitions, gender and psychological wellbeing: a life course ecological model. Journals of Gerontology Series B: Psychological Science and Social Sciences, vol. 57B(3), p. 212-222. Email: jungmeen@vt.edu	19.11.13	No	-
Laaksonen, M., Metsä-Simola, N., Martikainen, P., Pietiläinen, O., Rahkonen, O., Gould, R., Partonen, T., Lahelma, E. (2012). Trajectories of mental health before and after old-age and disability retirement: a register-based study on purchases of psychotropic drugs. Scandinavian Journal of Work and Environmental Health, vol. 38(5), p. 409-417. Email: mikko.laaksonen@etk.fi	19.11.13	No	-
Oksannen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjosten, N., Virtanen, M., Kawachi, I., Kivimaki, M. (2011). Is retirement beneficial for mental health? Antidepressant use before and after retirement. Epidemiology, vol. 22(4), p.553-559. Email: tuula.oksanen@ttl.fi	19.11.13	No	-
Lindeboom, M., Portrait, F., van den Berg, G.J. (2002). An econometric analysis of the mental-health effects of major events in the life of older individuals. Health Economics, vol. 11, p. 505-520. Email: m.lindeboom@vu.nl	19.11.13	20.11.3	No further data – author does not have access to data for further analysis
Clarke, P., Marshall, V., House, J., Lantz, P. (2011). The social structuring of mental health over the adult life course: advancing theory in the sociology of aging. Social Forces, vol. 89(4), p. 1287-1313. Email: pjclarke@umich.edu	19.11.13	20.11.13	Data provided, but still not clear; measure of depression considered unsafe

3.1 Cross-sectional studies sample

Dear Dr Drentea,

I am writing to you as the author of the flowing paper:
Drentea, P. (2002). Retirement and mental health. *Journal of Aging and Health*, vol. 14, p. 167-194.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible for you to provide, in relation to the sample data used in your paper. I realise that some time has elapsed between the paper's publication and now. Nevertheless, I outline below the additional information that, if possible, would be incredibly helpful to obtain.

Your paper reports on data from two surveys: Aging, Status and Sense of Control Survey and National Survey of Families and Households. Because of the inclusion criteria of my meta-analysis, I am only including data related to Aging, Status and Sense of Control Survey. Within that, I am also only interested in the data pertaining to the CES-D scores, as the dependent variable.

In your original paper, in page 179, you state: 'Difference-of-means t-tests indicated a significant difference
between the retired and non-retired with all of the mental health outcomes in ASOC (…). Retirement was associated with less distress, lower depressive symptomatology, lower anxiety, lower sense of control (in ASOC only), and a higher level of positive affect'. However, the descriptive and inferential statistical details of the relationship are not included in your paper.

I was therefore wondering if it would be at all possible for you to provide data related to the following categories, in relation to the ASOC survey only:
1. Retired group sample size
2. Retired sample mean score on the CES-D only
3. Retired sample standard deviation of the mean CES-D score
4. Non-retired group sample size
5. Non-Retired sample mean score on the CES-D only
6. Non-retired sample standard deviation of the mean CES-D score
7. t-test value and exact (if possible) probability value

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Dr Jang,

I am writing to you as the main and author of the flowing paper:
Jang, S-N., Cho, S-I., Chang, J., Boo, K., Shin, H-G., Lee, H., Berkman, L.F. (2009). Employment status and depressive symptoms in Koreans: results from a baseline survey of the Korean longitudinal

study of aging. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 64(5), p. 677-683.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible for you to provide. I realise that some time has elapsed between the paper's publication and now. Nevertheless, I outline below the additional information that, if possible, would be incredibly helpful to obtain.

Your paper reports on data from the baseline survey of the Korean Longitudinal Study of Aging, in relation to CES-D scores and employment status. I am specifically interested in this relationship. I was, therefore, wondering if it would be at all possible for you to provide the descriptive and inferential statistical details of this relationship that are not included in your paper:

1. Combined retired and employed sub-samples age mean
2. Combined retired and employed sub-samples age standard deviation
3. Retired sub-sample age mean
4. Retired sub-sample age standard deviation
5. Employed sub-sample age mean
6. Employed sub-sample age standard deviation
7. Retired sub-sample size
8. Retired sub-sample mean score on the CES-D
9. Retired sub-sample standard deviation of the mean CES-D score
10. Employed sub-sample size
11. Employed sub-sample mean score on the CES-D only
12. Employed sub-sample standard deviation of the mean CES-D score

If you performed any kind of inferential difference test (t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Prof Steptoe,

I am writing to you as the main author of the flowing paper:
Steptoe, A., Demakakos, P., Oliveira, C. (2012). Chapter 4: The psychological well-being, health and functioning of older people in England. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5), p. 98-182. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. I realise that some time has elapsed between the overall report's publication and now.

Your paper includes data from the 5th Wave of the English Longitudinal Study of Aging. I am specifically focusing on this aspect of your analysis, namely, the cross-sectional data that you provide of the relationship between CES-D scores (as indicator of presence of depressive symptoms) and employment status on the 5th Wave data. You provide some information on this relationship in Table A4.16 of the report (p. 168). In this, you dichotomised the CES-D variable (scores <4 'not-depressed', and >4 'depressed'), and presented proportions related to each category of employment status (working vs not-working).

As I am particularly interested in this relationship, I was wondering if it would be at all possible to obtain further descriptive and inferential statistical details of the data, using the CES-D scores as a continuous variable. I outline here the additional information that, if possible, would be incredibly helpful to have:

1. Combined working and not working sub-samples age mean
2. Combined working and not working sub-samples age standard deviation
3. Not-working sub-sample age mean
4. Not-working sub-sample age standard deviation
5. Working sub-sample age mean
6. Working sub-sample age standard deviation
7. Not-working sub-sample size (in the analysis of CES-D scores)
8. Not-working sub-sample mean score on the CES-D
9. Not-working sub-sample standard deviation of the mean CES-D score
10. Working sub-sample size (in the analysis of CES-D scores)
11. Working sub-sample mean score on the CES-D only
12. Working sub-sample standard deviation of the mean CES-D score

If you performed any kind of inferential difference test, either on the dichotomised variable or continuous variable values (chi square, t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

3.2 Cohort studies sample

Dear Prof Brugiavini,

I am writing to you as the main author of the flowing paper:
Brugiavini, A., Croda, E., Dewey, M. (2008). Retirement and mental health. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 247-254. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. However, I do realise that some time has elapsed between the overall report's publication and now.

Your paper includes data from the first two Waves of the SHARE. I am specifically focusing on the longitudinal aspect of your analysis, namely, data on the relationship between EURO-D scores (as indicator of presence of depressive symptoms) and changes in employment status. You provide some information on this relationship in Table 2 in the report (p. 252). In this, you dichotomised the EURO-D variable (scores <3 'not-depressed', and >3 'depressed'), and presented proportions related to each category of employment status (retired vs employed).

As I am particularly interested in this relationship, I was wondering if it would be at all possible to obtain further descriptive and inferential statistical details of the data, using the EURO-D as a continuous variable. I outline here the additional information that, if possible, would be incredibly helpful to have:

1. Sample age mean in Wave 1 and Wave 2 (of the sample used for the analysis of EURO-D scores)
2. Sample age mean standard deviation at Wave 1 and Wave 2 (of the sample used for the analysis of EURO-D scores)
3. Employed sample size in Wave 1 (in the analysis of EURO-D scores)
4. Retired sample size in Wave 2 of those who were employed in Wave 1 (in the analysis of EURO-D score)
5. Employed sample mean score on the EURO-D in Wave 1
6. Employed sample standard deviation of the mean EURO-D score in Wave 1
7. Retired sample mean score on the EURO-D in Wave 2, of those who were employed in Wave 1
6. Retired sample standard deviation of the mean EURO-D score in Wave 2, of those who were employed in Wave 1

If you performed any kind of inferential difference test in relation to the EURO-D scores, either on the dichotomised variable or continuous variable values (chi square, t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Finally, in page 247, you mentioned 'We construct a binary indicator which takes value one if the EURO-D scale is above three and zero otherwise, which has been demonstrated to indicate a clinically significant level of depression'. I was wondering if you could advise me on where in the EURO-D literature I might find test-retest reliability estimates for the binary use of the scale.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Prof Westerlund,

I am writing to you as the main author of the flowing paper:
Westerlund, H., Vahtera, J., Ferrie, J.E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M., Zins, M., Kivimaki, M. (2010). Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. British Medical Journal, BMJ 2010:341:c6149.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. However, I do realise that some time has elapsed between the overall report's publication and now.

Your paper includes data from the GAZEL cohort. I am specifically focusing on longitudinal changes of the relationship between CES-D scores (as indicator of presence of depressive symptoms) and employment status. You provide information on this relationship in Table B (web-based tables) of your paper, using odd-ratios. You present data on the main effect of time ($p<0.0001$) on the prevalence of depressive symptoms for the following time intervals: year -1 vs -7; year +1 vs -1; year +7 vs +1.

I was wondering if it would be at all possible to obtain the same information, related to the same time intervals, but in terms of means and standard deviations of the CES-D scores at each interval, and specific sample sizes for each time interval.

I was also wondering if you would perhaps have correlation values for the CES-D scores between the different points/intervals. This would be incredibly helpful, if at all possible.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Prof Midanik,

I am writing to you as the main author of the flowing paper:
Midanik, L.T., Soguikian, K., Ranson, L.J., Tekawa, I.S. (1995). The effect of retirement on mental health and health behaviours: the Kaiser Permanente Retirement Study. Journals of Gerontology Series B, Psychological and Social Sciences, vol. 50B(1), p. S59-S61.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. I do, however, realise that a long period of time has elapsed between the paper's publication and now, and that you may no longer have this information to hand.

Your paper includes data from the Kaiser Permanente cohort. I am specifically interested in the longitudinal changes of the relationship between CES-D scores (as indicator of presence of depressive symptoms) and employment status. You provide information on this relationship in Table 1 of your paper (p. S60). In this, you dichotomised the CES-D variable (scores <16 'not-depressed', and >16 'depressed'), and presented relative risk scores for retired vs not retired at Follow-up.

I am particularly interested in this relationship, and was wondering if it would be at all possible to obtain further descriptive and inferential statistical details of the data, using the CES-D as a continuous variable. I outline here the additional information that, if possible, would be incredibly helpful to have:

1. Sample age mean at Time 1 and Follow-up (of the sample used in the analysis of CES-D scores)
2. Sample age mean standard deviation at Time 1 and Follow-up (of the sample used in the analysis of EURO-D scores)
3. Employed sample size at Time 1 (in the analysis of CES-D scores)
4. Retired sample size at Follow up of those who were employed at Time 1 (in the analysis of CES-D scores)
5. Employed sample mean score on the CES-D at Time 1
6. Employed sample standard deviation of the mean CES-D score at Time 1
7. Retired sample mean score on the CES-D at Follow-up, of those who were employed at Time 1 only
6. Retired sample standard deviation of the mean CES-D score at Follow-up, of those who were employed at Time 1.

If you performed any kind of inferential difference test in relation to the CES-D scores, either on the dichotomised variable or continuous variable values (chi square, t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Prof Kim-Spoon,

I am writing to you as the main author of the flowing paper:
Kim, J.E., Moen, P. (2002). Retirement transitions, gender and psychological wellbeing: a life course ecological model. Journals of Gerontology Series B: Psychological Science and Social Sciences, vol. 57B(3), p. 212-222.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. I do realise, however, that a considerable amount of time has elapsed between the paper's publication and now, and that you may no longer have this information to hand.

Your paper includes data from the Cornell Retirement and Wellbeing Study. I am specifically focusing on longitudinal changes of the relationship between CES-D scores (as indicator of presence of depressive symptoms) and employment status, between measurements at Time 1 and Time 2. You provide descriptive information on this relationship in Table 1 (p. 216). However, I was wondering if it was at all possible to obtain the following specific information for the sample of 80 individuals who make the transition from employment to retirement between Time 1 and Time 2 only:

1. Employed sample mean score on the CES-D at Time 1 (men and women)
2. Employed sample standard deviation of the mean CES-D score at Time 1 (men and women)
3. Retired sample mean score on the CES-D at Time 2 of those who were employed at Time 1 only
4. Retired sample standard deviation of the mean CES-D score at Time 2, of those who were employed at Time

If you performed any kind of inferential difference test in relation to the CES-D scores between the two measurement times (t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Dr Laaksonen,

I am writing to you as the main author of the flowing paper:
Laaksonen, M., Metsä-Simola, N., Martikainen, P., Pietiläinen, O., Rahkonen, O., Gould, R., Partonen, T., Lahelma, E. (2012).Trajectories of mental health before and after old-age and disability retirement: a register-based study on purchases of psychotropic drugs. Scandinavian Journal of Work and Environmental Health, vol. 38(5), p. 409-417.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of

the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. I do, however, realise that some time has elapsed between the paper's publication and now, and that you may not have this information to hand.

Your paper includes data from your study of a cohort of City of Helsinki employees. I am specifically interested in the longitudinal changes of the relationship between your measure of purchase of psychotropic drugs and changes in employment status. Particularly, I would like to focus on this relationship in your sub-sample of 'old-age retirees', and the specific variable of purchase of anti-depressants.

I was therefore wondering if it would be at all possible to obtain further statistical details on the data for the 'old-age retirees' only (N=4456), and the purchase of anti-depressants as the dependent variable. The three time-intervals that you use in Table 1 in your paper (p. 413) are very useful – 5-1.5 years before retirement; 1.5-0 years before retirement; and 0-5 years after retirement. The data I would be looking to obtain are the following:
1. Sample size (of old age retirees only), mean and standard deviation of purchase of Daily Defined Doses of anti-depressants for Interval 1 (5-1.5 years before retirement)
2. Sample size (of old age retirees only), mean and standard deviation of purchase of Daily Defined Doses of anti-depressants for Interval 2 (1.5-0 years before retirement)
3. Sample size (of old age retirees only), mean and standard deviation of purchase of Daily Defined Doses of anti-depressants for Interval 3 (0-5-years after retirement)

If you performed any kind of inferential difference test in relation to these specific values over time (t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Finally, I was also wondering if you would perhaps have correlation value for your measure (purchase of anti-depressants) between the three time points. This would be incredibly helpful, if at all possible.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Dr Oksanen,

I am writing to you as the main author of the flowing paper:
Oksannen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjosten, N., Virtanen, M., Kawachi, I., Kivimaki, M. (2011). Is retirement beneficial for mental health? Antidepressant use before and after retirement. Epidemiology, vol. 22(4), p.553-559.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. I do, however, realise that some time has elapsed between the paper's publication and now, and that you may not have this information to hand.

Your paper analyses data of the Finish Public Sector Study Cohort. I am specifically interested in the longitudinal changes of the relationship between your measure of purchase of antidepressant drugs and changes in employment status. Particularly, I would like to focus on this relationship in the sub-sample of statutory age retirees (N= 7138) in your study (as opposed to the sub-sample of early retirees on health grounds).

I was therefore wondering if it would be at all possible to obtain further statistical details on the longitudinal changes for statutory age retirees and the purchase of anti-depressants. The time-points that you use in your e-Table 3 are really useful, i.e.: Pre-retirement period (years -4 to -2), Transition period (years -1 to +1), Post-retirement period (years +2 to +4). The data I would be looking to obtain are the following:
1. Sample size (of total statutory age retirees), mean and standard deviation of purchase of Daily Defined Doses of anti-depressants for pre-retirement period;
2. Sample size (of total statutory age retirees), mean and standard deviation of purchase of Daily Defined Doses of anti-depressants for transition period;
3. Sample size (of total statutory age retirees), mean and standard deviation of purchase of Daily Defined Doses of anti-depressants for post-retirement period.

If you performed any kind of inferential difference test in relation to these specific values over time (t-test, F-ratio, etc.), I was wondering if it would it be at all possible to also include this value and associated exact (if available) probability value.

Finally, I was also wondering if you would perhaps have correlation value for your measure (purchase of anti-depressants) between the three time intervals above or any of the time intervals you may have used. This would be incredibly helpful, if at all possible.

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Prof Lindeboom,

I am writing to you as the main author of the flowing paper:
Lindeboom, M., Portrait, F., van den Berg, G.J. (2002). An econometric analysis of the mental-health effects of major events in the life of older individuals. Health Economics, vol. 11, p. 505-520.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is some additional statistical information that I would be very grateful if it would be at all possible to obtain. I do, however, realise that a considerable amount of time has elapsed between the paper's publication and now, and that you may not have this information to hand.

Your paper analyses data from three Waves of the Longitudinal Aging Study Amsterdam (LASA). I am specifically interested in the longitudinal changes of the relationship between CES-D scores (as indicators of presence of depressive symptoms) and changes in employment status. In table 4 in your paper (p. 512), you include the frequency of complete transitions from 'employed' to retired ('new 'early' pensioners') in between 92/93 (Wave 1) and 95/96 (Wave 2) as 3.1%, and between 95/96 (Wave 2) and 98/99 (Wave 3) as 1.3% of 'the relevant group' (p. 511).

I was, therefore, wondering if it would be at all possible to obtain further statistical details on the longitudinal changes for these participants only, who make the transition from employed to retired, in relation to their CES-D scores (as a continuous variable, i.e. not dichotomised). I outline here the additional information that, if possible, would be incredibly helpful to have:

1. Employed sample size in Wave 1 (in the analysis of CES-D scores)
2. Retired sample size in Wave 2 of those who were employed in Wave 1 (in the analysis of CES-D scores)
3. Employed sample size in Wave 2 (in the analysis of CES-D scores)
4. Retired sample size in Wave 3 of those who were employed in Wave 2 (in the analysis of CES-D scores)
5. Employed sample mean score on the CES-D in Wave 1
6. Employed sample standard deviation of the mean CES-D score in Wave 1
7. Retired sample mean score on the CES-D in Wave 2, of those who were employed in Wave 1
8. Retired sample standard deviation of the mean CES-D score in Wave 2, of those who were employed in Wave 1
9. Employed sample mean score on the CES-D in Wave 2
6. Employed sample standard deviation of the mean CES-D score in Wave 2
7. Retired sample mean score on the CES-D in Wave 3, of those who were employed in Wave 2
8. Retired sample standard deviation of the mean CES-D score in Wave 3, of those who were employed in Wave 2

If you performed any kind of inferential difference test in relation to the CES-D scores, either on the dichotomised variable or continuous variable values (chi square, t-test, F-ratio, etc.), would it be at all possible to also include this value and associated exact (if available) probability value?

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Dear Dr Clarke,

I am writing to you as the main author of the flowing paper:
Clarke, P., Marshall, V., House, J., Lantz, P. (2011). The social structuring of mental health over the adult life course: advancing theory in the sociology of aging. Social Forces, vol. 89(4), p. 1287-1313.

I am currently undertaking a meta-analysis of studies that explore the relationship between retirement from the work force and the experience of depression or symptoms of depression, in older individuals. Specifically, I am interested in: a) the quantification of this relationship, as it has been examined with variable results in a large number of studies; and b) which factors (methodological characteristics of the study and characteristics of the sample) may moderate this relationship and account for some of the variability found in this area of research.

The meta-analysis is being carried out as doctoral work, in part-fulfilment of a Doctorate in Clinical Psychology, School of Health in Social Science, University of Edinburgh, United Kingdom (http://www.ed.ac.uk/schools-departments/health/clinical-psychology/studying), supervised by Dr Emily Newman (emily.newman@ed.ac.uk).

I have included your paper above in the sample of studies for this meta-analysis. However, there is additional statistical information that I would be very grateful if it would be at all possible to obtain. I do, however, realise that some time has elapsed between the paper's publication and now, and that you may not have this information to hand.

Your paper analyses data from four measurement time-points of the Americans' Changing Lives Survey. I am specifically interested in the longitudinal changes of the relationship between CES-D scores (as indicators of presence of depressive symptoms) and changes in employment status. I was, therefore, wondering if it would be at all possible to obtain further statistical details on these participants who make the transition from employed to retired during the four measurement time points of the study, in relation to changes in their CES-D scores, as a continuous variable, using raw (un-transformed) scores, if possible. I outline here the additional information that, if at all available, would be incredibly helpful to have:

1. Pre-retirement sample size, of those that make the transition to retirement (used in the analysis of CES-D scores)
2. Pre-retirement mean and standard deviation of CES-D scores, of those that make the transition to retirement
3. Post-retirement sample size, of those that made the transition to retirement (used in the analysis of CES-D scores)
4. Post-retirement mean and standard deviation of CES-D scores, of those that make the transition to retirement

If you performed any kind of inferential difference test in relation to the CES-D scores (t-test, F-ratio, etc.), would it be at all possible to also include this value and associated exact (if available) probability value?

Thank you so much for taking the time to read this request. I very much appreciate any help you may be able to provide.

Appendix 3

Data Selection Process (Sequential Sifts 1-6)

Table 1 – Electronic Searches (White and Grey Literature) Screening Sifts 1 to 4

Total Yield **Records identified from all searches**	**Sift 1** **Saved unique records from all sources after de-duplication**	**Sift 2** **Saved records after abstract inspection**	**Sift 3** **Saved records after abstract (brief full-text inspection when necessary)**	**Sift 4** **Saved records after abstract (brief full-text inspection when necessary)**
Total: 99355	Total: 9191 unique records (9.25% total records identified)	Total: 1121 unique records (1.12% of total records identified)	Total: 302 (0.30% of total records identified)	Total: 87 (0.087 of total records identified)
White Literature: 94013 Grey Literature: 5342 (5.4% of total records identified)	White Literature: 9060 Grey Literature: 131 (1.4% of Sift 1 total unique records)	White Literature: 1086 Grey Literature: 35 (3% of Sift 2 total unique records)	White literature: 177 Grey Literature: 29 + Hand Searches: 75 Large Studies (websites and other records): 21	--

Table 2 – Sift 5 – Excluded reports with reasons for exclusion noted, grouped by type of reason for exclusion

Note: reports are organised in the table by type of reason for exclusion, i.e. related to the relevant inclusion criteria, not alphabetically.

Reference	Reason(s) for exclusion
George, L.K., Fillenbaum, G.G., Palmore, E. (1984). Sex differences in the antecedents and consequences of retirement. *Journal of Gerontology*, vol. 39(3), p. 364-371.	Depression or depressive symptoms not identified as outcome variable – Affect Balance Scale used as a measure of 'subjective well-being'.
Palmore, E., Cleveland, W.P., Nowlin, J.B., Ramm, D., Siegler, I.C. (1979). Stress and adaptation in later life. *Journal of Gerontology*, vol. 34(6), p. 841-851.	Depression or depressive symptoms not identified as outcome variable – Affect Balance Scale used as a measure of 'life satisfaction'; African-Americans excluded from sampling.
Vahtera, J., Westerlund, H., Hall, M., Sjosten, N., Kivimaki, M., Salo, P., Ferie, J.E., Jokela, M., Pentti, J., Singh-Manoux, A., Goldberg, M., Zins, M. (2009). Effect of retirement on sleep disturbances: the GAZEL prospective cohort study. *Sleep,* vol. 32(11), p. 1459-1466.	Depression or depressive symptoms not identified as outcome variable
Saïas, T., Beck, F., Bodard, J., Guignard, R., du Roscoät, E. (2012). Social participation, social environment and death ideations in later life. *PLOS One*, vol. 7(10), p. e46u723.	Depression or depressive symptoms not identified as outcome variable.
Westerlund, H., Kivimaki, M., Singh-Manoux, A., Melchior, M., Ferrie, J.E., Pentti, J., Jokela, M., Leineweber, C., Goldberg, M., Zins, M., Vahtera, J. (2009). Self-rated health before and after retirement in France (GAZEL): a cohort study. *The Lancet*, vol. 374, p. 1889-1896.	No clear measure of depression or depressive symptoms; Depression or depressive symptoms not identified as outcome variable.
Zhan, Y., Wang, M., Liu, S., Shultz, K. (2009). Bridge employment and retirees' health: a longitudinal investigation. *Journal of Occupational Health Psychology*, vol. 14(4), p. 474-389	No clear measure of depression or depressive symptoms; Depression or depressive symptoms not identified as outcome variable.
Wang, M. (2007). Profiling retirees in the retirement transition and adjustment process: examining the longitudinal change patterns of retirees' psychological well-being. *Journal of Applied Psychology*, vol. 92(2), p. 455-474.	No clear measure of depression or depressive symptoms; Depression or depressive symptoms not identified as outcome variable. Variable concepts used – wellbeing, adaptation, satisfaction
Bosse, R., Aldwin, C.M., Levenson, M.R., Workman-Daniels, K. (1991). How stressful is retirement? Findings from the normative aging study. *Journals of Gerontology Series B*: *Psychological Sciences and Social Sciences*, vol. 46(1), p. 9-I4	No standardised measure of depression or depressive symptoms; Depression or depressive symptoms not identified as outcome variable
Braithwaite, V.A., Gibson, D.M., Bosly-Craft, R. (1986). An exploratory study of poor adjustment styles among retirees. *Social Science and Medicine*, vol. 23(5), p. 493-499.	No standardised measure of depression or depressive symptoms – only 1 *ad hoc* item to measure depression.
Mojon-Azzi, S., Sousa-Poza, A., Widmer, R.W. (2007). The effect of retirement on health: a panel analysis using data from the Swiss Household Panel. *Swiss Medical Weekly*, vol. 137, p. 581-585.	No standardised measure of depression or depressive symptoms – only 1 *ad hoc* item to measure depression.
Anderson, W.F., Cowan, N.R. (1956). Work and retirement: influences on the health of older men. *The Lancet*, vol. 29, p. 1344-1347	No standardised measure of depression or depressive symptoms – only 1 *ad hoc* item to measure 'happiness'
Ostberg, H., Samuelsson, S.-M. (1994). Occupational retirement in women due to age. *Scandinavian Journal of Social Medicine*, vol. 22(2), p. 90-96.	No standardised measure of depression or depressive symptoms – very unclear process of assessment or diagnosis.
Blazer, D. (1980). Life events, mental health functioning and the use of health care services by the elderly. *American Journal of Public Health*, vol. 70(11), p. 1174-1179.	No standardised measure of depression or depressive symptoms; Depression or depressive symptoms not identified as outcome variable – 'modified form of the Duke-OARS Multidimensional Functional Assessment Questionnaire'

	used as a measure of 'mental health functioning'.
Shieman, S., van Gundy, K., Taylor, J. (2001). Status, role and resource explanations for age patterns in psychological distress. Journal of Health and Social Behaviour, vol. 42(1), p. 80-96.	In 1996 sample (only sample of interest), no standardised measure of depression or depressive symptoms – 'depression index uses items similar to the CES-D' (p. 84).
Christ, S.L., Lee, D.J., Fleming, L.E., LeBlanc, W.G., Arheart, K.L., Chung-Bridges, K., Caban, A.J., McCollister, K.E. (2007). Employment and occupation effects on depressive symptoms in older Americans: does working past the age of 65 protect against depression? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(6), p. S399-S403.	No standardised measure of depression or depressive symptoms – symptoms of depression derived by Factor Analysis of general measure of 'generalised psychological distress'. Retired status of sample not clear – can only be inferred by the category 'not working' within the +60 age interval;
Demakakos, P., McMunn, A., Steptoe, A. (2010). Chapter 4: Well-being in older age: a multidimensional perspective. In, Banks, J, Lessof, C., Nazroo, J., Rogers, N., Stafford, M., Steptoe, A. (Eds.), Financial circumstances, health and well-being of the older population in England: The 2008 English Longitudinal Study Of Ageing (Wave 4), p. 115-177. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave4	No direct analysis of the relationship between retirement and outcomes
Tokuda, Y., Ohde, S., Takahashi, O., Shakudo, M., Yanai, H., Shimbo, T., Fukuhara, S., Hinohara, S., Fukui, T. (2008). Relationships between working status and health or health-care utilisation among Japanese elderly. *Geriatric Gerontology International*, vol. 8, p. 32-40.	Retired status of sample not clear – can only be inferred by the category 'not working' within the +60 age interval; No direct analysis of the relationship between retirement and outcomes.
Butterworth, P., Gill, S.C., Rodgers, B., Anstey, K.J., Villamil, E., Melzer, D. (2006). Retirement and mental health: analysis of the Australian national survey of mental health and wellbeing. Social Science and Medicine, vol. 63, p. 1179-1191.	Retired sample of the sample not clear; 'proxy for retirement' used – 'absence from the labour force between the ages of 45 and 74'
Alpass, F., Towers, A., Stephens, C., Fitzgerald, E., Stevenson, B., Davey, J. (2007). Independence, well-being, and social participation in an aging population. *Annals of the New York Academy of Sciences*, vol. 1114, p. 241-250.	Retired status of the sample not clear; No direct analysis of relationship between retirement and outcomes.
Blay, S.L., Andreoli, S.B., Fillenbaum, G.G., Gastal, F.L. (2007). Depression morbidity in later life: correlates in a developing country. *American Journal of Geriatric Psychiatry*, vol. 15(9), p. 790-799.	Retired status of the sample not clear – category 'no employment' includes unemployed and retired individuals; retired status can only be inferred by the category 'no employment' within the +60 age interval; No direct analysis of relationship between retirement and outcomes.
Jeffris, B.J., Nazareth, I., Marston, L., Moreno-Kustner, B., Bellón, J.A., Svab, I., Rotar, D., Geerlings, M.I., Xavier, M., Goçalves-Pereira, M., Vicente, B., Saldivia, S., Aluoja, A., Kalda, R., King, M. (2011). Associations between unemployment and major depressive disorder: evidence from an international, prospective study (the PREDICT cohort). *Social Science and Medicine,* vol. 73, p. 1627-1634.	Retired status of the sample not clear; No direct analysis of relationship between retirement and outcomes.
King, M., Nazareth, I., Levy, G., Walker, C., Morris, R., Weich, S., Bellón-Saameño, J.A., Moreno, B., Svab, I., Rotar, D., Rifel, J., Maaroos, H-I., Aluoja, A., Kalda, R., Neeleman, J., Geerlings, M.L., Xavier, M., Caldas de Almeida, M., Correa, B., Torres-Gonzalez, F. (2008). Prevalence of common mental disorders in general practice attendees across Europe. *British Journal of Psychiatry*, vol. 192, p. 362-367.	Retired status of the sample not clear; No direct analysis of relationship between retirement and outcomes.
Andrade, L.H., Benseñor, I.M., Viana, M.C., Andreoni, S., Wang, Y-P. (2010). Clustering of psychiatric and somatic illnesses in the general population: multimorbidity and socio-economic correlates.	Retired status of the sample not clear; No direct analysis of relationship between retirement and outcomes.

Brazilian Journal of Medical and Biological Research, vol. 43(5), p. 483-491.	
Melzer, D., Buxton, J., Villamil, E. (2004). Decline in common mental disorder prevalence in men in the sixth decade of life. *Social Psychiatry and Psychiatric Epidemiology*, vol. 39, p. 33-38.	Retired status of the sample not clear; Very heterogeneous group classed as 'economically inactive'
Buxton, J.W., Singleton, N., Melzer, D. (2005). The mental health of early retirees: National Interview Survey in Britain. *Social Psychiatry and Psychiatric Epidemiology*, vol. 40, p. 99-105	Retired status of the sample not clear; Very heterogeneous group classed as 'economically inactive'
Villamil, E., Huppert, F.A., Melzer, D. (2006). Low prevalence of depression and anxiety is linked to statutory retirement ages rather than personal work exit: a national survey. Psychological Medicine, vol 36, p. 999-1009.	Retired status of the sample not clear; Very heterogeneous group classed as 'economically inactive'
Tuohy, A., Knussen, C., Wrenall, M.J. (2005). Effects of age on symptoms of anxiety and depression in a sample of retired police officers. *Psychology and Aging*, vol. 20(2), p. 202-210.	Population with very specific characteristics – outcomes not generalizable to general population.
Lindeboom, M., Lindegaard, H. (2010). The impact of early retirement on health. Retrieved on the 8th of February 2013 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1672025	Population with very specific characteristics (early retirement window sample) – outcomes not generalizable to general population.
Atchley, R.C. (1976). Selected social and psychological differences between men and women in later life. *Journal of Gerontology*, vol 31(2), p. 204-211.	Cross-sectional study includes only comparisons *within* a retired sample.
Elgarresta, I.L., de Miguel, M.S., Arruabarrena, L. R. (2009). Diferentes formas de aceder a la jubilación y su relación con la salud psicológica. *Revista Española de Gerontologia*, vol. 44(6), p. 311-316.	Cross-sectional study includes only comparisons *within* a retired sample.
Sharpley, C.F., Layton. R. (1998). Effects of age of retirement, reason for retirement and pre-retirement training on psychological and physical health during retirement. *Australian Psychologist*, vol. 33(2), p. 119-124.	Cross-sectional study includes only comparisons *within* a retired sample.
Zenger, M., Brähler, E., Berth, H., Stöbel-Richter, Y. (2011). Unemployment during working life and mental health of retirees: results of a representative survey. *Aging and Mental Health*, vol. 15(2), p. 178-185.	Cross-sectional study includes only comparisons *within* a retired sample.
McMunn, A., Nazroo, J., Wahrendorf, M., Breeze, E., Zaninotto, P. (2009). Participation in socially productive activities, reciprocity and wellbeing in later life: baseline results in England. Aging and Society, vol. 29, 765-782.	Cross-sectional study includes only comparisons *within* a retired sample.
Gall, T.L., Evans, D.R., Howard, J. (1997). The retirement adjustment process: changes in the wellbeing of male retirees across time. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 52B(3), p. 110-117.	Studies do not use a specific measure of depression. Study used the SCL-90R (Derogatis, 1983) Global Severity Index, with no specific use of the Depression sub-scale. Therefore, the outcomes do not provide a specific measure of depression, and do not allow the paper to meet the inclusion criterion. Also, Clark and Friedman (1983), among others, have questioned the dimensionality (i.e. factorial structure) of the SCL-90 and consider it valid only as a measure of overall psychological distress.
Gall, T.L., Evans, D.R. (2000). Preretirement expectations and the quality of life of male retirees in later retirement. Canadian Journal of Behavioural Science, vo. 32(3), p. 187-197.	
Gill, S.C., Butterworth, P., Rodgers, B., Anstey, K.J. , Villamil, E., Melzer, D. (2006). Mental health and the timing of men's retirement. *Social Psychiatry and Psychiatric Epidemiology*, vol. 41, p. 515-522.	Studies do not use a specific measure of depression. Studies and reports used the SF-36 (Ware and Sherbourne, 1992). The SF-36 is not a specific measure of depression. The factorial structure of the SF-36 reveals only a
Stephens, C., Noone, J. (2008). *Health, Work and Retirement Survey: Summary report for the 2006*	

data wave. Health. Retrieved on the 7th of April 2013 from http://hwr.massey.ac.nz/resources/social-support_Stephens-Noone.pdf	general mental health factor which maps on to the mental health scale (Keller and Ware et al., 1998). Therefore this raises concerns about its specificity as a measure of depression or depressive symptoms. Furthermore, the explicit use of the scale in these studies/reports as a general measure of mental health or mental health functioning, do not allow them to meet the inclusion criterion.
Alpass, F. (2008). *Health, Work and Retirement Survey: Summary report for the 2006 data wave. Work and Retirement.* Retrieved on the 7th of April 2013 from http://hwr.massey.ac.nz/resources/Work_Fiona%20Alpass.pdf	
Jokela, M., Ferrie, J.E., Gimeno, D., Chandola, T., Shipley, M.J., Head, J., Vahtera, J., Westerlund, H., Marmot, M.G., Kivimaki, M. (2010). From midlife to early old age: health trajectories associated with retirement. *Epidemiology*, vol. 21(3), p. 284-290.	
Mein, G., Martikainen, P., Hemingway, H., Stansfels, S., Marmot, M. (2003). Is retirement good or bad for mental and physical health functioning? Whitehall II longitudinal study of civil servants. *Journal of Epidemiology and Community Health*, vol. 57, p. 46-49.	
Dulin, P., Stephens, C., Alpass, F., Hill, R.D., Stevenson, B. (2011). The impact of socio-cultural, physical and lifestyle variables on measures of physical and psychological wellbeing among Maori and non-Maori: the New Zealand Health, Work and Retirement Study. *Ageing and Society*, vol 31, p. 1406-1424.	
Talala, K., Huurre, T., Aro, H., Martelin, Tuija, Prattala, R. (2008). Socio-Demographic differences in self-reported psychological distress among 25 to 64 year-old Finns. *Social Indices Research*, vol. 86, p. 323-335	
Salokangas, R.K.R., Joukamaa, M. (1991). Physical and mental health changes in retirement age. *Psychotherapy and Psychosomatics*, vol. 55, p. 100-107.	Studies do not use a specific measure of depression. Studies used the General Health Questionnaire (GHQ) (Goldberg, 1972) 30 and 36-item versions. This is not a specific measure of depression, and was designed to screen a wide range of common psychiatric morbidity (Goldberg and Gater, et al., 1997), covering symptoms of anxiety and depression, social dysfunction and loss of confidence (Aalto, Elovainio, Kivimäki, Uutela, Pirkola, 2012). Although the GHQ has been used in survey studies as indicator of depression, concerns about its specificity remain. Furthermore, these studies used this measure as a general indicator of mental health or mental health functioning. Therefore, these do not meet the inclusion criterion.
Hyde, M., Ferrie, J., Higgs, P., Mein, G., Nazroo, J. (2004). The effects of pre-retirement factors and retirement route on circumstances in retirement: findings form the Whitehall II Study. *Ageing and Society*, vol 24(2), p. 279-296.	
Mattila, V.J., Joukamaa, M.I., Salokangas, R.K.R. (1989). Retirement, aging, psychosocial adaptation and mental health: findings of the TURVA project. *Acta Psychiatrica Scandinavica*, vol. 80, p. 356-367.	
Artazcoz, L., Cortés, I., Borrell, C., Escribá-Agüir, V., Cascant, L. (2010). Gender and social class differences in the association between early retirement and health in Spain. *Women's Health Issues*, 20(6), p. 441-447.	Studies do not use a specific measure of depression. Studies used the GHQ 12. As above. Aalto et al. (2012) conclude that, although the GHQ 12 can function as a reasonable measure of depressive disorder, the scale did not differentiate well between depressive and other psychiatric disorders. Therefore, again, concerns about its specificity and the explicit use of the scale in the paper as a general measure of mental health or mental health functioning, do not allow these studies to meet the inclusion criterion.
Lindstrom, M., Ali, S.M., Rosvall, M. (2012). Socioeconomic status, labour market connection and self-rated psychological health: the role of social capital and economic stress. Scandinavian Journal of Public Health, vol. 40, p. 51-60.	
Potočnik, K., Tordera, N., Peiró, J.M. (2010). The influence of the early retirement process on satisfaction with early retirement and psychological wellbeing. *International Journal of Aging and Human Development*, vol. 70(3), p. 251-273.	
Fe, E., Hollingsworth, B. (2012). *Estimating the effect of retirement on mental health via panel*	

discontinuity designs. Retrieved on the 7th of April 2013 from http://mpra.ub.uni-muenchen.de/38162/	
Total Excluded: 49	
Total Retained: 38	

Table 3 – Sift 6 (Step 2) in selection process – 1 report per study selected, with reasons for exclusion of other reports

Note: One report per study was selected, the table is organised in reference to 'parent' studies with more than one report.

Reference	Reasons for exclusion
HRS	
Calvo, E., Sarkisian, N. (2011). *Retirement and well-being: examining the characteristics of life-course transitions.* Retrieved on the 10th of February 2013 from www.politicaspublicas.udp.cl	Excluded for Calvo et al. (2013) Less clear descriptive statistic Not peer reviewed report
Coe, N.B., Lindeboom, M. (2008). Does Retirement Kill You? Evidence from Early Retirement Windows. Retrieved on the 8th of February 2013 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295315	Excluded for Calvo et al. (2013) Uses early retirement windows methodology – specific sample, not comparable.
Dave, D., Rashad, I., Spasojevic, J. (2008a). The effects of retirement on physical and mental health outcomes. Retrieved on the 25th of March 2013 from http://aysps.gsu.edu/uwrg-research.html	Excluded for Calvo et al. (2013) Less clear descriptive statistics
Dave, D., Rashad, I., Spasojevic, J. (2007). The effects of retirement on physical and mental health outcomes. Retrieved on the 8th of February 2013 from http://aysps.gsu.edu/publications/2007/index.htm	Excluded for Calvo et al. (2013) Less clear descriptive statistics
Dave, D., Rashad, I., Spasojevic, J. (2008b). The effects of retirement on physical and mental health outcomes. *Southern Economic Journal*, vol 75(2), p. 497-523.	Excluded for Calvo et al. (2013) Less clear descriptive statistics
Fondnow, M.D.M. (2007). Effects of retirement on health among men and women in the Health and Retirement Study. Unpublished PhD Thesis, Ohio State University. Retrieved on the 8th of February 2013 from https://etd.ohiolink.edu/ap:0:0:APPLICATION_PROCESS=DOWNLOAD_ETD_SUB_DOC_ACCNUM:::F1501_ID:osu1180107602,attachment	Excluded for Calvo et al. (2013) Less clear descriptive statistics Smaller data set Not a peer reviewed report
Hao, Y. (2008). Productive activities and psychological well-being among older adults. *Journals of Gerontology Series B*: *Psychological Sciences and Social Sciences*, vol. 63B(2), p. S64-S72	Excluded for Calvo et al. (2013) Less clear descriptive statistics Smaller data set
Mandal, B., Roe, B. (2008). Job loss, retirement and the mental health of older Americans. *The Journal of Mental Health Policy and Economics,* vol. 11, p. 167-176	Excluded for Calvo et al. (2013) Less clear descriptive statistics Smaller data set
Silver, M.P. (2010): Women's retirement and self-assessed well-being: an analysis of three measures of well-being among recent and long-term retirees relative to homemakers. *Women & Health*, vol. 50, p. 1-19	Excluded for Calvo et al. (2013) Less clear descriptive statistics

	Smaller data set
Szinovacz, M.E., Davey, A. (2004a). Honeymoons and joint lunches: effects of retirement and spouse's employment on depressive symptoms. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* vol. 59B(5), p. 233-245.	Excluded for Calvo et al. (2013) Less clear descriptive statistics Smaller data set
Szinovacz, M.E., Davey, A. (2004). Retirement transitions and spouse disability: effects on depressive symptoms. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* vol. 59B(6), p. S333-S342.	Excluded for Calvo et al. (2013) Less clear descriptive statistics Smaller data set
Szinovacz, M.E., Davey, A. (2006). Effects of retirement and grandchild care on depressive symptoms. *International Journal of Aging and Human Development,* vol. 62(1), p. 1-20.	Excluded for Calvo et al. (2013) Less clear descriptive statistics Smaller data set
SHARE	
Alavinia, S.M., Burdorf, A. (2008). Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. *International Archives of Occupational and Environmental Health*, vol. 82, p. 39-45	Excluded for Brugiavini et al. (2008b) Smaller data set (1 Wave of data only – 2004)
Coe, N.B., Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, vol. 30, p. 77-86.	Excluded for the Brugiavini et al. (2008b) Smaller data set (1 Wave of data only – 2004)
Dewey, M.E., Prince, M.J. (2005). Mental Health. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) Health, Ageing and Retirement in Europe: First Results from the Survey of Health, Ageing and Retirement in Europe, p. 108-117. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/SHARE_FirstResultsBookWave1.pdf	Excluded for Brugiavini et al. (2008b) Report does not include specific analysis of the relationship between retirement and depression.
Kolodziej, I. (2011). The relationship between retirement and mental health: investigating the causal relationship in eleven European countries using SHARE. Unpublished MSc Thesis, Erasmus University Rotterdam. Retrieved on the 8th of February 2013 from http://arno.uvt.nl/show.cgi?fid=122185	Excluded for Brugiavini et al. (2008b) Less clear descriptive statistics
Sahlgren, G.H. (2012). Work 'til you drop: short- and longer-term health effects of retirement in Europe. Research Institute of Industrial Economics (IFN, Stockholm, Sweden) Working Paper No. 928. Retrieved on the 10th of February 2013 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2153191	Excluded for Brugiavini et al. (2008b) Less clear descriptive statistics
Sahlgren, G.H. (2013). Work longer, live healthier: the relationship between economic activity, health and government policy. Institute of Economic Affairs (London) Discussion Paper No. 46. Retrieved on the 20th of May 2013 from http://www.iea.org.uk/sites/default/files/publications/files/Work%20Longer,%20Live_Healthier.pdf	Excluded for Brugiavini et al. (2008b) Less clear descriptive statistics
ELSA	
Behncke, S. (2012). Does retirement trigger ill health? Health Economics, vol. 21, p. 282-300.	Excluded for Steptoe et al. (2012) Very unclear descriptive statistics
Other Studies	
Herzog, A.R., House, J.S., Morgan, J.N. (1991). Relation of work and retirement to health and well-being in older age. *Psychology and Aging*, vol. 6(2), p. 202-211.	Study: American Changing Lives Survey Excluded for Clarke et al. (2011) Cross sectional design Smaller N

Moen, P., Erickson, W.A., Agarwal, M., Fields, V., Todd, L. (2000). The Cornell Retirement and Well-Being Study: Final Report. Ithaca, New York: Bronfenbrenner Life Course Center, Cornell University. Retrieved on the 4th of March 2013 from http://worlddatabaseofhappiness.eur.nl/hap_bib/freetexts/moen_p_2000.pdf	Study: Cornell Retirement and Wellbeing Study (USA) Excluded for Kim and Moen (2002) Less clear descriptive statistics Unpublished
Ross, C. E., Drentea, P. (1998). Consequences of retirement activities for distress and sense of personal control. *Journal of Health and Social Behavior*, vol. 39(4). p. 317-334.	Study: Survey of Aging, Status and Sense of Control (USA) Excluded for Drentea (2002) Fewer data points/observations
Mixed Samples	
Charles, K.K. (2002). Is retirement depressing? Labour force inactivity and psychological wellbeing in later life. Working Paper, National Bureau of Economic Research. Retrieved on the 9th of March 2013 from http://www.nber.org/papers/w9033	Cross-sectional analysis of longitudinal data; Report uses a combination of 3 samples from 3 different studies. Data combined, sample/study provenance unclear.
Crimmins, E.M., Kim, J.K., Solé-Auró, A. (2010). Gender differences in health: results from the SHARE, ELSA and HRS. *European Journal of Public Health*, vol. 21(1), p. 81-91.	Cross-sectional analysis of longitudinal data; Report uses a combination of 3 samples from 3 different studies. Data combined, sample/study provenance unclear.
Total Excluded: 24 **Total Retained: 15**	

Appendix 4a

Coding Manual – Cohort Studies

(Lipsey and Wilson, 2001; Shamliyan and Kane et al., 2011; Luhmann et al. 2012)

Full Reference		**STUDY ID**

		Variable Name and Description	**Measurement level**	**Coding Key and Instructions**	**Variable ID**
Study Level	Research Design Descriptors				
		1. Overall design of the study	Categorical	1 = Cohort (a study in which a cohort is followed over time, to examine associations between different interventions received and subsequent outcomes; a 'prospective' cohort study recruits participants before any intervention and follows them into the future' 2 = Cross-sectional (a study that collects information on interventions [past or present] and current health outcomes, i.e. restricted to health states, for a group of people at a particular point in time, to examine associations between the outcomes and exposure to interventions)	DESIGNTYPE
		2. Time orientation of investigation	Categorical	1 = Prospective pre-retirement (baseline assessment occurred before retirement began) 2 = Prospective post-retirement (baseline assessment occurred after retirement began) 3 = Retrospective (baseline assessment occurred after retirement, i.e. data were collected	TIMEOR

		retrospectively on the index variable) 4 = Prospective and Retrospective (retirement is the reference point; data on the index variable is collected before and after retirement)	
3. Is this a secondary analysis of previously collected data?	Categorical	1 = Yes 2 = No 99 = Cannot be determined	2ANALYSIS
4. Overall confidence of judgement on design type	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONDESIGN
5. Location of the first author's institutional affiliation (the location of the institution to which the main author of the study is affiliated)	Categorical	1 = USA 2 = Canada 3 = South America 4 = Europe 5 = UK 6 = Australia and New Zealand 7 = Asia	AULOCAL
6. Subject area of the first author	Categorical	1 = Psychology 2 = Medicine (including Public Health and Psychiatry) 3 = Sociology 4 = Economics	AUSUBJECT
7. Type of publication	Categorical	1 = Published in peer-reviewed journal 2 = Published in institutional website (not peer-reviewed) 3 = Unpublished	PUBTYPE
Sample Descriptors			
8. Geographical origin of the study sample (Where the study participants are from)	Categorical	1 = North America – USA 2 = North America – Canada 3 = South America	SAMPLOCAL

		4 = Europe – multiple countries 5 = Europe – UK 6 = Central/Northern European Countries (France, Netherlands, Germany) 7 = European Scandinavian Countries (Finland, Norway, Sweden, Denmark) 8 = Southern European Countries (Spain, Greece, Portugal, Italy) 9 = Australia and New Zealand 10 = Asia (China, Japan, Korea, India)	
9. Sampling method – general population (Sampling method for the study sample, if this was selected from the general population)	Categorical	1 = Random 2 = Non-random 3 = Random multistage, stratified 4 = Random restricted to geographic area 6 = Other sampling of the general population (specify) 88 = Not applicable 99 = Cannot be determined	SAMPGP
10. Sampling method – non-general population (Sampling method for the study sample, if this was selected from a non-general population)	Categorical	1 = Random 2 = Convenient with automatic enrolment 3 = Convenient with self-selection 4 = Other (specify) 88 = Not applicable 99 = Cannot be determined	SAMPNGP
11. Sampling frame (Sampling frame for the study sample, from general or non-general populations)	Categorical	1 = Nationally representative registries or databases 2 = Medical, social, or insurance records 3 = Work place settings 4 = Health Care settings (Clinics, Hospitals) 6 = Proxy selection (parents, relatives, legal representatives, care takers...) 7 = Other (specify) 8 = Multiple sources 88 = Not applicable 99 = Cannot be determined	SAMPFRA

12. Overall confidence of judgement on sampling method and frame	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	RCONSAMP
13. Method used to determine retired status	Categorical	1 = Self-assessed (the individuals describes him/herself as retired) 2 = Independent data (employment, insurance, health or social securing records are used to identify retired individuals) 3 = Both methods are used 99 = Cannot be determined	RSTATUS
14. Overall confidence of judgement on method used to determine retired status	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONRSTATUS
15. Total sample size at start of the study (Total N1)	Metric	99.99 = Cannot be determined	TOTALN1
16. Total sample size at final time point (Total N2)	Metric	99.99 = Cannot be determined	TOTALN2
17. Sample attrition percentage (Percentage of lost data or participants)	Metric	99.99 = Cannot be determined	ATTRITION
18. Evidence of systematic dropout?	Categorical	1 = Yes 0 = No 99 = Cannot be determined	SYSTDROP
19. Proportion of men in the sample (percentage)	Metric	99.99 = Cannot be determined	MALE%
20. Proportion of women in the sample (percentage)	Metric	99.99 = Cannot be determined	FEMALE%
21 Sample age mean (M)	Metric	99.99 = Cannot be determined	AGEMEAN
22. Sample age standard deviation (SD)	Metric	99.99 = Cannot be determined	AGESD
23. Sample time in retirement mean (M)	Metric	99.99 = Cannot be determined	TIMERMEAN
24. Sample time in retirement standard deviation (SD)	Metric	99.99 = Cannot be determined	TIMERSD
25. Sample age at retirement mean (M)	Metric	99.99 = Cannot be determined	AGERMEAN
26. Sample age at retirement standard deviation (SD)	Metric	99.99 = Cannot be determined	AGERSD

Effect Size Level

Dependent Measure Descriptors			
27. Type of outcome measure	Categorical	1 = Diagnostic indicator of presence of depression (pre-existing diagnosis or outcome of a clinical diagnostic interview, e.g. CIDI interview) 2 = Indicator of severity of depressive symptoms (e.g. CES-D, EURO-D) 3 = Other indicator of presence of depression or depressive symptoms (e.g. purchase of anti-depressive medication)	TYPEMEASURE
28. If 2, instrument/scale used to measure severity of depressive symptoms	Categorical	1 = CES-D (indicate version) 2 = EURO-D 3 = SCL 90 88 = Not applicable	MEASURE
29. If 2, Positive vs Negative coding	Categorical	-1 = High values indicate high depression or depressive symptoms 1 = High values indicate low depression or depressive symptoms 88 = Not applicable	CODING
30. If 1 or 2, source of reported reliability estimate	Categorical	0 = Not reported 1 = Reported but reference to another publication 2 = Reported and calculated for sample of this study 88 = Not applicable 99 = Cannot be determined	SOURRELI
31. If 1 or 2, reliability estimate	Metric	88.88 = Not applicable 99.99 = cannot be determined	RELI
32. If 1 or 2, source of reported validity estimate	Categorical	0 = Not reported 1 = Reported but reference to another publication 2 = Reported and calculated for sample of this study 88 = Not applicable 99 = Cannot be determined	SOURVALID
33. If 1 or 2, validity estimate	Metric	88.88 = Not applicable 99.99 = not reported	VALID

34. Time frame of measure	Categorical	1 = General state 2 = Depression with respect to the event, i.e. retirement 3 = Precise time frame (e.g. 'last week') 88 = Not applicable 99 = Cannot be determined	TIMEFRAME
35. Number of measurement time points (How many times was the dependent variable measured in the study)	Metric	99.99 = Cannot be determined	TIMEPOINTSN
36. Year data collection started	Metric	99.99 = cannot be determined	STARTYEAR
37. Year data collection ended	Metric	99.99 = cannot be determined	ENDYEAR
38. Delay between retirement and first measurement occasion (in months)	Metric	99.99 = cannot be determined	MEASUREDELAY
39. Time points used for effect size calculation (At what time points was the dependent variable measured)	Categorical	1 = Pre-retirement and retirement transition 2 = Retirement transition and post-retirement period 3 = Time points during post-retirement period 4 = Pre-retirement and post retirement period 5 = Pre-retirement and retirement transition and post-retirement period 99.99 = cannot be determined	TIMEPOINTSES
Effect Size Data			
40. Statistic the study's effect size is based on	Categorical	1 = Means and standard deviations 2 = *t*-value or *F*-value 3 = chi-square 4 = Frequencies or proportions dichotomous 5 = Frequencies or proportions polycotomous 6 = Correlation value (R_P or R_S) 6 = Other (specify)	ESSTAT
41. Page number where the data for the effect size is found	--		PAGE
42. Raw difference favours (lower or fewer depression indicators or symptoms) which time point?	Categorical	1 = Pre-retirement 2 = Retirement transition	SUCCESS

		3 = Post-retirement period 4 = None (equal) 5 = Cannot tell	
When means and standard deviations are reported or can be estimated:			
43. Pre-retirement sample size (n)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	PRERN
44. Retirement transition sample size (n)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	TRANSRN
45. Post-retirement sample size (n)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	POSTRN
46. Pre-retirement mean (M)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	PRERMEAN
47. Retirement transition mean (M)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	TRANSRMEAN
48. Post-retirement mean (M)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	POSTRMEAN
49. Pre-retirement standard deviation (SD)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	PRERSD
50. Retirement transition standard deviation (SD)	Metric	88.88 = Not applicable 99.99 = Cannot be determined	TRANSRSD
51. Post-retirement standard deviation (SD)	Metric	88.88 = Not applicable 99.99 = cannot be determined	POSTRSD
When proportions or frequencies are reported or can be estimated			
52. n of pre-retirement sample with a successful outcome	Metric	88.88 = Not applicable 99.99 = cannot be determined	PRERSUCCN
53. n of transition sample with a successful outcome	Metric	88.88 = Not applicable 99.99 = cannot be determined	TRANSRSUCCN
54. n of post-retirement sample with a successful outcome	Metric	88.88 = Not applicable 99.99 = Cannot be determined	POSTRSUCCN
55. Proportion of pre-retirement sample with a successful outcome	Metric	88.88 = Not applicable 99.99 = Cannot be determined	PRERSUCCPROP
56. Proportion of transition sample with a successful outcome	Metric	88.88 = Not applicable 99.99 = cannot be determined	TRNSRSUCCPROP

57. Proportion of post-retirement sample with a successful outcome	Metric	88.88 = Not applicable 99.99 = cannot be determined	POSTRSUCCPROP
When significance test information is reported			
58. t-value	Metric	99.99 = cannot be determined	T-VALUE
59. F-value (df for the numerator must equal 1)	Metric	99.99 = cannot be determined	F-VALUE
60. Chi-Square value (df = 1)	Metric	99.99 = cannot be determined	CHISQUARE
Statistical Power Calculation			
61. Power calculation presented in study	Metric	99.99 = Cannot be determined	POWER1
62. Statistical Power (post-hoc calculated by rater)	Metric	99.99 = Cannot be determined	POWER2
Calculated Effect Size			
63. Effect size calculated using the procedures outlined in Lipsey and Wilson (2001) (Report two decimals with an algebraic sign in front)	Metric	(plus) - = difference favours post-retirement (minus) + = difference favours pre-retirement 99.99 = cannot be determined	TOTALES
64. Overall confidence rating in effect size computation	Categorical	1 = highly estimated (have N and crude p-value only, such as p<.10 and must reconstruct via rough t-test equivalence) 2 = Moderate estimation (have complex but relatively complete statistics, such as multifactor ANOVA as basis for estimation) 3 = Some estimation (have unconventional statistics and must convert to equivalent t-values; or have conventional statistics but incomplete, such as exact p-level) 4 = Slight estimation (must use significance testing statistics, rather than descriptive statistics, but have complete statistics of the conventional sort) 5 = No estimation (have descriptive data such as means, standard deviations, frequencies, proportions, etc, and can calculate the effect size directly)	CONTOTALES

Appendix 4b

Coding Manual – Cross-Sectional Studies

(adapted from Lipsey and Wilson, 2001; Shamliyan and Kane et al., 2011; Luhmann et al. 2012)

Full Reference		**STUDY ID**

	Variable Name and Description	**Measurement level**	**Coding Key and Instructions**	**Variable ID**
Study Level	Research Design Descriptors			
	1. Overall design of the study	Categorical	1 = Cohort (a study in which a cohort is followed over time, to examine associations between different interventions received and subsequent outcomes; a 'prospective' cohort study recruits participants before any intervention and follows them into the future' 2 = Cross-sectional (a study that collects information on interventions [past or present] and current health outcomes, i.e. restricted to health states, for a group of people at a particular point in time, to examine associations between the outcomes and exposure to interventions)	DESIGTYPE
	2. Overall confidence of judgement on design type	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONDESIGN

3. Is this a secondary analysis of previously collected data?	Categorical	1 = Yes 2 = No 99 = Cannot be determined	2ANALYSIS
4. Location of first author's institutional affiliation (the location of the institution to which the main author of the study is affiliated)	Categorical	1 = USA 2 = Canada 3 = South America 4 = Europe 5 = UK 6 = Australia and New Zealand 7 = Asia	AULOCAL
5. Subject area of the first author	Categorical	1 = Psychology 2 = Medicine (including Public Health and Psychiatry) 3 = Sociology 4 = Economics 99 = Cannot be determined	AUSUBJECT
6. Type of publication	Categorical	1 = Published in peer-reviewed journal 2 = Published in institutional website (not peer-reviewed) 3 = Unpublished	PUBTYPE
Sample Descriptors			
7. Geographical origin of the study sample (Where the study participants are from)	Categorical	1 = North America – USA 2 = North America – Canada 3 = South America 4 = Europe – multiple countries 5 = Europe – UK 6 = Central/Northern European Countries (France, Netherlands, Germany) 7 = European Scandinavian Countries (Finland, Norway, Sweden, Denmark) 8 = Southern European Countries (Spain, Greece, Portugal, Italy)	SAMPLOCAL

		9 = Australia and New Zealand 10 = Asia (China, Japan, Korea, India)	
8. Retired sample sampling method – general population (Sampling method for the retired sample, if this was selected from the general population)	Categorical	1 = Random 2 = Non-random 3 = Random multistage, stratified 4 = Random restricted to geographic area 5 = Other sampling of the general population (specify) 88 = Not applicable 99 = Cannot be determined	RSAMGP
9. Retired sample sampling method – non-general population (Sampling method for the retired sample, if this was selected from a non-general population)	Categorical	1 = Random 2 = Convenient with automatic enrolment 3 = Convenient with self-selection 4 = Other (specify) 88 = Not applicable 99 = Cannot be determined	RSAMNGP
10. Retired sample sampling frame	Categorical	1 = Nationally representative registries or databases 2 = Medical, social, or insurance records 3 = Work place settings 4 = Health Care settings (Clinics, Hospitals) 6 = Proxy selection (parents, relatives, legal representatives, care takers...) 7 = Other (specify) 8 = Multiple sources 88 = Not applicable 99 = Cannot be determined	RSAMPFRA
11. Overall confidence of judgement on sampling method and frame for retired sample	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONRSAMP
12. Non-retired sample sampling method – general population (Sampling method for the non-retired sample, if this was selected	Categorical	1 = Random 2 = Non-random	NRSAMGP

from the general population)		3 = Random multistage, stratified 4 = Random sampling restricted to geographic area 5 = Other sampling of the general population (specify) 88 = Not applicable 99 = Cannot be determined	
13. Non-retired sample sampling method – non-general population (Sampling method for the non-retired sample, if this was selected from a non-general population)	Categorical	1 = Random 2 = Convenient with automatic enrolment 3 = Convenient with self-selection 4 = Other (specify) 88 = Not applicable 99 = Cannot be determined	NRSAMPNGP
14. Non-retired sample sampling frame	Categorical	1 = Nationally representative registries or databases 2 = Medical, social, or insurance records 3 = Work place settings 4 = Health Care settings (Clinics, Hospitals) 6 = Proxy selection (parents, relatives, legal representatives, care takers...) 7 = Other (specify) 8 = Multiple sources 88 = Not applicable 99 = Cannot be determined	NRSAMPFRANGP
15. Overall confidence of judgement on sampling method and frame for non-retired sample	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONNRSAMP
16. Were samples selected from the same population? (Were the retired and non-retired samples selected from the same population?)	Categorical	1 = Non-retired sample is from the same population as retired sample 2 = Non-retired sample is from different population as retired sample 99 = Cannot be determined	POPSAME
17. Was the equivalence of the samples tested before any other	Categorical	1 = Yes	PREEQUIV

analysis?		2 = No 99 = Cannot be determined	
18. If tested, what differences were found? (Key: Important differences – a difference on several variables, or on a major variable, or a large (size) difference; Major variables – variables likely to be related to retirement, e.g. time in retirement, age at time of retirement, sex, age)	Categorical	1 = Negligible differences, judged unimportant 2 = Some differences, judged of uncertain importance 3 = Some differences, judged important 88 = Not applicable 99 = Cannot be determined	PREDIF
19. Overall confidence of judgement on the equivalence of samples or judgement on differences between samples	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONEQUIV
20. Method used to determine retired status (How participants' retired status was determined)	Categorical	1 = Self-assessed (the individual describes him/herself as retired) 2 = Independent data (employment, insurance, health or social security records are used to identify retired individuals) 3 = Both methods were used 99 = Cannot be determined	RSTATUS
21. Overall confidence of judgement on method used to determine retired status	Categorical	1 = Very low (little basis) 2 = Low (guess) 3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	CONRSTATUS
22. Method used to determine non-retired status (How participants' worker status was determined)	Categorical	1 = Self-assessed (the individual describes him/herself as not retired) 2 = Independent data (employment, insurance, health or social security records are used to identify not retired individuals) 3 = Both methods are used 99 = Cannot be determined	NRSTATUS
23. Overall confidence of judgement on method used to determine non-retired status	Categorical	1 = Very low (little basis) 2 = Low (guess)	CONNRSTATUS

		3 = Moderate (weak inference) 4 = High (strong inference) 5 = Very High (explicitly stated)	
24. Total sample size at start of study (N Total)	Metric	99.99 = Cannot be determined	TOTALN
25. Retired sample size at start of study (n retired)	Metric	99.99 = Cannot be determined	RN
26. Non-retired sample size at start of study (n non-retired)	Metric	99.99 = Cannot be determined	NRN
27. Total sample attrition percentage (Percentage of lost data or participants)	Metric	99.99 = Cannot be determined	ATTRITION
28. Total sample proportion of men in sample (percentage)	Metric	99.99 = Cannot be determined	MALEN
29. Total sample proportion of women in sample (percentage)	Metric	99.99 = Cannot be determined	FEMALEN
30. Total sample age mean (M)	Metric	99.99 = Cannot be determined	SAMPAGEMEAN
31. Total sample age standard deviation (SD)	Metric	99.99 = Cannot be determined	SAMPAGESD
32. Proportion of men in retired sample (percentage)	Metric	99.99 = Cannot be determined	RMALE
33. Proportion of women in retired sample (percentage)	Metric	99.99 = Cannot be determined	RFEMALE
34. Retired sample mean age (M)	Metric	99.99 = Cannot be determined	RAGEMEAN
35. Retired sample age standard deviation (SD)	Metric	99.99 = Cannot be determined	RAGESD
36. Proportion of men in non-retired sample (percentage)	Metric	99.99 = Cannot be determined	NRMALE
37. Proportion of women in non-retired sample (percentage)	Metric	99.99 = Cannot be determined	NRFEMALE
38. Non-retired sample mean age (M)	Metric	99.99 = Cannot be determined	NRAGEMEAN
39. Non-retired sample age standard deviation (SD)	Metric	99.99 = Cannot be determined	NRAGESD
40. Retired sample time in retirement mean (M) (in months)	Metric	99.99 = Cannot be determined	RTIMEMEAN
41. Retired sample time in retirement standard deviation (SD)	Metric	99.99 = Cannot be determined	RTIMESD
42. Retired sample age at retirement mean (M) (in years)	Metric	99.99 = Cannot be determined	RETAGEMAEAN
43. Retired sample age at retirement standard deviation (SD)	Metric	99.99 = Cannot be determined	RETAGESD
Effect Size Level Dependent Measure Descriptors			
44. Type of outcome measure	Categorical	1 = Diagnostic indicator of presence of depression (pre-existing diagnosis or outcome of a clinical diagnostic interview, e.g. CIDI interview)	MEASTYPE

		2 = Indicator of severity of depressive symptoms (e.g. CES-D, EURO-D) 3 = Other indicator of presence of depression or depressive symptoms (e.g. purchase of anti-depressive medication)	
45. If 2, instrument/scale used to measure severity of depressive symptoms	Categorical	1 = CES-D (indicate version) 2 = EURO-D 3 = SCL 90 88 = Not applicable	MEASURE
46. If 2, Positive vs Negative coding	Categorical	-1 = High values indicate high depression or depressive symptoms 1 = High values indicate low depression or depressive symptoms 88 = Not applicable	CODING
47. If 1 or 2, source of reported reliability estimate	Categorical	0 = Not reported 1 = Reported but reference to another publication 2 = Reported and calculated for sample of this study 88 = Not applicable	SOURRELI
48. If 1 or 2, reliability estimate	Metric	88 = Not applicable 99.99 = Cannot be determined	RELI
49. If 1 or 2, source of reported validity estimate	Categorical	0 = Not reported 1 = Reported but reference to another publication 2 = Reported and calculated for sample of this study 88 = Not applicable	SOURVALID
50. If 1 or 2, validity estimate	Metric	88 = Not applicable 99 = Cannot be determined	VALID
51. If 1 or 2, time frame of measure	Categorical	1 = General State 2 = Depression with respect to the event, i.e. retirement 3 = Precise time frame (e.g. last week) 88 = Not applicable 99 = Cannot be determined	TIMEFRAME

52. Year data collection started	Metric	99.99 = Cannot be determined	STARTYEAR
53. Year data collection ended	Metric	99.99 = Cannot be determined	ENDYEAR
54. Delay between retirement and measurement occasion (In months)	Metric	99.99 = Cannot be determined	MEASUREDELAY
Effect Size Data			
55. Statistic the study's effect size is based on	Categorical	1 = Means and standard deviations 2 = *t*-value or *F*-value 3 = Chi-square 4 = Frequencies or proportions dichotomous 5 = Frequencies or proportions polycotomous 6 = Correlation value (R_P or R_S) 7 = Other (specify)	ESSTAT
56. Page number where the data for the effect size is found	--		PAGE
57. Raw difference favours (lower or fewer depression indicators of symptoms) which sample?	Categorical	1 = Retired sample 2 = Non-retired sample 3 = Neither (equal) 4 = Cannot tell	SUCCESS
When means and standard deviations are reported or can be estimated:			
58. Retired sample size	Metric	99.99 = Cannot be determined	RN
59. Non-retired sample size	Metric	99.99 = Cannot be determined	NRN
60. Retired sample mean (M)	Metric	99.99 = Cannot be determined	RMEAN
61. Non-Retired sample mean (M)	Metric	99.99 = Cannot be determined	NRMEAN
62. Retired sample standard deviation (SD)	Metric	99.99 = Cannot be determined	RSD
63. Non-retired sample standard deviation (SD)	Metric	99.99 = Cannot be determined	NRSD
When proportions or frequencies are reported or can be estimated			
64. *n* of retired sample with a successful outcome	Metric	99.99 = Cannot be determined	RSUCCESS
65. *n* of non-retired sample with a successful outcome	Metric	99.99 = Cannot be determined	NRSUCCESS
66. Proportion of retired sample with a successful outcome	Metric	99.99 = Cannot be determined	RPROP
67. Proportion of non-retired sample with successful outcome	Metric	99.99 = Cannot be determined	NRPROP
When significance test information is reported			
68. t-value	Metric	99.99 = Cannot be determined	T-VALUE

69. F-value (df for the numerator must equal 1)	Metric	99.99 = Cannot be determined	F-VALUE
70. Chi-Square value (df = 1)	Metric	99.99 = Cannot be determined	CHISQUARE
Statistical Power Calculation			
71. Power calculation presented in study	Metric	99.99 = Cannot be determined	POWER1
72. Statistical Power (post-hoc calculated by rater)	Metric	99.99 = Cannot be determined	POWER2
Calculated Effect Size			
73. Effect size calculated using the procedures outlined in Lipsey and Wilson (2001). (Report two decimals with an algebraic sign in front)	Metric	(plus) + = difference favours retired group (minus) - = difference favours non-retired group +99.99 = cannot be determined	ES
74. Overall confidence rating in effect size computation		1 = highly estimated (have N and crude p-value only, such as p<.10 and must reconstruct via rough t-test equivalence) 2 = Moderate estimation (have complex but relatively complete statistics, such as multifactor ANOVA as basis for estimation) 3 = Some estimation (have unconventional statistics and must convert to equivalent t-values; or have conventional statistics but incomplete, such as exact p-level) 4 = Slight estimation (must use significance testing statistics, rather than descriptive statistics, but have complete statistics of the conventional sort) 5 = No estimation (have descriptive data such as means, standard deviations, frequencies, proportions, etc, and can calculate the effect size directly)	CONES

Appendix 5

Risk of Bias Assessment Forms

1. Cohort Studies

Reference: **ID: CO**

Domain	Instructions	Judgement: +; -; ? (cannot tell)	Support for Judgement
External Validity (the extent to which the result of the study provides a correct basis for applicability to other circumstances)			
1. Representativeness of the sample for the target population	Study receives a + if the sample included all eligible workers/retirees over a defined period of time, or in a defined catchment area or context/setting (e.g. employer, school graduates), or if sample is a random or systematic sample.		
2. What percentage of selected individuals agreed to participate (at baseline)	Study receives a + if the percentage participation was 80% or more of the initially recruited/selected sample		
Internal validity (the extent to which systematic error or bias is minimised in a study)			
Performance bias (systematic difference in measurement of exposure)			
3. How was retirement status ascertained?	Study receives a + if participants' retired status is ascertained via independent record(s) (e.g. employment, medical or social security records); or, if self-assessed retired status is confirmed by independent records.		
Detection bias (systematic differences in outcome assessment)			
4. How was depression status ascertained?	Study receives a + if the measure of depression was adequate (valid) for the population		
Attrition bias (systematic differences in follow-up)			
5. All participants followed up for the same amount of time	Study receives a + if all participants were followed-up for the same amount of time		
6. Percentage of participants included in the final analysis (i.e. adequacy of follow-up of cohorts)	Study receives a + if the percentage participants in the final analysis was 80% or more (of baseline participants); or, if a full description of those lost-to-follow-up was not suggestive of bias.		
Selection bias/control of confounding (systematic differences in comparison groups)			
7. Clear worker/retired status at start of the study	Study receives a + if no participant was retired at the start of the study; or, if those participants that were retired at the start of the study can be clearly identified throughout the analysis.		

8. Pre-specified potential confounding factors: a. Age (at baseline) b. Gender c. Region/country of origin d. Socio-economic status e. Physical health status f. Marital status	This is a list of potential confounders (i.e. characteristics of samples or groups). For each of the factors listed (a. to h.), the item receives a + if the factor was balanced between workers and retired groups (10% or less difference) or adjusted for in analysis. An absence of a + indicates that the factor was either not measured, or, if it was, it did not meet the + requirements.	a. b. c. d. e. f.
Reporting bias (systematic differences in the way outcomes or analyses are reported)		
9. All pre-specified primary outcomes are reported	Study receives a + if the pre-specified outcome related to depression is reported	
10. The outcome of interest in the review is reported completely so that it can be entered in a meta-analysis?	Study receives a + if the outcome related to depression is reported completely (including descriptive statistics) so that it can enter in a meta-analysis	
11. There is no evidence that an outcome that was partially reported or not reported was statistically insignificant or low in magnitude	Study receives a + if there is no evidence that an outcome that was partially reported or not reported was statistically insignificant or related to an effect low in magnitude	
12. There is no evidence that multiple adjusted analyses were carried out but only one or a subset of one was fully reported	Study receives a + if there is no evidence that multiple adjusted analyses were carried out but only one or a subset of one was fully reported	
13. There is no evidence of methods being applied to deal with missing data that were not pre-specified	Study receives a + if there is no evidence of methods being applied to deal with missing data that were not pre-specified	

2. Cross-sectional studies

Reference: **ID: CS**

Domain	Instructions	Judgement: +; - ; ? (cannot tell)	Support for Judgement
External Validity (the extent to which the result of the study provides a correct basis for applicability to other circumstances)			
1. Representativeness of the sample for the target population	Study receives a + if the sample included all eligible workers/retirees over a defined period of time, or in a defined catchment area or context/setting (e.g. employer, school graduates), or if sample is a random or systematic sample.		
2. What percentage of selected individuals agreed to participate (at baseline)	Study receives a + if the percentage participation was 80% or more of the initially recruited/selected sample		
Internal validity (the extent to which systematic error or bias is minimised in a study)			
Performance bias (systematic difference in measurement of exposure)			
3. How was retirement status ascertained**?**	Study receives a + if participants' retired status is ascertained via independent record(s) (e.g. employment, medical or social security records); or, if self-assessed retired status is confirmed by independent records.		
Detection bias (systematic differences in outcome assessment)			
4. How was depression status ascertained?	Study receives a + if the measure of depression was adequate (valid) for the population		
Attrition bias (systematic differences in follow-up)			
5. Percentage of participants included in the final analysis	Study receives a + if the percentage participants in the final analysis was 80% or more (of baseline participants); or, if a full description of those lost-to-follow-up was not suggestive of bias.		
Selection bias/control of confounding (systematic differences in comparison groups)			
6. Pre-specified potential confounding factors: a. Age (at baseline) b. Gender c. Region/country of origin d. Socio-economic status e. Physical health status f. Marital status	This is a list of potential confounders (i.e. characteristics of samples or groups). For each of the factors listed (a. to h.), the item receives a + if the factor was balanced between workers and retired groups (10% or less difference) or adjusted for in analysis. An absence of a + indicates that the factor was either not measured, or, if it was, it did not meet the + requirements.	a. b. c. d. e. f.	
Reporting bias (systematic differences in the way outcomes or analyses are reported)			
7. All pre-specified primary outcomes are reported	Study receives a + if the pre-specified outcome related to depression is reported		

8. The outcome of interest in the review is reported completely so that it can be entered in a meta-analysis?	Study receives a + if the outcome related to depression is reported completely (including descriptive statistics) so that it can enter in a meta-analysis
9. There is no evidence that an outcome that was partially reported or not reported was statistically insignificant or low in magnitude	Study receives a + if there is no evidence that an outcome that was partially reported or not reported was statistically insignificant or related to an effect low in magnitude
11. There is no evidence that multiple adjusted analyses were carried out but only one or a subset of one was fully reported	Study receives a + if there is no evidence that multiple adjusted analyses were carried out but only one or a subset of one was fully reported
12. There is no evidence of methods being applied to deal with missing data that were not pre-specified	Study receives a + if there is no evidence of methods being applied to deal with missing data that were not pre-specified

Appendix 6

Characteristics of Included and Excluded Studies

1. Included studies

Table 1 – List of Included Studies

Full Reference	Study ID	• Author contacted • Outcome
Cross-Sectional Studies – CS		
Bosse, R., Aldwin, C.M., Levenson, M.R., Ekerdt, D.J. (1987). Mental health differences among retirees and workers: findings from the Normative Aging Study. Psychology and Aging, vol. 2(4), p. 383-389.	CS1	• No (author deceased) • Some statistics missing; ES can be estimated
Drentea, P. (2002). Retirement and mental health. Journal of Aging and Health, vol. 14, p. 167-194.	CS3	• Yes • Requested data provided
Steptoe, A., Demakakos, P., Oliveira, C. (2012). Chapter 4: The psychological well-being, health and functioning of older people in England. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5), p. 98-182. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5	CS4	• Yes • Requested data provided
Cohort Studies – CO		
Reitzes, D.C., Mutran, E.J., Fernandez, M.E. (1996). Does retirement hurt wellbeing? Factors influencing self-esteem and depression among retirees and workers. The Gerontologist, vol. 36(5), p. 649-656.	CO4	• No • Complete statistics
Kubicek, B., Korunka, C., Raymo, J.M., Hoonakker, P. (2011). Psychological wellbeing in retirement: the effects of personal and gendered contextual resources. Journal of Occupational Health Psychology, vol. 16(2), p. 230-246.	CO6	• No • Complete statistics
Westerlund, H., Vahtera, J., Ferrie, J.E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M., Zins, M., Kivimaki, M. (2010). Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. British Medical Journal, BMJ 2010:341:c6149.	CO8	• Yes • No reply • Incomplete but usable statistics
Clarke, P., Marshall, V., House, J., Lantz, P. (2011). The social structuring of mental health over the adult life course: advancing theory in the sociology of aging. Social Forces, vol. 89(4), p. 1287-1313.	CO9	• Yes • Some requested data provided; unclear but usable statistics;
Calvo, E., Sarkisian, N., Tamborini, C.R., (2013). Causal effects of retirement timing on subjective physical and emotional health. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, vol. 68(1), p. 73–84	CO10	• No • Complete statistics

1.1. Characteristics of included studies[13]

1.1.1 Cohort studies

CO4 Reitzes et al. (1996)

Methods	• Cohort, prospective, telephone interview-based study, carried out between 1992 and 1994. Participants followed-up in 6 months' intervals over two years. Only data from two measurement points (initial and final) are used in this analysis. • DVs: depression measured by CES-D 20 (full scale), keeping the original response scale properties (alpha reliability .89 and .85 in the two waves); self-esteem measured by Rosenberg's Self Esteem Scale. • IVs/predictor variables: age; gender; employment status self-assessed (retired, working); poor health/functional limitation; ethnicity; marital status; income; education; occupation; worker commitment; worker identity.
Participants	• Analysis used data from the Carolina Health and Transitions Study (USA), a geographically defined, gender weighed sample of full-time workers between the ages of 58 and 64. Sampling frame was local motor vehicles register/database, which is estimated to include 'over 80% of the entire population in the age group'. 62% of all eligible individuals consented to take part. • The sample included 826 participants at Time 1, with a loss of 11% of participants at Time 2 (N at Time 2=737). Between Time1 and 2, of the 737 respondents who provided usable follow-up data, 438 were still employed full-time and 299 retired.
Exposure	• Retirement, self-assessed. • All participants enrolled in the study while working; some participants transitioned to retirement during follow up period. • Follow-up equal for all participants
Outcomes	• Analysis of difference pre and post-retirement: • Raw sample means indicate that participants reported fewer depressive symptoms post-retirement (pre-retirement M=5.705, SD=6.576; post-retirement M= 4.673, SD=5.557); difference is reported as not significant, thought data is not presented. • Multiple regression analysis: For participants who retired, depression was stable, and depression scores at Time 1 had a modest effect on depression scores at Time 2 (β= .380). For participants who continued working, depression scores at Time 1 had a low effect on depression scores at Time 2 (β= .192), no changes reported either.
Notes	• Sample is not likely to be representative of local retiree and worker populations. • No data is presented in relation to the mean age of sample, though age is included in regression models. • The use of the intact and unchanged full CES-D scale allows greater confidence in the validity of the measure. • Outcomes are reported in very unclear way; difference in means and respective tests are not clearly distinguished between-groups or within-groups differences. Multiple regression analyses also unclearly reported.

CO6 Kubicek et al. (2011)

Methods	• Secondary analysis of cohort, prospective, interview-based (mail and telephone) study, carried out between 1957 and 2004. Participants followed-up in 1975, 1993, and 2004. Only data from the two later measurement points (1993 and 2004) are used in this analysis. • DVs: depression measured by CES-D 20, full scale (alpha reliability .87 and .86), though not keeping original response scale properties; psychological well-being measured by short version of Ryff's Psychological Wellbeing Scale. • IVs/predictor variables: employment status, gender, pre-retirement resources (closeness to spouse, social contacts, social group involvement, net assets, wages, flexible goal adjustment, tenacious goal pursuit, importance of work, job satisfaction, poor health, spouse's poor health, monotonous work). • Control variable: time spent in retirement.

[13] Format adapted from Siegfried, Muller, Deeks, Volmink (2009); ordered by study ID.

Participants	• Analysis used data from the Wisconsin Longitudinal Study (USA), cohort study of a random sample of 10,317 men and women. Sampling frame was the registry of graduates from Wisconsin high schools in 1957. • As a secondary analysis of data, the sample analysed included 1609 participants who satisfied pre-established selection conditions: completed both 1994 and 2003 surveys; retired between 1993 and 2004; had a spouse in 1993; had no missing data in the self-report of depressive symptoms. • Information on the comparison between full sample and selected sample indicated significant differences only on two financial indexes (wages and assets). • Loss of data between two time points is considerable (45.5%), and not justified. • The sample age mean at Time 1 was 53 (SD not reported), and at Time 2 was 65 (SD= .66).
Exposure	• Retirement, self-assessed. • All participants enrolled in the study while working; all selected participants transitioned to retirement during follow up period. • Follow-up equal for all participants
Outcomes	• Analysis of difference pre and post-retirement: Raw sample means (1993 and 2004) indicate a decline in self-reported depressive symptoms post-retirement (pre-retirement M= 9.62, SD=7.75, post-retirement M= 7.68, SD=7.07, t(1608)=11.46, p<.01, d=0.26). Also small but significant gender difference, with women reporting more frequent symptoms of depression than men (t(1607)=4.68, p< .01, d=0.23) • Multiple regression analysis: Gender invariant and gender specific models estimated; outcomes suggest: fewer pre-retirement social contacts increased depressive symptoms in female retirees, but not male; greater pre-retirement social group involvement had a detrimental effect on depressive symptoms in retirement for males, but not for females; pre-retirement job satisfaction increased report of depressive symptoms in men, but not in women.
Notes	• Sample is not likely to be representative of general retiree and worker populations. • Selection bias is likely to be a concern. No information is provided response rate or self-selection of participants. As a secondary analysis, the sample was selected to minimise missing data, and this introduces additional selection bias. • Loss of data between two time points is considerable (45.5%), and not justified. • The study uses well-established, valid measure of depression (CES-D full scale), although though the original study used an extended response scale (0-8, instead of the original 0-3). These responses were then re-coded into the original scale in this analysis,

CO8 Westerlund et al. (2010)

Methods	• Secondary analysis of cohort, prospective, panel survey, based on annual questionnaires and register data. The analysis related to depression outcomes used data collected between 1996 and 2005. Initial (-7), transition into retirement (-1 to +1), and final points (+7), are used as measurement points use in this analysis. • DVs: depression measured by CES-D, full scale (20 items, alpha reliability value not provided); checklist of chronic diseases (heart, stroke, respiratory, diabetes); mental and physical fatigue, measured by *ad hoc* single question. • IV: retirement status measured by work-place records • Covariates: demographic variables, marital status, occupational grade/category (managerial, technical, etc.).
Participants	• Analysis used data from the GAZEL cohort, this is a study cohort established in 1989, comprising of workers of the French national gas and electricity company (EDF-GDF). At baseline 20624 participants consented to take part. • As a secondary analysis, the sample used included participants who retired between 1990 and 2006, but excluded those who retired on health grounds. • Data included in the analysis was chosen on the basis of completeness of measures, with final N=1404. There is an indication of loss of data after this (in results tables), but this is not fully justified or quantified. • The sample age distribution is unclearly presented, and the gender proportion is heavily skewed, with 80% of men in the analysis sample.
Exposure	• Retirement both self-assessed and objectively assessed (independent registers). • All participants enrolled in the study while working; participants transitioned to retirement during

	follow up period, as statutory retirement. • Follow-up equal for all participants.
Outcomes	• To examine changes in the likelihood of depressive symptoms, repeated measures logistic regression analysis was used, and annual odds calculated. • Odds-ratio of for depressive symptoms were: pre-retirement (year -7 to -1, 1.06, 95% CI 0.91 to 1.25), at transition to retirement (year -1 to year 1, 0.60, 95% CI 0.53 to 0.67) and post-retirement (year +1 to +7, 1.05, 95% CI 0.92 to 1.19). • Authors conclude that retirement was associated with a substantial decrease in the prevalence of mental fatigue, physical fatigue and depressive symptoms.
Notes	• Overall distribution and procedures for handling missing data are reported, but this is not entirely clear. • Efforts were made to control for endogeneity bias, by focusing on participants who retired not for physical health reasons. • As a secondary analysis, the sample was selected to minimise missing data, and this introduces additional selection bias. • The study uses well-established, valid measure of depression (CES-D 20).

CO9 Clarke et al. (2011)

Methods	• Secondary analysis of data from Americans' Changing Lives survey, which is a stratified, multi-stage area probability sample of non-institutionalized adults age 25 and older, living in the coterminous United States, followed over a 15-year period (1986 to 2002). • Time points in this analysis were 1989 and 2001. • DVs: depression measured by CES-D, 7 items; the scale was heavily transformed and shortened and the authors claim to have split the scale into two dimensions (depressive affect and somatic symptoms), which were then averaged to '*produce an index of depressive symptoms ranging from 1 to 3, and log transformed to correct positive skew'*. No reliability values provided. • IVs: employment status (retired, working, unemployed, 'homemaker'), time varying factors (wealth, marital status, physical health); time invariable (gender, ethnicity, education, socio-economic position at young adulthood).
Participants	• As a secondary analysis, the sample used included 3616 participants at baseline (1986); with 1549 who were working at first measurement point, • Loss of overall data between the two time points (Time 2 N= 1787), is considerable (50.6%), and not fully justified. • The overall sample age mean at Time 1 was 47.05, with a proportion of 37.52% male participants.
Exposure	• Retirement, self-assessed. • All participants enrolled in the study while working; participants transitioned to retirement during follow up period; however, only 352 participants made the transition to retirement during follow up. • Follow-up equal for participants included in the analysis (i.e. who made the transition to retirement), though this is not entirely clear.
Outcomes	• Analysis of difference pre and post-retirement: Raw sample means (1989 and 2001) indicate a slight reduction, though not significant, in self-reported depressive symptoms post-retirement (pre-retirement M= 0.305, SD=0.343, post-retirement M=0.301, SD=0.312). • The authors used growth curve models to examine life-course trajectories of depressive symptoms, with separate models for men and women: authors report that self-defined retirement was not associated with an increase in depressive symptoms.
Notes	• Authors report that the full sample is representative of the USA population. • Overall procedures for handling missing data are very unclear: loss of data (50.6%) of initial sample is not accounted for clearly; • Further, the inclusion or exclusion of participants in specific analyses or specific variables (in this case transition to retirement) is also obscured (only 352 participants made the transition to retirement at the final Time-point, but this is not clear in data presentation). • The study uses a highly transformed version of the CES-D, which compromises the measurement properties of this version and its comparability with other studies.

CO10 Calvo et al. (2013)

Methods	• Secondary analysis of cohort, prospective, interview-based panel survey, carried out between 1992 and 2010. Initial and final measurement points use in this analysis. • DVs: depression measured by CES-D, 8 items, scale reversed (i.e. higher values indicate better health), and dimensional scale transformed into yes/no response (scale range 0-8) (alpha reliability value not provided); subjective physical health measured by *ad hoc* single question (health rating poor-excellent) • IVs: employment status (retired vs working), retirement timing (measured by interactions of retirement status with current age, divided by 10 and centred at 60) • Instrumental variables: (to control for endogeneity bias, variation in the timing of retirement that is exogenous to health.) changes to Social Security's full retirement age, unexpected early retirement window offers. • Control variables: time variable (wealth, income, marital status, spouse employment status); time invariable (gender, ethnicity, education, occupation type).
Participants	• Analysis uses data from the Health and Retirement Study (USA), using a random, multistage, stratified sample of older Americans and spouses. • As a secondary analysis, the sample used included 6275 participants who met pre-defined criteria: born between 1931 and 1941; enrolled in 1992; who were working full or part-time at first measurement point. • Loss of data between two time points is considerable (21.31%), and not justified. • The sample age mean at Time 1 was 55.67 (SD=3.12), and at Time 2.
Exposure	• Retirement both self-assessed and objectively assessed (independent indicators). • All participants enrolled in the study while working; participants transitioned to retirement during follow up period, except 333 participants who still worked in 2010. • Follow-up equal for all participants
Outcomes	• Analysis of difference pre and post-retirement: Raw sample means (1992 and 2010) indicate an increase in self-reported depressive symptoms post-retirement (pre-retirement M= 7.41, SD=1.11, post-retirement M=6.86, SD=1.9), scale inverted. • Multiple regression analysis: Analysis of the effects of retirement timing on depression are limited to the effects of the portion of variation in retirement timing related to the instrumental variables; long term and short term, fixed and random effects regression models for self-reported physical health and symptoms of depression presented. Outcome variables have significantly non-normal distributions. Negative coefficients for retirement status were found in all four models. Suggests that retirement has 'a negative effect on emotional health (i.e. promotes depressive symptoms) at 60 years'. However, significant positive interaction between retirement status and age also found, and 'the negative effect is reduced if a person retires later'.
Notes	• Authors report that the full sample is representative of the age, gender and ethnic group distributions of the population (25 years old and older) living in the United States in 1986. • Overall distribution and procedures for handling missing data are not clearly reported; loss of data (21.31% of initial sample is also not accounted for. • Efforts to control for endogeneity bias, by focusing on variation in retirement timing that is related to the two instrumental variables. • As a secondary analysis, the sample was selected to minimise missing data, and this introduces additional selection bias. • The study uses well-established, valid measure of depression (CES-D 8). However, the transformation of a dimensional response format, the shortening of the scale to selected items, and missing reliability value raise question related to measurement comparability with other studies.

1.1.2 Cross-sectional studies

CS1 Bosse et al. (1987)

Methods	• Cross-sectional, mail survey study (82.2% response rate). • DVs: depression, measured by SCL-90-R, depression subscale (full-scale, keeping the original response scale properties); physical health measured by *ad hoc* single question on presence of health problems; • IVs: Retirement status self-assessed on 6 point categorical index (3 retirement categories, 3 working categories; unemployed excluded from study); time elapsed since retirement; timing of retirement.
Participants	• Participants recruited from the sample of the Boston Veterans Administration Normative Aging Study (cohort study of community dwelling individuals, born between 1884 and 1945). • The self-selected (consenting to respond) 1513 older men (age mean 61), included 673 retired (mean age 66.35), and 840 working participants (mean age 56). • The researchers note that the socioeconomic, educational and occupational levels are 'slightly higher than the national population'.
Exposure	• Retirement, self-assessed
Outcomes	• Retired sample time elapsed since retirement (M=6.1 and SD=4.6 years). Mean age of the workers lower than mean age of the retirees ($F(3, 1506) = 530.82$, $p < .001$), but overlap in the age range between the workers (39-80) and the retirees (46-88). 5% of workers classified as late workers, whereas 13% of retirees classified as late retirees (retiring after 65). • Analysis of difference between workers and retirees: Retirees reported more depressive symptoms (adjusted means) than workers (retirees M=.36, workers M=.28; $F(1, 1449)=14.62$, $p<.001$); the difference was still significant controlling for physical health status (retirees M=.34 and workers M=.29; $F(1, 1407)=6.10$, $p<.001$). Significant group differences in reported depressive symptoms reported. • Regression analysis revealed no significant association between time elapsed since retirement and psychological symptoms, but no data provided. Timing of retirement did affect the association between retirement and psychological symptoms; early and late retirement associated with 'more emotional distress', but no data provided.
Notes	• Entirely male samples are not representative of retiree or worker population. • Self-selection into study raises concerns, though authors note that the group of respondents compared to non-respondents is 'reasonably' similar. • The measure of depression used is valid for this population. • Potential confounders: analysis of variance did not control for differences in age between the two samples, though these were found to be significant; previous mental and physical health status not controlled for (potential endogeneity bias); differences in relation to other samples' characteristics not reported, though both samples are selected from the same population.

CS2 Drentea (2002)

Methods	• Cross-sectional analysis, based on telephone survey study (response rate 71.6%). • DVs: depression measured by 7 items from CES-D (alpha reliability .82), with the scale transformed from the original 0-3 response format to 0-7 response format; distress measured by 13 items of CES-D (alpha reliability of .87); anxiety measured by 3 items of CES-D (alpha reliability .79); positive affect measured by *ad hoc* question; sense of control measured by Mirowsky-Ross scale. • IVs: age; gender; employment status self-assessed including 6 categories (retired, full-time worker, part-time worker; homemaker; unable to work; in education; unemployed); work and activity characteristics.
Participants	• Analysis used data from the Aging, Status and Sense of Control Survey (ASOC), USA national, telephone probability sample survey (age range 18-95). No information on sampling frame. • The full self-selected (consenting to respond) survey sample was used (N=2587, age M=47.6, SD=17.68). The retired group included 672 participants and the non-retired (collated) group included 1915 participants. • The mean ages of the retired and non-retired groups are not reported. No information is given on sample equivalence.
Exposure	• Retirement, self-assessed
Outcomes	• Analysis of difference between retirees and non-retirees:

	• Raw group means indicate that retirees reported fewer depressive symptoms (retirees M=0.79, SD=1.21; non-retired M=0.96, SD=1.33); difference significant t(1282.479)=2.986, p=.003, equal variances not assumed; • However, multiple regression analysis reports no association between retirement and depressive symptoms, but no data is presented.
Notes	• As a national probability sample, this is likely to be representative of retiree or worker populations. Self-selection into study, however, raises concerns. Sample lower age range is younger compared to other studies, as is the sample's mean age. • Potential confounders: samples are unbalanced in size, no information is given in relation to the equivalence of samples; the non-retired sample (as it is used in the analysis of group differences) is a composite of 5 different work statuses, including unemployed participants. This is theoretically likely to influence outcomes (i.e. increase the report of depressive symptoms in the non-retired group) • The CES-D measurement properties are likely to be affected by splitting of a global measure do depression into ad-hoc indices, and the transformation of the scale response format. • No information how age, gender and other variables were considered in the multiple regression analysis of depression outcomes.

CS2 Steptoe et al. (2012)

Methods	• Cross-sectional analysis, based on cohort, interview-based study (study response rate 68.7%). • DVs: depression is measured by CES-D, 8 items, with the original dimensional response scale transformed into yes/no response (range 0-8); affective well-being measured by the Life Satisfaction Scale; enjoyment of life, measured by CASP-19 questionnaire; positive affect, measured by items of the Positive and Negative Affect Scale; eudemonic well-being, measured by CASP-19 • IVs (in the analysis of psychological wellbeing measures): labour market participation (working full-time, working part-time, retired), age, sex, wealth, marital status, volunteering, self-rated health, limiting long-standing illness, disability, health behaviours, cognitive and physical functioning, coronary heart disease.
Participants	• Analysis used data from the 5th wave (2010-11) of the English Longitudinal Study of Ageing (ELSA), UK probability sample survey (age range 50-80+). The original ELSA sample was drawn from households previously responding to the Health Survey for England (1998, 1999 and 2001). • The full self-selected (consenting to respond) sample used in this analysis included 8414 participants, both men and women, (age M=67.76, SD=9.5), with 5680 retired individuals (age M=71.32, SD=9.68), and 2734 non-retired individuals (age M=59.23, SD=4.6). • No information is given on sample equivalence.
Exposure	• Retirement both self-assessed and objectively assessed (independent indicators).
Outcomes	• Analysis of difference between retirees and non-retirees: raw group means suggest that retirees report more depressive symptoms compared to workers (retired group M=1.77, SD=2.08, worker group M=1.16, SD=1.77). No information is given in relation to the statistical significance of this difference. Reports that this difference is increased in younger retired individuals (age 52-59), no data provided.
Notes	• As a national probability sample, this is likely to be representative of retiree or worker populations. • Retirement assessed both subjectively and independently. • Potential confounders: Samples are unbalanced in size, no information is given in relation to the equivalence of samples.

2. Excluded studies due to lack of data

Table 2 – List of excluded studies due to lack of data (ordered by study ID)

Full Reference	Study ID	• Author contacted • Outcome • Reason for exclusion (Final)
Cross-Sectional Studies - CS		
Jang, S-N., Cho, S-I., Chang, J., Boo, K., Shin, H-G., Lee, H., Berkman, L.F. (2009). Employment status and depressive symptoms in Koreans: results from a baseline survey of the Korean longitudinal study of aging. Journal of Gerontology Series B: Psychological Sciences and Social Sciences, vol. 64(5), p. 677-683.	CS2	• Yes • No reply • All usable statistics missing
Cohort Studies - CO		
Kim, J.E., Moen, P. (2002). Retirement transitions, gender and psychological wellbeing: a life course ecological model. Journals of Gerontology Series B: Psychological Science and Social Sciences, vol. 57B(3), p. 212-222.	CO1	• Yes • No reply • Incomplete/unclear usable statistics
Laaksonen, M., Metsä-Simola, N., Martikainen, P., Pietiläinen, O., Rahkonen, O., Gould, R., Partonen, T., Lahelma, E. (2012). Trajectories of mental health before and after old-age and disability retirement: a register-based study on purchases of psychotropic drugs. Scandinavian Journal of Work and Environmental Health, vol. 38(5), p. 409-417.	CO2	• Yes • No reply • All usable statistics missing
Oksannen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjosten, N., Virtanen, M., Kawachi, I., Kivimaki, M. (2011). Is retirement beneficial for mental health? Antidepressant use before and after retirement. Epidemiology, vol. 22(4), p.553-559.	CO3	• Yes • No reply • Incomplete/unclear usable statistics
Midanik, L.T., Soguikian, K., Ranson, L.J., Tekawa, I.S. (1995). The effect of retirement on mental health and health behaviours: the Kaiser Permanente Retirement Study. Journals of Gerontology Series B, Psychological and Social Sciences, vol. 50B(1), p. S59-S61.	CO5	• Yes • Author retired; no further data, dataset not available • Incomplete/unclear usable statistics
Lindeboom, M., Portrait, F., van den Berg, G.J. (2002). An econometric analysis of the mental-health effects of major events in the life of older individuals. Health Economics, vol. 11, p. 505-520.	CO7	• Yes • No further data, dataset not available to author • Incomplete/unclear usable statistics
Brugiavini, A., Croda, E., Dewey, M. (2008). Retirement and mental health. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension, p. 247-254. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf Additional information sought from: De Luca, G., Rossetti, C. (2008). Sampling Design and Weighting Strategies in the Second wave of SHARE. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension, p. 333-338. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf Zamarro, G., Meijer, E., Fernandes, M. (2008). Labor force participation and retirement. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension, p. 48-55. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-	CO11	• Yes • No reply • Incomplete/unclear usable statistics

project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf
Brugiavini, A., Pasini, G., Peracchi, F. (2008). Exits from the labour force. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension, p. 206-214. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf
Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., Bula, C., Reisches, F., Wancata, J., Ritchie, K., Tsolaki, M., Mateos, R., Prince, M. (2008). Ascertaining late-life depressive symptoms in Europe: an evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE project. International Journal of Methods in Psychiatric Research, vol. 17(1), p. 12-29.

Table 3 – Summary characteristics of excluded cohort studies due to lack of data

	Design						*Sample/Sampling*						*Measurement*		*Outcomes*
Study ID	*Study design*	*Time points*	*Interval*	*2^{dary} analysis*	*Analysis design*	*Discipline*	*Origin*	*Method*	*Total N T1*	*Data loss %*	*Male % (T1)*	*Age (M) T1*	*Measure*	*Other DVs*	*Summary of outcomes re. depression*
CO1 Kim and Moen (2002)	Cohort, prospective	2	1994-1996	Yes	Multiple Regression	Psychology	USA	Random Non-general population	458	6	61.57	60	CES-D, 12 items	. Morale	1. Positive relationship between retirement length and (increased) depressive symptoms for men, but not for women
CO2 Laaksonen et al. (2012)	Cohort, prospective + retrospect.	41	1995-2009	Yes	Growth curve modelling	Public Health	Finland	Random Non-general population	4456	No data	No data	No data	Purchase of psychotrop. meds.	. None	1. No relationship between retirement and purchase of psychotropic meds.
CO3 Oksannen et al. (2011)	Cohort, prospective + retrospect.	9	1995-2004	Yes	Multiple regression	Public Health	Finland	Random Non-general population	7138	0.9	24	No data	Purchase of antidepres. meds.	. Purchase of diabetes meds.	1. Reduction of purchase of antidepres. meds in retirement transition 2. Similar trend in post-retirement for women only.
CO5 Midanik et al. (1995)	Cohort, prospective	2	1985-1987	Yes	Multiple regression	Medicine	USA	Convenient, self-selected Non-general population	595	0	57.47	62.05	CES-D, 20 items	. Stress . Coping . Health behaviours	1. No effect of retirement on relative risk of depression
CO7 Linbeboom et al. (2002)	Cohort, prospective	3	1992-1999	Yes	Multiple regression	Economics	Holland	Random multistage, stratified General population	2253	27	50	69.9	CES-D, 20 items	. Cognitive function . Physical functioning . Chronic conditions	1. Unclear relationship between retirement and depressive symptoms
CO11 Brugiavini et al. (2008)	Cohort, prospective	2	2004-2006	Yes	Multiple regression	Economics	Europe (multi-country)	Random General population	4673	No data	85.66	No data	EURO-D, 12 items	. Physical health	1. No relationship between retirement and symptoms of depression

Key: Time Points = number of measurement time points used in the analysis; Interval = interval between measurement time points ; Secondary Analysis = is the record a secondary analysis of previously collected data?; Analysis design = the design of the statistical analysis performed on data; Discipline = scientific area of the study's first author or affiliated department/institution; Origin = geographical origin of the sample; Method = sampling method and target population (general or non-general); T1 = Measurement time point 1; Data loss = percentage of data lost between T1 and T2 (measurement time point 2); Measure = Measure of depression used; Other DVs = other outcome variables included in the analysis

Table 4 – Summary characteristics of excluded cross-sectional studies due to lack of data

	Design					*Samples/Sampling*						*Measure*		*Outcomes*
Study ID	*Study design*	*2^{dary} Analysis*	*Analysis design*	*Compare.*	*Discipline*	*Origin*	*Method*	*N*	*Equivalent?*	*Male %*	*Age (M)*	*Measure*	*Other DVs*	*Summary of outcomes re. depression*
CS3 Jang et al. (2009)	Cohort, prospective	Yes	Cross-sectional; multiple regression	Retired vs Workers	Public Health	China	Random multistage, stratified. General population	N:8821 R:1675 C:3652	No data Samples from same population	44.29	61.89	CES-D, 10 items	None	1. Positive independent association between retirement and increase in depressive symptoms.

Key: 2^{adry} Analysis = is the record a secondary analysis of previously collected data?; Analysis design = the design of the statistical analysis performed on data; Compare. = comparison between what groups; Discipline = scientific area of the study's first author or affiliated department/institution; Origin = geographical origin of the sample; Method = sampling method and target population (general or non-general); R = retired sample; C = comparison sample (non-retired or working); Equivalent? = were equivalence tests performed on the two samples; Measure = Measure of depression used; Other DVs = other outcome variables included in the study.

Appendix 7

Effect size estimation formulas and procedures

1. Effect size estimation formulas

1.1 Cross-sectional studies

1.1.1 Standardised Mean Difference (d)

The following basic formula was used[14], with accompanying Standard Error (SE_d) and Variance (V_d) (Borenstein at al, 2009):

$$d = \frac{\overline{X_1} - \overline{X_2}}{S_{within}}$$

Where $\overline{X_1}$ is the mean for the Retired group (Group 1), $\overline{X_2}$ is the mean for Non-retired group (Group 2), and S_{within} is the within groups standard deviation, defined as:

$$S_{within} = \sqrt{\frac{(n_1 - 1)S_1^2 \quad + \quad (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

With S_1 as the standard deviation for the Retired group (Group 1) and S_2 is the standard deviation for Non-retired group (Group 2), n_1 is the number of participants in the Retired group and n_2 is the number of participants in the Non-retired group.

$$V_d = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2\,(n_1 + n_2)}$$

$$SE_d = \sqrt{V_d}$$

1.1.2 Unstandardised (raw) mean difference (D)

The following formulas were used for this purpose (Borenstein et al., 2009):

$$D = \bar{X}_1 - \bar{X}_2$$

$V_D = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$, not assuming the two populations standard deviations to be the same.

$$SE_D = \sqrt{V_D}$$

[14] Given the use of large samples in the included studies, Hedges's correction for upward bias (g) was not required.

1.1 Cohort studies

1.2.1 Standardised mean difference

The following formulas were used (Borenstein et al., 2009):

$d = \frac{\overline{Y_2} - \overline{Y_1}}{S_{within}}$

$S_{within} = \frac{S_{diff}}{\sqrt{2(1-r)}}$

Where $\overline{Y_1}$ is the mean at Time 1, $\overline{Y_2}$ is the mean at Time 2, S_{diff} is the standard deviation of the difference in scores, and S_{within} is the standard deviation within groups and *r* is the correlation between Time 1 and Time 2 scores. According to Lipsey and Wilson (2001), the correlation value is commonly missing from reports and is likely to be found in external sources (e.g. test-retest reliability value), but caution should be taken when using '*very uncertain estimates of the correlation*' (p. 43). As described in the results section, a significant proportion of studies used modified versions of the scales referenced, implying even greater uncertainty in the use of literature-based reliability estimates.

$V_d = \left(\frac{1}{n} + \frac{d^2}{2n}\right) 2\,(1\text{-}r)$

$SE_d = \sqrt{V_d}$

Where *n* is the number of pairs.

1.2.2 Unstandardised (raw) mean difference (D)

Defined as above.

2. Further effect size estimation procedures

2.1 Bose et al. (1987)

Standard Deviations for each group were missing, as was its whole sample equivalent. Therefore, individual group Standard Deviations (SD) for both the Standardised Mean Difference (*d*) and Raw Mean Difference (*D*) were estimated using the available univariate F-ratio data as a starting point (values in brackets). Higgins and Deeks (2011, Section 7.7.3.3) provide a 2-step procedure for the estimation of independent group SDs using:

a) the square root of the univariate F-value as an acceptable approximation of a t-value

$t = \sqrt{F}$; *(3.8236)*

b) from that, estimating the standard error (SE), where the t-value is the ratio of the difference in means (MD) to the SE of the difference in means; and finally estimating SDs

$SE = \frac{MD}{t}$; *(0.0209)*

c) and finally, from the SE, estimating SDs for each group using the Lipsey and Wilson's formula (2001:200), where this s_1 is the estimated standard deviation for the Retired group and s_2 is the estimated standard deviation for the Non-retired group:

$s_1 = \text{se}\sqrt{n_1 - 1}$;

$s_2 = \text{se}\sqrt{n_2 - 1}$

2.2 Reitzes et al. (1996), Westerlund et al. (2010), Clarke (2011), Calvo et al. (2013)

The correlation value between Time 1 and Time 2 was missing in both studies. Lipsey and Wilson (2001) advise that, in such cases, this value *'should approximate the test-retest reliability (...) which may be available from other sources'* (Lipsey and Wilson, 2001:43). For Reitzes et al. (1996) and Westerlund et al. (2010), which use the full CES-D scale (20 items) in its original form, the value was obtained from the test-retest reliability value from the original validation studies of the CES-D (Radloff, 1977), using the longest tested time interval (12 months, *r*=0.49).

For Calvo et al. (2013), efforts were made to obtain a specific value for the CES-D 8 scale, as it is used in the Health and Retirement Study (HRS), from the study's documentation on methodology and measurement (e.g. Steffick, 2000). However, this was also missing. Further efforts were made to obtain this value from other publications using the same data set (e.g. Wang, 2007; Gallo, Bradley, Siegel, Kasl, 2000), but none were available. Finally, a wider search revealed that, although (internal) alpha reliability values for the CES-D 8 were widely available (as verified in the present meta-analysis sample and set of excluded studies), the same is not true for its factorial or stability (test-retest) properties (e.g. Van de Velde, Levecque, Bracke, 2009). Therefore, the decision was made to use the same test-retest reliability value as above (*r*=0.49). The same decision was made in relation to Clarke (2011), who uses an idiosyncratic 3 item selection. Although this value will not affect the mean effect size estimate significantly, it will affect the estimation a) of the confidence interval around the mean effect size, and b) of the value of effect size heterogeneity (Lipsey and Wilson, 2001). Therefore, caution was used in the interpretation of these values.

Combined References for the Full Thesis

Aalto, A.M., Elovainio, M., Kivimäki, M., Uutela, A., Pirkola, S. (2012). The Beck Depression Inventory and General Health Questionnaire as measures of depression in the general population: A validation study using the Composite International Diagnostic Interview as the gold standard. *Psychiatry Research*, vol. 197, p. 163-171.

Alavinia, S.M., Burdorf, A. (2008). Unemployment and retirement and ill-health: a cross-sectional analysis across European countries. *International Archives of Occupational and Environmental Health*, vol. 82, p. 39-45

Americans' Changing Lives Survey, http://www.isr.umich.edu/acl/data.htm

Angelini, V., Brugiavini, A., Webber, G. (2009). Aging and unused capacity in Europe: is there an early retirement trap? *Economic Policy*, vol. 24(59), p. 463-504.

Atchley, R.C. (1971). Retirement and leisure participation: continuity or crisis? *The Gerontologist*, vol. 11(1), p. 13-17.

Atchley, R.C. (1979). Issues in retirement research. *The Gerontologist*, vol. 19(1), p. 44-54.

Atchley, R.C. (1982a). Retirement: leaving the world of work. *Annals of the American Academy of Political and Social Science*, vol. 464, p. 120-131.

Atchley, R.C. (1982b). Retirement as a social institution. *Annual Review of Sociology*, vol. 8, p. 263-287.

Atchley, R.C. (1989). A Continuity theory of normal aging. *The Gerontologist*, vol. 29(2), p. 183-190.

Ayalom, L. (2011). The prevalence and predictors of passive death wishes in Europe: a 2-year follow-up of the survey of health, ageing, and retirement in Europe. *International Journal of Geriatric Psychiatry*, vol. 26, p. 923–929.

Bai, A., Shukla, V.K., Bak, G., Wells, G. (2012). *Quality Assessment Tools Project Report*. Ottawa: Canadian Agency for Drugs and Technologies in Health. Retrieved on the 28th of April 2013 from http://www.cadth.ca

Baltes P.B., Staudinger U.M., Lindenberger, U. (1999). Lifespan psychology: theory and application to intellectual functioning. *Annual Review of Psychology*, vol. 50, p. 471-507.

Baltes, P. (1987). Theoretical propositions of life-span developmental psychology: on the dynamics between growth and decline. *Developmental Psychology*, vol. 23(5), p. 611-626.

Barnes, B., Bloor, D., Henry, J. (1996). *Scientific knowledge: a sociological analysis*. London: Athlone Press.

Beck, S.H. (1982). Adjustment to and satisfaction with retirement. *Journal of Gerontology*, vol. 37(5), p. 616-624.

Behncke, S. (2012). Does retirement trigger ill health? *Health Economics*, vol. 21, p. 282-300.

Biswas-Diener, R., Diener, E. (2001). Making the best of a bad situation: satisfaction in the slums of Calcutta', *Social Indicators Research,* vol. 55, p.329-352.

Blazer D.G. (2003). Depression in late life: review and commentary. *Journal of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 58(3), p. 249–265.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R. (2009). *Introduction to Meta-Analysis*. Chichester: John Wiley & Sons, Ltd.

Bosse, R., Aldwin, C.M., Levenson, M.R., Ekerdt, D.J. (1987). Mental health differences among retirees and workers: findings from the Normative Aging Study. *Psychology and Aging*, vol. 2(4), p. 383-389.

Braithwaite, V.A., Gibson, D.M., Bosly-Craft, R. (1986). An exploratory study of poor adjustment styles among retirees. *Social Science and Medicine*, vol. 23(5), p. 493-499.

Brugiavini, A., Croda, E., Dewey, M. (2008b). Retirement and mental health. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 247-254. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

Brugiavini, A., Croda, E., Mariuzzo, F. (2005). Labour force participation of the elderly: unused capacity? In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *Health, Ageing and Retirement in Europe: First Results from the Survey of Health, Ageing and Retirement in Europe*, p. 236-240. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/SHARE_FirstResultsBookWave1.pdf

Brugiavini, A., Pasini, G., Peracchi, F. (2008a). Exits from the labour force. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 206-214. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

Buber, I., Engelhardt, H. (2011). The association between age and depressive symptoms among older men and women in Europe. findings from SHARE. *Comparative Population Studies*, vol. 36(1), p. 103-126.

Bunce, D., Batterhamb, P.J., Mackinnon, A.J., Christensen, H. (2012). Depression, anxiety and cognition in community-dwelling adults aged 70 years and over. *Journal of Psychiatric Research*, vol. 46, p. 1662-1666.

Busseri, M. A., Sadava, S.W. (2011). A review of the tripartite structure of subjective well-being: implications for conceptualization, operationalization, analysis, and synthesis. *Personality and Social Psychology Review*, vol. 15(3), p. 290-314.

Bussing, A., Bissels,T., Fuchs, V., Perrari, K.-M. (1999). A dynamic model of work satisfaction: qualitative approaches. *Human Relations*, vol. 52, p. 999-1028.

Butterworth, P., Gill, S.C., Rodgers, B., Anstey, K.J., Villamil, E., Melzer, D. (2006). Retirement and mental health: analysis of the Australian national survey of mental health and wellbeing. *Social Science and Medicine*, vol. 63, p. 1179-1191.

Buxton, J., Singleton, N., Melzer, D. (2005). The mental health of early retirees: national interview survey in Britain. *Social Psychiatry and Psychiatric Epidemiology*, vol. 40, p. 99-105.

Cahill, K. E., Giandrea, M. D., Quinn, J. F. (2006). Retirement patterns from career employment. *The Gerontologist*, 46(2), p. 514–523.

Calasanti, T. (1996). Gender and life satisfaction in retirement: an assessment of the male model. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 51B(1), p. S18-S29.

Calvo, E., Haverstick, K., Sass, S.A. (2007). *What makes retirees happier: a gradual or 'cold turkey' retirement?* Center for Retirement Research at Boston College Document CRRWP2007-18. Retrieved on the 10th of February 2013 from http://www.bc.edu/crr

Calvo, E., Sarkisian, N. (2011). *Retirement and well-being: examining the characteristics of life-course transitions*. Retrieved on the 10th of February 2013 from www.politicaspublicas.udp.cl

Calvo, E., Sarkisian, N., Tamborini, C.R., (2013). Causal effects of retirement timing on subjective physical and emotional health. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 68(1), p. 73–84

Campion, J., Bhui, K, Bhugra, D. (2012). European Psychiatric Association (EPA) guidance on prevention of mental disorders. *European Psychiatry*, vol. 27, p. 68–80

Carleton, R.N., Thibodeau, M.A., Teale, M.J.N., Welch, P.G., Abrams, M.P., Robinson, T.,. Asmundson, G.J.G. (2013). The Center for Epidemiologic Studies Depression Scale: a review with a theoretical and empirical examination of item content and factor structure. *PLoS ONE*, 8(3): e58067. doi:10.1371/journal.pone.0058067

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., Bula, C., Reisches, F., Wancata, J., Ritchie, K., Tsolaki, M., Mateos, R., Prince, M. (2008). Ascertaining late-life depressive symptoms in Europe: an evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE project. *International Journal of Methods in Psychiatric Research*, vol. 17(1), p. 12-29.

Charles, K.K. (2002). *Is retirement depressing? Labour force inactivity and psychological wellbeing in later life.* Working Paper, National Bureau of Economic Research. Retrieved on the 9th of March 2013 from http://www.nber.org/papers/w9033

Cheshire, H., Hussey, D., Phelps, A., Wood, N. (2012). Chapter 5: Methodology. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), *The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5)*, p. 183-205. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

Christ, S.L., Lee, D.J., Fleming, L.E., LeBlanc, W.G., Arheart, K.L., Chung-Bridges, K., Caban, A.J., McCollister, K.E. (2007). Employment and occupation effects on depressive symptoms in older Americans: does working past the age of 65 protect against depression? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(6), p. S399-S403.

Christ, S.L., Lee, D.J., Fleming, L.E., LeBlanc, W.G., Arheart, K.L., Chung-Bridges, K., Caban, A.J., McCollister, K.E. (2007). Employment and occupation effects on depressive symptoms in older Americans: does working past the age of 65 protect against depression? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(6), p. S399-S403.

Clark, A., Friedman, M.J. (1983). Factor structure and discriminant validity of the SCL-90 in a veteran psychiatric population. *Journal of Personality Assessment*, vol. 47, p. 396-404.

Clark, R. L., Quinn, J. F. (2002). Patterns of work and retirement for a new century. *Generations*, 26(2), p. 17-24.

Clarke, M., Oxman, A.D., Paulsen, E., Higgins, J.P.T., Green, S. (2011). Appendix A: Guide to the contents of a Cochrane Methodology protocol and review. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 25th of September 2012 from www.cochrane-handbook.org.

Clarke, P, Marshall, V.W., Weir, D. (2012). Unexpected retirement from full time work after age 62: consequences for life satisfaction in older Americans. *European Journal of Ageing*, vol. 9, p. 207–219.

Clarke, P., Marshall, V., House, J., Lantz, P. (2011). The social structuring of mental health over the adult life course: advancing theory in the sociology of aging. *Social Forces*, vol. 89(4), p. 1287-1313.

Coe, N.B., Lindeboom, M. (2008). *Does retirement kill you? Evidence from early retirement windows*. Retrieved on the 8th of February 2013 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295315

Coe, N.B., Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, vol. 30, p. 77-86.

Coelho, C., Newman, E. (2014). *A systematic review of constructs and measures of adjustment and adaptation to retirement*. Unpublished Doctoral Thesis, University of Edinburgh.

Coelho, C., Newman, E., Huxtable, D. (2014). *A meta-analysis of the relationship between retirement and depression in later life.* Unpublished Doctoral Thesis, University of Edinburgh.

Conwell Y. (2009). Suicide prevention in later life: a glass half full, or half empty? *American Journal of Psychiatry*, vol 166, p. 845–848.

Conwell, Y., Duberstein, P. R., Caine, E. D. (2002). Risk factors for suicide in later life. *Biological Psychiatry*, vol. 52(3), p. 193-204.

Cooke, M. (2006). Policy changes and the labour force participation of older workers: Evidence from six countries. *Canadian Journal on Aging/La Revue Canadienne du Vieillissement*, vol. 25(4), p. 387-400.

Crawford, M. (1972). Retirement as a psycho-social crisis. *Journal of Psychosomatic Research*, vol. 16, p. 375-380.

Cyr, J.J., Mckenna-Foley, J.M., Peacock, E. (1985). Factor structure of the SCL-90-R: Is there one? *Journal of Personality Assessment*, vol. 49(6), p. 571-578.

Davies, A.R., Sherbourne, C.D., Peterson, J.R., Ware, J.E. (1998). *Scoring manual: adult health status and patient satisfaction measures used in RAND's Health Insurance Experiment*. Santa Monica: RAND Corporation.

De Luca, G., Rossetti, C. (2008). Sampling design and weighting strategies in the second wave of SHARE. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 333-338. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

De Vaus, D., Wells, Y., Kendig, H., Quine, S. (2007). Does gradual retirement have better outcomes than abrupt retirement? Results from an Australian panel study. *Ageing and Society*, vol. 27(5), p. 667-682.

Deeks, J.J., Dinnes, J., D'Amico, R., Sowden, A.J., Sakarovitch, C., Song, F., Petticrew, M., Altman, D.G. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, vol. 7, p. 27-36.

Denton, F.T., Spencer, B.G. (2009). What is retirement?: A review and assessment of alternative concepts and measures. *Canadian Journal on Aging /La Revue Canadienne du Vieillissement*, vol. 28(1), p. 63-76.

Derogatis, L. R., Cleary, P. A. (1977). Confirmation of the dimensional structure of the SCL-90: a study in construct validation. *Journal of Clinical Psychology*, vol. 33, p. 981-989.

Derogatis, L.R. (1983). *SCL-90-R: Administration, scoring and procedures manual-II for the R(evised) version and other instruments of the psychopathology rating scale series (2nd Edition)*. Towson: Clinical Psychometric Research.

Dick, H.R., Friedsam, H.J. (1964). Adjustment of residents of two homes for the aged. *Social Problems*, vol. 11(3), p. 282-290.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, vol. 95, p. 542-575.

Diener, E., Emmons, R., Larsen, R.J., Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, vol. 49(1), p. 71-75.

Diener, E., Eunkook, M.S., Lucas, R.E., Smith, H., L. (2009). Subjective well-being: three decades of progress. *Psychological Bulletin*, vol. 125(2), p. 276-302.

Donaldson, T., Earl, J.K, Muratore, A.M. (2010). Extending the integrated model of retirement adjustment: incorporating mastery and retirement planning. *Journal of Vocational Behavior*, vol. 77, p. 279–289.

Dorfman, L. (1989). Retirement preparation and retirement satisfaction in the rural elderly. *Journal of Applied Gerontology*, vol. 8, p. 432-450.

Drentea, P. (2002). Retirement and mental health. *Journal of Aging and Health*, vol. 14, p. 167-194.

Dulin, P.L, Stephens, C., Alpass, F., Hill, R.D., Stevenson, B. (2011). The impact of sociocontextual physical and lifestyle variables on measures of physical and psychological wellbeing among Māori and nonMāori: the New Zealand Health, Work and Retirement Study. *Ageing and Society*, vol. 31, p. 1406-1424.

Eichhorst, W., Gerard, M., Kendzia, M.J., Mayrhuber, C., Nielsen, C., Rünstler, G., Url, T. (2011). *Pension systems in the EU – contingent liabilities and assets in the public and private sector*. Retrieved on the 2nd of March 2013 from http://www.europarl.europa.eu/activities/committees/studies.do?language=EN.

Ekerdt, D. J., Bossé, R., & Levkoff, S. (1985). An empirical test of phases of retirement: findings from the normative aging study. *Journal of Gerontology*, vol. 40(1), p. 95-101.

Ekerdt, D.J. (2010). Frontiers of research on work and retirement. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 65B(1), p. 69-80.

Ekerdt, D.J., DeViney, S. (1990). On defining persons as retired. *Journal of Aging Studies*, vol. 4(3), p. 211-229.

Elder, G.H. (1994). Time, human agency, and social change: perspectives on the life course. *Social Psychology Quarterly*, vol. 57(1), p, 4-15.

Elder, G.H., Johnson, M.K., Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer, J.T and Shanahan, M.J. (Eds.), *Handbook of the life course*, p. 3-19. New York: Kluwer.

Elwood, M. (2007). *Critical Appraisal of Epidemiological Studies and Clinical Trials, 3rd Edition*. Oxford: Oxford University Press.

English Longitudinal Study of Ageing (ELSA), http://www.ifs.org.uk/ELSA

Farmer, R., Sundberg, N.D. (1986). Boredom proneness – the development and correlates of a new scale. *Journal of Personality Assessment*, vol. 50(1), p. 4-17.

Fe, E., Hollingsworth, B. (2012). *Estimating the effect of retirement on mental health via panel discontinuity designs*. Retrieved on the 7th of April 2013 from http://mpra.ub.uni-muenchen.de/38162/

Fiske, A., Wetherell, J.L., Gatz, M. (2009). Depression in older adults. *Annual Review of Clinical Psychology*, vol. 5, p. 363–389.

Floyd, F.J., Haynes, S.N., Doll, E.R., Winemiler, D., Lemsky, C., Burgy, T.M., Werle, M., Heilman, N. (1992). Assessing retirement satisfaction and perceptions of retirement experience. *Psychology and Aging*, vol. 7(4), p. 609-621.

Fonner V.A., Denison J., Kennedy C.E., O'Reilly K., Sweat M. (2012). Voluntary counselling and testing (VCT) for changing HIV related risk behaviour in developing countries. *Cochrane Database of Systematic Reviews*, Issue 9. Art. No.: CD001224. DOI:10.1002/14651858.CD001224.pub4.

Fonseca, A.M. (2007). Determinants of successful retirement in a Portuguese population. *Reviews in Clinical Gerontology*, vol. 17(3), p. 219-224.

Fouquereau, E., Fernandez, A., Fonseca, A.M., Paul, M.C., Uotinen, V. (2005). Perceptions of and satisfaction with retirement: a comparison of six European Union countries. *Psychology and Aging*, vol. 20(3), p. 524-528.

Gall, T.L., Evans, D.R. (2000). Preretirement expectations and the quality of life of male retirees in later retirement. *Canadian Journal of Behavioural Science*, vo. 32(3), p. 187-197.

Gall, T.L., Evans, D.R., Howard, J. (1997). The retirement adjustment process: changes in the wellbeing of male retirees across time. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 52B(3), p. 110-117.

Gallagher, D., Savva, G.M., Kenny, R.A., Lawlor, B.A. (2013). What predicts persistent depression in older adults across Europe? Utility of clinical and neuropsychological predictors from the SHARE study. *Journal of Affective Disorders*, vol. 147, p. 192-197.

Gallagher, M.W., Lopez, S.J., Preacher, K.J. (2009). The hierarchical structure of well-being. Journal of Personality, vol. 77(4), p. 1025-1050.

Gallo, W.T., Bradley, E.H., Siegel, M., Kasl, S.V. (2000). Health effects of involuntary job loss among older workers: findings from the health and retirement survey. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 55B(3), p. S131–S140.

Gana, K., Akremi, M. (1998). L'échelle de disposition à l'ennui (EDE): adaptation Française et validation du Boredom Proneness Scale (BP). *L'année psychologique*, vol. 98(3), p. 429-450.

Gautier, X. (2002). Emploi, retraites et cycles de vie. *Retraite et Société*, vol. 37, p. 163-207.

George, L. K. (1993). Sociological perspectives on life transitions. *Annual Review of Sociology*, vol. 19, p. 353-73.

George, L.K., Fillenbaum, G.G., Palmore, E. (1984). Sex differences in the antecedents and consequences of retirement. *Journal of Gerontology*, vol. 39(3), p. 364-371.

George, L.K., Maddox, G.L. (1977). Subjective adaptation to loss of the work role: a longitudinal study. *Journal of Gerontology*, vol. 32(4), p. 456-462.

Gilhooly, M.L.M. (1984). The impact of care-giving on care-givers: factors associated with the psychological well-being of people supporting a dementing relative in the community. *British Journal of Medical Psychology*, vol. 57, 35-44.

Gill, S.C., Butterworth, P., Rodgers, B., Anstey, K.J. , Villamil, E., Melzer, D. (2006). Mental health and the timing of men's retirement. *Social Psychiatry and Psychiatric Epidemiology*, vol. 41, p. 515-522.

Goldberg, D.P., 1972. *The Detection of Psychiatric Illness by Questionnaire*. Oxford: Oxford University Press.

Goldberg, D.P., Gater, R., Sartorius, N., Ustun, T.B., Piccinelli, M., Gureje, O., Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, vol. 27, p. 191-197.

Goldberg, D.P., Oldehinkel, T., Ormel, J. (1998). Why GHQ threshold varies from one place to another. *Psychological Medicine*, vol. 28, p. 915-921.

Golder, S., Loke, Y.K. (2009). Search strategies to identify information on adverse effects: a systematic review. *Journal of the Medical Library Association*, vol. 97(2), p. 84-92.

Guillemard, A.M., Rein, M. (1993). Comparative patterns of retirement: recent trends in developed societies. *Annual Review of Sociology*, vol. 19, p. 469-503.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social*

Han, S.K., Moen, P. (1999). Clocking out: temporal patterning of retirement. *American Journal of Sociology*, vol. 105, p. 191-236.

Han, S.K., Moen, P. (1999). Clocking out: temporal patterning of retirement. *American Journal of Sociology*, vol. 105, p. 191-236.

Hanisch, K.A., Hulin, C.L. (1990). Job attitudes and organizational withdrawal: an examination of retirement and other voluntary withdrawal behaviors. *Journal of Vocational Behavior*, vol. 37, p. 60-78.

Hardy, M. A. (2002). The transformation of retirement in twentieth-century America: from discontent to satisfaction. *Generations*, vol. 26(2), p. 9-16.

Hartling, L., Bond, K., Santaguida, P.L., Viswanathan, M., Dryden, D.M. (2011). Testing a tool for the classification of study designs in systematic reviews of interventions and exposures showed moderate reliability and low accuracy. *Journal of Clinical Epidemiology*, vo. 64, p. 861-871.

Hauser, R.M., Weir, D. (2010). Recent developments in longitudinal studies of aging in the United States. *Demography*, vol. 47 (Supplement), p. S111-S130.

Havinghurst, R.J. (1961). Successful aging. *The Gerontologist*, vol. 1, p. 8-13.

Helmes, E., Goffi, R.D., Chrisjohn, R.D. (2010). Confirmatory analysis of the Bradburn Affect Balance Scale and its relationship with morale in older Canadian adults. *Canadian Journal on Aging / La Revue canadienne du vieillissement*, vol 29 (2), p. 259-266.

Hershey, D., Henkens, K., Van Dalen, H. (2007). Mapping the minds of retirement planners. A cross-cultural perspective. *Journal of Cross-Cultural Psychology*, vol. 38, 361-382.

Hervé, C., Bailly, N., Joulain, M., Alaphilippe, D. (2012). Comparative study of the quality of adaptation and satisfaction with life of retirees according to retiring age. *Psychology*, vol. 3(4), p. 322-327.

Higgins, J. P. T., Deeks, J.J. (2011). Chapter 7: Selecting studies and collecting data. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Altman, D.G., Sterne, J.A.C. (2011). Chapter 8: Assessing risk of bias in included studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*. The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Green, S. (Editors) (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Higgins, J.P.T., Ramsay, C., Reeves, B.C., Deeks, J.J., Shea, B., Valentine, J.C., Tugwell, P., Wells, G. (2013) Issue relating to study design and risk of bias when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 12-25.

Higgins, J.P.T., Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, vol. 21, p. 1539-1558.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, vol. 327, p. 557-560.

Higo, M., Williamson, J.B. (2008). Retirement. In, D. Carr, R. Crosnoe,, M. E. Hughes, A. M. Pienta (Editors), *Encyclopedia of the life course and human development vol. 3 (Social Sciences)*, p. 328-336. Detroit: MacMillan Reference.

Hofer, S.M., Piccinin, A.M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods*, vol. 14(2), p. 150-164.

Holi, M. (2003). *Assessment of psychiatric symptoms using the SCL-90*. Unpublished Doctoral Thesis, Helsinki University.

Hooker, K., Ventis, D.G. (1984). Work ethic, daily activities, and retirement satisfaction. *Journal of Gerontology*, vol. 39(4), p. 478-484.

Horner, E.M. (2012). Subjective well-being and retirement: analysis and policy recommendations. *Journal of Happiness Studies.* Published Online on 1st December 2012, DOI 10.1007/s10902-012-9399-2

Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A., Moschetti, I., Phillips, B., Thornton, H. (2011). *The 2011 Oxford CEBM Evidence Levels of Evidence:*

Introductory Document. Oxford Centre for Evidence-Based Medicine. Retrieved on 29th of April 2013 from http://www.cebm.net/index.aspx?o=5653

Hyde, M., Ferrie, J., Higgs, P., Mein, G., Nazroo, J. (2004). The effects of pre-retirement factors and pre-retirement route on circumstances in retirement: findings from the Whitehall II Study. *Aging and Society*, vol. 24, p. 274-296.

Ioannidis, J.P. (1998). Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *Journal of the American Medical Association*, vol. 279, p. 281-286.

Isaksson, K., Johansson, G. (2000). Adaptation to continued work and early retirement following downsizing: long-term effects and gender differences. *Journal of Occupational and Organizational Psychology*, vol. 73, p. 241-256.

Jajodia, A., Borders, A. (2011). Memory predicts changes in depressive symptoms in older adults: a bidirectional longitudinal analysis. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 66(5), p. 571-581.

Jang, S-N., Cho, S-I., Chang, J., Boo, K., Shin, H-G., Lee, H., Berkman, L.F. (2009). Employment status and depressive symptoms in Koreans: results from a baseline survey of the Korean longitudinal study of aging. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 64(5), p. 677-683.

Jeffris, B.J., Nazareth, I., Marston, L., Moreno-Kustner, B., Bellón, J.A., Svab, I., Rotar, D., Geerlings, M.I., Xavier, M., Goçalves-Pereira, M., Vicente, B., Saldivia, S., Aluoja, A., Kalda, R., King, M. (2011). Associations between unemployment and major depressive disorder: evidence from an international, prospective study (the PREDICT cohort). *Social Science and Medicine*, vol. 73, p. 1627-1634.

Jokela, M., Ferrie, J.E., Gimeno, D., Chandola, T., Shipley, M.J., Head, J., Vahtera, J., Westerlund, H., Marmot, M.G., Kivimaki, M. (2010). From midlife to early old age: health trajectories associated with retirement. *Epidemiology*, vol. 21(3), p. 284-290.

Joyce, K., Pabayo, R., Critchley, J.A., Bambra C. (2010). Flexible working conditions and their effects on employee health and wellbeing. *Cochrane Database of Systematic Reviews*, Issue 2. Art. CD008009. DOI: 10.1002/14651858.CD008009.pub2.

Katz, I. R. (1996). On the inseparability of mental and physical health in aged persons: Lessons from depression and medical comorbidity. The American Journal of Geriatric Psychiatry, vol. 4, p. 1–16.

Keller, S.D, Ware, J.E., Bentler, P.M., Aaronson, N.K., Alonso, J., Apolone, G., Bjomer, J.B., Brazier, J., Bullinger, M., Kaasa, S., Leplège, A., Sullivan, M., Gandek, B. (1998). Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology*, vol. 51(11), p. 1179-1188.

Kerlinger, F.N., Lee, H.B. (2000). *Foundations of behavioural research (Fourth Edition).* Fort Worth: Harcourt College Publishers.

Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly*, 61, 121–140.

Kim, J. E., Moen, P. (2001). Is retirement good or bad for subjective well-being? *Current Directions in Psychological Science*, vol. 10(3), p. 83-86.

Kim, J.E., Moen, P. (2002). Retirement transitions, gender and psychological wellbeing: a life course ecological model. *Journals of Gerontology Series B: Psychological Science and Social Sciences*, vol. 57(3), p. 212-222.

Kim-Prieto, C., Diener, E., Tamir, M., Scollon, C., Diener, M. (2006). Integrating the diverse definitions of happiness: a time-sequential framework of subjective well-being. *Journal of Happiness Studies*, vol. 6, p. 261-300.

King, D.L., Haagsma, M.C., Delfabbro, P.H., Gradisar, M., Griffiths, M.D. (2013). Toward a consensus definition of pathological video-gaming: a systematic review of psychometric assessment tools. *Clinical Psychology Review*, vol. 33, p. 331-342.

Kleine-Budde, K., Muller, R., Kawohl, W., Bramesfeld, A., Moock, J., Rossler, W. (2013). The cost of depression: a cost analysis from a large database. *Journal of Affective Disorders*, vol. 147, p. 137–143.

Kohli, M., Künemund, H. (2002). La fin de carrière et la transition vers la retraite: Les limites d'âge chronologiques sont-elles un anachronism? *Retraite et société*, vol. 36, p. 81-107.

Kohout, F.J., Berkman, L.F., Evans, D.A., Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D Depression Symptoms Index. *Journal of Aging and Health*, vol. 5, p. 179-193.

Kolodziej, I. (2011). *The relationship between retirement and mental health: investigating the causal relationship in eleven European countries using SHARE*. Unpublished MSc Thesis, Erasmus University Rotterdam. Retrieved on the 8th of February 2013 from http://arno.uvt.nl/show.cgi?fid=122185

Kubicek, B., Korunka, C., Raymo, J.M., Hoonakker, P. (2011). Psychological wellbeing in retirement: the effects of personal and gendered contextual resources. *Journal of Occupational Health Psychology*, vol. 16(2), p. 230-246.

Kuerbis, A., Sacco, P. (2012). The impact of retirement on the drinking patterns of older adults: a review. *Addictive Behaviors*, vol. 37, p. 587-595.

Laaksonen, M., Metsä-Simola, N., Martikainen, P., Pietiläinen, O., Rahkonen, O., Gould, R., Partonen, T., Lahelma, E. (2012). Trajectories of mental health before and after old-age and disability retirement: a register-based study on purchases of psychotropic drugs. *Scandinavian Journal of Work and Environmental Health*, vol. 38(5), p. 409-417.

Ladin, K. (2012). Decomposing differences in utilization of health services between depressed and non-depressed elders in Europe. *European Journal of Ageing*, vol. 9(1), p. 51-64.

Larson, R. (1978). Thirty years of research on the subjective well-being of older Americans. *Journal of Gerontology*, vol. 33(1), p. 109-125.

Lawton, M.P. (1975). The Philadelphia Geriatric Centre Morale Scale: a revision. *Journal of Gerontology*, vol. 30(1), p. 85-89.

Lawton, M.P. (2003). *Lawton's PGC Morale Scale* [Morale Scale created by M. Powell Lawton (1923-2001) while at the Polisher Research Institute of the Philadelphia Geriatric Center (now known as the Abramson Center for Jewish Life)]. Retrieved on the 6th of November 2013, from http://www.abramsoncenter.org/PRI/ (Scales page).

Leber, U., Wagner, A. (2007). *Early and phased retirement in European companies: establishment survey on working time 2004–2005*. Retrieved on the 2nd of March 2013 from http://www.eurofound.europa.eu/publications/htmlfiles/ef0688.htm

Lebowitz, B.D., Pearson, J.L., Schneider, L.S., Reynolds, C.F. III, Alexopoulos, G.S., Bruce ,M.L., Conwell, Y., et al. (1997). Diagnosis and treatment of depression in late life: consensus statement update. *Journal of the American Medical Association*, vol. 278, p. 1186–1190.

Lefebvre, C., Manheimer, E., Glanville, J. (2011). Chapter 6: Searching for studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Lindeboom, M., Portrait, F., van den Berg, G.J. (2002). An econometric analysis of the mental-health effects of major events in the life of older individuals. *Health Economics*, vol. 11, p. 505-520.

Lipsey, M.W., Wilson, D.B. (2001). *Practical Meta-Analysis*. Thousand Oaks: Sage Publications, Inc.

Löckenhoff, C.E. (2012). Understanding retirement: the promise of life-span developmental frameworks. *European Journal of Ageing*, vol. 9, p. 227-231.

Lohmann, N. (1977). Correlations of life satisfaction, morale and adjustment measures. *Journal of Gerontology*, vol. 32(1), p. 73-75.

Longitudinal Aging Study Amsterdam (LASA), http://www.lasa-vu.nl/index.htm

Loretto, W. (2010). Work and retirement in an ageing world: the case of older workers in the UK. *21st Century Society*, vol. 5(3), p. 279-294.

Luborski, M. R., LeBlanc, I.M. (2003). Cross-cultural perspectives on the concept of retirement: an analytic redefinition. *Journal of Cross-Cultural Gerontology*, vol. 18, p. 251-271.

Lucas, R.E. (2007). Adaptation and the set-point model of subjective well-being: does happiness change after major life events? *Current Directions in Psychological Science*, vol. 16(2), p. 75-79.

Lucas, R.E., Clark, A.E., Georgellis, Y., Diener, E. (2003). Re-examining adaptation and the set-point model of happiness: reactions to changes in marital status, *Journal of Personality and Social Psychology*, vol. 84(3), p. 527-539.

Luhmann, M., Hoffman, W., Eid, M., Lucas, R.E. (2012). Subjective wellbeing and adaptation to life-events: a meta-analysis. *Journal of Personality and Social Psychology*, vol. 102(3), p. 592-615.

Luppa, M., Sikorski, C., Luck, T., Weyerer, S., Villringer, A., König, H-H., Riedel-Heller, S. G. (2011). Epidemiology of depression in old age: results of the Leipzig Longitudinal Study of the Aged (LEILA 75+). *Journal of Epidemiology and Community Health*, vol. 65, Suppl 1, p. 127.

Luppa, M., Sikorski, C., Luck, T., Weyerer, S., Villringer, A., König, H-H., Riedel-Heller, S. G. (2012). Prevalence and risk factors of depressive symptoms in latest life – results of the Leipzig Longitudinal Study of the Aged (LEILA 75+). *International Journal of Geriatric Psychiatry*, vol. 27, p. 286-295.

Maimaris, W., Hogan, H., Lock, K. (2010). The impact of working beyond traditional retirement ages on mental health: implications for public health and welfare policy. *Public Health Reviews*, vol. 32, p. 532-548.

Marshall, V.R., Taylor, P. (2005). Restructuring the life course: work and retirement. In, Johnson, M.L., Bengston, V.L., Coleman, P.G. and Kirkwood, T.B.L. (Editors), *The Cambridge Handbook of Age and Ageing*, p. 572-582. Cambridge: Cambridge University Press.

Mattila, V.J., Joukamaa, M.I., Salokangas, R.K.R. (1989). Retirement, aging, psychosocial adaptation and mental health: findings of the TURVA project. *Acta Psychiatrica Scandinavica*, vol. 80. p. 356-367.

McDowell, I. (2010). Measures of self-perceived well-being. *Journal of Psychosomatic Research*, vol. 69, p. 69-79.

McDowell, I., Praught, E. (1982). On the measurement of happiness: an examination of the Bradburn Scale in the Canada Health Survey. *American Journal of Epidemiology*, vol. 116(6), p. 949-958.

McKee-Ryan, F.M., Song, Z., Wanberg, R., Kinicki, A.J. (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of Applied Psychology*, vol. 90(1), p. 53-76.

McVittie, C., Goodall, K. (2012). The ever-changing meanings of retirement. *American Psychologist*, vol. 67(1), p. 75-77.

McVittie, C., McKinlay, A., Widdicombe, S. (2008). Passive and active non-employment: age, employment and the identities of older non-working people. *Journal of Aging Studies*, 22(3), 248-255.

Mein, G., Martikainen, P., Hemingway, H., Stansfels, S., Marmot, M. (2003). Is retirement good or bad for mental and physical health functioning? Whitehall II longitudinal study of civil servants. *Journal of Epidemiology and Community Health*, vol. 57, p. 46-49.

Midanik, L.T., Soguikian, K., Ranson, L.J., Tekawa, I.S. (1995). The effect of retirement on mental health and health behaviours: the Kaiser Permanente Retirement Study. *Journals of Gerontology Series B, Psychological and Social Sciences*, vol. 50B(1), p. S59-S61.

Mitchell, H., Guled, G. (2010). *Average age of withdrawal from the labour market: A methodology update*. Retrieved on the 3rd of January 2013 from www.ons.gov.uk/ons/rel/pensions/average-age-of-withdrawal-from-the-labour-market/2010/index.html

Moen, P. (1996). A life course perspective on retirement, gender and wellbeing. *Journal of Occupational Health Psychology*, vol. 1(2), p. 131-144.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6), p. e1000097.

Moher, D., Pham, B., Lawson, M.L., Klassen, T.P. (2003). The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment*, vol. 7(41).

Mojon-Azzi, S., Sousa-Poza, A., Widmer, R.W. (2007). The effect of retirement on health: a panel analysis using data from the Swiss Household Panel. *Swiss Medical Weekly*, vol. 137, p. 581-585.

Moriwaki, S.Y. (1974). The Affect Balance Scale: a validity study with aged samples. *Journal of Gerontology*, vol. 29(1), p. 73-78.

Morris, J.N., Sherwood, S. (1975). A retesting and modification of the Philadelphia Geriatric Center Morale Scale. *Journal of Gerontology*, vol. 30(1), 77-84.

Morris, S.B., DeShon, R.P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, vol. 7(1), p. 105-125.

Neugarten, B.L., Havighurst, R.J., Tobin, S.S. (1961). The measurement of life satisfaction. *Journal of Gerontology*, vol. 16, p. 134-143.

Nordenmark, M., Stattin, M. (2009). Psychosocial wellbeing and reasons for retirement in Sweden. *Ageing & Society*, vol. 29, p. 413-430.

Norris, S.L., Moher, D., Reeves, B.C., Shea, B., Loke, Y., Garner, S., Anderson, L., Tugwell, P., Wells, G. (2013). Issues relating to selective reporting when including non-randomized studies in systematic reviews on the effects of healthcare interventions. *Research Synthesis Methods*, vol. 4, p. 36-47.

Nuttman-Shwartz, O. (2004). Like a high wave: adjustment to retirement. *The Gerontologist*, vol. 44(2), p. 229-236.

Oksannen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjosten, N., Virtanen, M., Kawachi, I., Kivimaki, M. (2011). Is retirement beneficial for mental health? Antidepressant use before and after retirement. *Epidemiology*, vol. 22(4), p.553-559.

Oldfield, Z. (2012). Economics domain tables. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), *The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5)*, p. 214-258. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

Organisation for Economic Co-operation and Development (OECD) (2011). Pensionable years, in *Society at a Glance 2011. OECD Social Indicators*. OECD Publishing. Retrieved on 31st of December 2012 from http://dx.doi.org/10.1787/soc_glance-2011-14-en

Palmore, E., Cleveland, W.P., Nowlin, J.B., Ramm, D., Siegler, I.C. (1979). Stress and adaptation in later life. *Journal of Gerontology*, vol. 34(6), p. 841-851.

Paul, K.I., Moser, K. (2009). Unemployment impairs mental health: meta-analyses. *Journal of Vocational Behaviour*, vol. 79, p. 264-282.

Petticrew, M., Roberts, H. (2006). *Systematic Reviews in the Social Sciences: a practical guide*. Oxford: Blackwell Publishing.

Peytremann-Bridevaux, I., Voellinger, R., Santos-Eggimann, B. (2008). Healthcare and preventive services utilization of elderly Europeans with depressive symptoms. *Journal of Affective Disorders*, vol. 105, p. 247-252.

Pinquart, M., Schindler, I. (2007). Changes of life satisfaction in the transition to retirement: a latent class approach. *Psychology and Aging*, vol. 22(3), p. 442-455.

Potocnik, K., Tordera, N., Peiró, J.M. (2010). The influence of the early retirement process on satisfaction with early retirement and psychological well-being. *International Journal of Aging and Human Development*, vol. 70(3), p. 251-273.

Potocnik, K., Tordera, N., Peiró, J.M. (2013).Truly satisfied with your retirement or just resigned? pathways toward different patterns of retirement satisfaction. *Journal of Applied Gerontology*, vol. 32, p. 164.

Price, C.A. (2000). Women and retirement: relinquishing professional identity. *Journal of Aging Studies*, vol. 14(1), p. 81-101.

Price, C.A. (2003). Professional women's retirement adjustment: the experience of reestablishing order. *Journal of Aging Studies*, vol. 17, p. 341-355.

Price, C.A., Nesteruck, O. (2010). Creating retirement paths: examples from the lives of women. *Journal of Women & Aging*, vol. 22, p. 136-149.

Quick, H. E., Moen, P. (1998). Gender, employment, and retirement quality: a life-course approach to the differential experiences of men and women. *Journal of Occupational Health Psychology*, vol. 3, p. 44-64.

Radloff, L.S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, vol. 1, p. 385-401.

Reeves, B.C., Deeks J.J., Higgins, J.P.T., Wells, G.A. (2011). Chapter 13: Including non-randomized studies. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Reeves, B.C., Higgins, J.P.T., Ramsay, C., Shea, B., Tugwell, P., Wells. G. (2013). An introduction to methodological issues when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 1-11.

Reitzes, D.C., Mutran, E.J. (2004). The transition to retirement: stages and factors that influence retirement adjustment. International Journal of Aging and Human Development, vol. 59(1), p. 63-84.

Reitzes, D.C., Mutran, E.J. (2006). Lingering identities in retirement. *The Sociological Quarterly*, vol. 47, p. 333-359.

Reitzes, D.C., Mutran, E.J., Fernandez, M.E. (1996). Does retirement hurt wellbeing? Factors influencing self-esteem and depression among retirees and workers. *The Gerontologist*, vol. 36(5), p. 649-656.

Robbins, S.B., Payne, E.C., Chartrand, J.M. (1990). Goal instability and later life adjustment. *Psychology and Aging*, vol. 5(3), p. 447-450.

Rosenthal, R. (1979). The file-drawer problem and tolerance for null results. *Psychological Bulletin*, vol. 86, p. 638-641.

Ross, C. E., Drentea, P. (1998). Consequences of retirement activities for distress and sense of personal control. *Journal of Health and Social Behavior*, vol. 39(4). p. 317-334.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of eudaimonic well-being. *Journal of Personality and Social Psychology*, vol. 57, p. 1069-1081.

Ryff, C.D., Keyes, C.L.M. (1995). The Structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, vol. 69(4), p. 719-727.

Sahlgren, G.H. (2013). *Work longer, live healthier: the relationship between economic activity, health and government policy*. Institute of Economic Affairs (London) Discussion Paper No. 46. Retrieved on the 20th of May 2013 from http://www.iea.org.uk/sites/default/files/publications/files/Work%20Longer,%20Live_Healthier.pdf

Sanderson, S., Tatt, I.D., Higgins, J.P.T. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *International Journal of Epidemiology*, vol. 36, p. 666-676.

Sargent, L.D., Bataille, C.D., Vough, H.C., Lee, M.D. (2011). Metaphors for retirement: unshackled from schedules. *Journal of Vocational Behaviour*, vol. 79, p. 315-324.

Sargent-Cox, K.A., Anstey, K.J., Kendig, H., Skladzien, E. (2012). Determinants of retirement timing expectations in the United States and Australia: A cross-national comparison of the effects of health and retirement benefit policies on retirement timing decisions. *Journal of Aging & Social Policy*, vol. 24(3), p. 291-308.

Schalk, R., van Veldhoven, M., de Lange, A.H., de Witte, H., Kraus, K., Stamov-Roßnagel, C., Tordera, N., van der Heijden, B., Zappalà, S., Bal, M., Bertrand, F., Claes, R., Crego, A., Dorenbosch, L., de Jong, J., Desmette, D., Gellert, F.J., Hansez, I., Iller, C., Kooij, D., Kuipers, B., Linkola, P., van den Broeck, A., van der Schoot, E., Zacher, H. (2010). Moving European research on work and ageing forward: overview and agenda. *European Journal of Work and Organizational Psychology*, vol. 19(1), p. 76-101.

Schiaffino, K.M. (2003). Other measures of psychological well-being: The Affect Balance Scale (ABS), General Health Questionnaire (GHQ-12), Life Satisfaction Index-A (LSI-A), Rosenberg Self-Esteem Scale, Satisfaction with Life Scale (SWLS), and State-Trait Anxiety Index (STAI). *Arthritis & Rheumatism (Arthritis Care & Research)*, vol. 49(5S), p S165-S174.

Schrank, B., Bird, V., Tylee, A., Coggins, T., Rashid, T., Slade, M. (2013). Conceptualising and measuring the well-being of people with psychosis: systematic review and narrative synthesis. *Social Science & Medicine*, vol. 92, p. 9-21.

Schwarzbach, M., Luppa, M., Sikorski, C., Fuchs, A., Maier, W., van den Bussche, H., Pentzek, M., Riedel-Heller, S.G. (2013). The relationship between social integration and depression in non-demented primary care patients aged 75 years and older. *Journal of Affective Disorders*, vol. 145, p. 172–178.

Shamliyan, T.A., Kane, R.L., Ansari, M.T. Raman, G. Berkman, N.D., Grant, M., Janes, G., Maglione, M., Moher, D., Nasser, M., Robinson, K.A., Segal, J.B., Tsouros, S. (2011). *Development of Quality Criteria to Evaluate Nontherapeutic Studies of Incidence, Prevalence, or Risk Factors of Chronic Diseases: Pilot Study of New Checklists*. Retrieved on 7th of October 2013 from http://effectivehealthcare.ahrq.gov/ehc/products/281/617/Methods_Research_Report--Shamliyan_final_1-28-11.pdf

Shapiro, A., Yarborough-Hayes, R. (2008). Retirement and older men's health. *Generations*, vol 32(1), p. 49-53.

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: why validity issues in meta-analyses will not go away. *Clinical Psychology Review*, vol. 17(8), p. 881-901.

Shills, T. (2008). Early retirement in Germany, the Netherlands, and the United Kingdom: a longitudinal analysis of individual factors and institutional regimes. *European Sociological Review*, vol. 24(3), p. 315-329.

Shultz, K.S., Wang, M. (2011). Psychological perspectives on the changing nature of retirement. *American Psychologist*, vol. 66(3), p. 170-179.

Shünemann, H.J., Tugwel, P., Reeves, B.C., Akl, E.A., Santesso, N., Spencer, F.A., Shea, B., Wells, G., Helfand, M. (2013). Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews of on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 49-62.

Siegfried, N., Muller, M., Deeks, J., Volmink, J., Egger, M., Low, N., Walker, S., Williamson, P. (2005). HIV and male circumcision – a systematic review with assessment of the quality of studies. *Lancet Infectious Diseases*, vol. 5, p. 165-173.

Siegfried, N., Muller, M., Deeks, J.J., Volmink, J. (2009). Male circumcision for prevention of heterosexual acquisition of HIV in men. *Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD003362. DOI: 10.1002/14651858.CD003362.pub2.

Siegfried, N., Muller, M., Volmink, J., Deeks, J.J., Egger, M., Low, N., Weiss, H.H., Walker, S.A., Williamson, P.R. (2003). Male circumcision for prevention of heterosexual acquisition of HIV in men. *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD003362. DOI: 10.1002/14651858.CD003362.

Silver, M.P. (2010): Women's retirement and self-assessed well-being: an analysis of three measures of well-being among recent and long-term retirees relative to homemakers. *Women & Health*, vol. 50, p. 1-19.

Sklar, M., Groessl, E.J., O'Connell, M., Davidson, L., Aarons, G.A. (2013). Instruments for measuring mental health recovery: a systematic review. *Clinical Psychology Review*, vol. 33, p. 1082-1095.

Smit, F., Ederveen, A., Cuijpers, P., Deeg, D., Beekman, A. (2006). Opportunities for cost-effective prevention of late-life depression: an epidemiological approach. *Archives of General Psychiatry*, vol. 63, p. 290-296.

Smith, J.M., Alloy, L.B. (2009). A roadmap to rumination: a review of the definition, assessment, and conceptualization of this multifaceted construct. *Clinical Psychology Review*, vol. 29, p. 116-128.

Stang, A. (2008). Appropriate epidemiologic methods as a prerequisite for valid study results. *European Journal of Epidemiology*, vol 23, p. 761-765.

Steffens, D.C., Fisher, G.G., Langa, K.M., Potter, G.G., Plassman, B.L. (2009). Prevalence of depression among older Americans: the Aging, Demographics and Memory Study. *International Psychogeriatrics*, vol. 21(5), p. 879-888.

Steffick, D.E. (2000). *Documentation of affective functioning measures in the Health and Retirement Study*. HRS Health Working Group, Survey Research Center, University of Michigan. Retrieved on the 23rd of December 2013 from http://hrsonline.isr.umich.edu/sitedocs/userg/dr-005.pdf

Steptoe, A., Demakakos, P., Oliveira, C. (2012). Chapter 4: The psychological well-being, health and functioning of older people in England. In, Banks, J., Nazroo, J., Steptoe, A. (Eds.), *The Dynamics of Ageing: Evidence from The English Longitudinal Study Of Ageing 2002–10 (Wave 5)*, p. 98-182. London: The Institute for Fiscal Studies. Retrieved on the 1st of April 2013 from http://www.ifs.org.uk/ELSA/reportWave5

Sterne, J.A.C., Egger, M., Moher, D. (2011). Chapter 10: Addressing reporting biases. In, Higgins, J.P.T., Green, S. (Editors), *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).* The Cochrane Collaboration, 2011. Retrieved on 16th of July 2012 from www.cochrane-handbook.org.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A., Thacker, S.B. (2000). Meta-Analysis of Observational Studies in Epidemiology: A Proposal for Reporting. *Journal of the American Medical Association*, vol. 283(15), p. 2008-2015.

Szinovacz, M. E. 1980. Female retirement: Effects on spousal roles and marital adjustment. *Journal of Family Issues*, vol. 1, p. 423-440.

Szinovacz, M.E., Davey, A. (2004). Honeymoons and joint lunches: effects of retirement and spouse's employment on depressive symptoms. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* vol. 59B(5), p. 233-245.

Szinovacz, M.E., DeViney, S. (1999). The retiree identity: gender and race differences. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 54B(4), p. S2O7-S218.

Taylor , Shore (1995). Predictors of planned retirement age: an application of Beehr's model. *Psychology and Aging*, vol. 10(1), p. 76-83.

Terrin, N., Schmid, C. H. Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, vol. 58(9), p. 894-901.

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

The United Kingdom Government (2011). Department of Business Innovation and Skills: Phasing out the default retirement age: government response to consultation. Retrieved on the 8th of April 2013 fromhttp://www.bis.gov.uk

The United Kingdom Government (2013). *HM Treasury: Autumn Statement*. Retrieved on the 5th of December from https://www.gov.uk/government/news/autumn-statement-2013-date-announced-by-chancellor-as2013

Thelin, A., Holmberg, S. (2010). Farmers and retirement: A longitudinal cohort study. *Journal of Agromedicine*, vol. 15(1), p. 38-46.

Tilghman-Osborne, C., Cole, D.A., Felton, J.W. (2010). Definition and measurement of guilt: implications for clinical research and practice. *Clinical Psychology Review*, vol. 30, p. 536-546.

Tooth, L., Ware, R., Bain, C., Purdie, D.M., Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology*, vol. 161, p. 280-288.

Topa, G., Moriano, J.A., Depolo, M., Alcover, C.M., Morales, J.F. (2009). Antecedents and consequences of retirement planning and consequences: a meta-analysis and model. *Journal of Vocational Behaviour*, vol. 75, p. 38-55.

Tugwell, P., Petticrew, M., Kristjansson, E., Welch, V., Ueffing, E., Waters, E., Bonnefoy, J., Morgan, A., Doohan, E., Kelly, M.P. (2010). Assessing equity in systematic reviews: realising the recommendations of the Commission on Social Determinants of Health. *British Medical Journal*, vol. 341, p. c4739.

Turvey, C., Wallace, R.B., Herzog, R. (1999). Depression: a revised CES-D measure of depressive symptoms and a DSM-based measure of major depressive episodes in the elderly. *International Psychogeriatrics*, vol. 11(2), p. 139-148.

Vahtera, J., Westerlund, H., Hall, M., Sjosten, N., Kivimaki, M., Salo, P., Ferie, J.E., Jokela, M., Pentti, J., Singh-Manoux, A., Goldberg, M., Zins, M. (2009). Effect of retirement on sleep disturbances: the GAZEL prospective cohort study. *Sleep*, vol. 32(11), p. 1459-1466.

Valentine, J.C., Thomson, P. (2013). Issue relating to confounding and meta-analysis when including non-randomized studies in systematic reviews on the effects of interventions. *Research Synthesis Methods*, vol. 4, p. 26-35.

Van de Velde, S., Levecque, K., Bracke, P. (2009). Measurement equivalence of the CES-D 8 in the general population in Belgium: a gender perspective. *Archives of Public Health*, vol. 67, p. 15-29.

van Solinge, H., Henkens, K. (2005). Couples' adjustment to retirement: a multi-actor panel study. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 60B(1), S11–S20.

van Solinge, H., Henkens, K. (2007). Involuntary retirement: the role of restrictive circumstances, timing and social embeddedness. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 62B(5), p. S295-S303.

van Solinge, H., Henkens, K. (2008). Adjustment to and satisfaction with retirement: two of a kind? *Psychology and Aging*, vol. 23(2), p. 422-434.

van Solinge, H., Henkens, K. (2009). Living longer, working longer? The impact of subjective life expectancy on retirement intentions and behaviour. *European Journal of Public Health*, vol. 20(1), p. 47-51.

Vandewater, E.A., Ostrove, J.M., Stewart, A.J. (1997). Predicting women's well-being in midlife: the importance of personality development and social role involvements. *Journal of Personality and Social Psychology,* vol. 72, 1147–1160.

Veit, C.T., Ware, J.E. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology*, vol. 51(5), p. 730-742.

Vickerstaff. S., Cox, J. (2005). Retirement and risk: the individualization of retirement experiences? *Sociological Review*, vol. 53(1), p. 77-95.

Villamil, E., Huppert, F.A., Melzer, D. (2006). Low prevalence of depression and anxiety is linked to statutory retirement ages rather than personal work exit: a national survey. *Psychological Medicine*, vol. 36, p. 999-1009.

Villanueva, A.B. (2000). Review Essay: Population aging and retirement issues in the next millennium. *The Social Science Journal*, vol. 37(2), p. 321-325.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines of reporting observational studies. *Epidemiology*, vol. 18(6), p. 800-804.

Wang, M. (2007). Profiling retirees in the retirement transition and adjustment process: examining the longitudinal change patterns of retirees' psychological well-being. *Journal of Applied Psychology*, vol. 92(2), p. 455-474.

Wang, M., Bodner, T.E. (2007). Growth mixture modeling identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, vol. 10(4), p. 635-656.

Wang, M., Henkens, K., van Solinge, H. (2011). Retirement adjustment: a review of theoretical and empirical advancements. *American Psychologist*, vol. 66(3), p. 204-213.

Ware, J.E., Sherbourne, C.D. (1992). The MOS 36-Item Short-Form Health Survey (SF-36) I: Conceptual framework and item selection. *Medical Care*, vol. 30(6), p. 473-483.

Warner, D.F., Hayward. M.D., Hardy, M.A. (2010). The retirement life course in America at the dawn of the twenty-first century. *Population Research Policy Review*, vol. 29, p. 893–919.

Warr, P., Butcher, V., Robertson, I., Callinan, M. (2004). Older people's well-being as a function of employment, retirement, environmental characteristics and role preference. *British Journal of Psychology*, vol. 95, p. 297-324.

Watson, L.C., Lewis, C.L., Kistler, C.E., Amick, H.R., Boustani, M. (2004). Can we trust depression screening instruments in healthy 'old-old' adults? *International Journal of Geriatric Psychiatry*, vol. 19, p. 278-285.

Weller, C.E. (2010). Introduction: retirement security in the great recession. *Journal of Aging & Social Policy*, vol. 22, p. 95-98.

Wells, G.A., Shea, B., Higgins, J.P.T., Sterne, J., Tugwell, P., Reeves, B.C. (2013). Checklists of methodological issues for authors to consider when including non-randomized studies in systematic reviews. *Research Synthesis Methods*, vol.4, p. 63-77.

Werneke, U., Goldberg, D.P., Yalcin, I., Ustun, B.T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, vol. 30, p. 823-829.

Westerlund, H., Kivimaki, M., Singh-Manoux, A., Melchior, M., Ferrie, J.E., Pentti, J., Jokela, M., Leineweber, C., Goldberg, M., Zins, M., Vahtera, J. (2009). Self-rated health before and after retirement in France (GAZEL): a cohort study. *The Lancet*, vol. 374, p. 1889-1896.

Westerlund, H., Vahtera, J., Ferrie, J.E., Singh-Manoux, A., Pentti, J., Melchior, M., Leineweber, C., Jokela, M., Siegrist, J., Goldberg, M., Zins, M., Kivimaki, M. (2010). Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. *British Medical Journal*, BMJ 2010:341:c6149.

Williams, S.E., Watts, T.K.O., Wade, T.D. (2012). A review of the definitions of outcome used in the treatment of bulimia nervosa. *Clinical Psychology Review*, vol. 32, p. 292-300.

Wilson, K., Mottram, P.G., Vassilas, C. (2008). Psychotherapeutic treatments for older depressed people. *Cochrane Database of Systematic Reviews*, Art. No: CD004853. doi:10.1002/14651858.CD004853.pub2

Wilson, S.H., Walker, G.M. (1993). Unemployment and health: a review. *Public Health*, vol. 107, p. 153-162.

Wong, J.Y., Earl, J.K. (2009). Towards an integrated model of individual, psychosocial, and organizational predictors of retirement adjustment. *Journal of Vocational Behavior*, vol. 75, p. 1-13.

World Health Organisation (1948). Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946. *Official Records of the World Health Organization*, no. 2, p. 100. Retrieved on the 31st of October 2013 from http://www.who.int/about/definition/en/print.html

World Health Organisation (1990). *Composite International Diagnostic Interview (CIDI).* Geneva: World Health Organisation.

World Health Organization (2001). *World Health Report 2001. Mental Health: New Understanding, New Hope.* Retrieved on the 3rd of September 2013 from http://www.who.int/whr/2001/en/whr01_en.pdf

Young T, Hopewell S. (2011). Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews*, Issue 11, Art. No.: MR000027. DOI: 10.1002/14651858.MR000027.pub2.

Zamarro, G., Meijer, E., Fernandes, M. (2008). Labor force participation and retirement. In, Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (Eds.) *First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension*, p. 48-55. Mannheim: Mannheim Research Institute for the Economics of Aging (MEA). Retrieved on the 8th of April 2013 from http://www.share-project.org/uploads/tx_sharepublications/BuchSHAREganz250808.pdf

Zhang, J-P., Kahana, B., Kahana, E., Hu, B., Pozuelo, L. (2009). Joint modeling of longitudinal changes in depressive symptoms and mortality in a sample of community-dwelling elderly people. *Psychosomatic Medicine*, vol. 71, p. 704–714.

Zsembik, B.A., Singer, A. (1990). The Problem of defining retirement among minorities: The Mexican Americans. *The Gerontologist*, vol. 30(6), p. 749-757.