A Systematic Review of the Psychometric Properties of Self-Report Measures of Attitudes to Ageing

Farah Nadia M.Faudzi¹, Christopher J. Armitage¹,²,³ , Christina Bryant⁴, Laura J.E. Brown*¹

This is a pre-copyedited, author-produced PDF of an article accepted for publication in Research on Aging following peer review. The final published version of the article Faudzi et al (2019) is available at: https://doi.org/10.1177/0164027518825117

¹ Manchester Centre for Health Psychology, Division of Psychology and Mental Health, University of Manchester, UK.
² NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, UK
³NIHR Greater Manchester Patient Safety Translational Research Centre
⁴ Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, Australia

*Address correspondence to Laura J.E. Brown, Manchester Centre for Health Psychology, Division of Psychology and Mental Health, and The University of Manchester, Zochonis Building, Manchester, M13 9PL. E-mail: laura.brown@manchester.ac.uk
Abstract

The aim of this study was to describe and systematically evaluate the psychometric properties of attitudes to ageing measures that have been validated in adults younger than 60 years. Studies were assessed for methodological quality using the COnsensus-based Standards for the selection of health Measurement INstruments. Best evidence synthesis was performed to summarize levels of evidence. Systematic searching yielded 21 articles, relating to 10 different measures. Some instruments were validated only in middle-aged and older people, whilst others were validated in a wider age range. Measures differed in whether their items related to participants’ own experiences of ageing, their anticipated future ageing, and/or ageing in general. None of the measures had had all of its psychometric properties assessed. The ERA-12 and the AAS received positive ratings for the greatest number of psychometric properties, but capture different constructs, and may be differentially suited to different age groups of younger adults.

**Keywords:** Systematic Review, Measurement Properties, Psychometric Assessment, Attitudes to Ageing, younger adults.
Introduction

Attitudes to ageing can be defined as expectations, experiences or feelings about the process of ageing (Hess, 2006), or how an individual envisages old age (Kornadt & Rothermund, 2015). There is consistent evidence for long-term consequences of attitudes to ageing on individuals’ health and well-being outcomes (Wurm, Diehl, Kornadt, Westerhof & Wahl, 2017). For example, negative attitudes to ageing held in early or middle adulthood are predictive of future cognitive, physical or mental health, longevity and mortality (Levy, Slade, Kunkel & Kasl, 2002; Sargent-Cox, Anstey & Luszcz, 2013; Wurm, Tesch-Römer & Tomasik, 2007). Reliable and valid methods for assessing attitudes towards ageing in younger populations are therefore important for supporting healthy ageing.

Measures of attitudes to ageing differ in terms of the specific construct(s) that they assess. For instance, attitudes to ageing likely take on different meanings at different life stages (Barret & Montepare, 2015), as those held by younger adults are most likely based on observations of other people’s ageing (Gilbert & Ricketts, 2008) or cultural stereotypes of old age (Levy, 2009), whereas those of older adults (i.e., those over 60 years of age, see United Nations, 2013) are additionally influenced by their personal experiences of age-related changes (O’Hanlon & Colemen, 2008). This is reflected in the content of some measures of attitudes to ageing, whereby those designed for older adults may draw on reflections of the individual’s experience of ageing. For example, the Attitudes to Ageing Questionnaire (AAQ: Laidlaw, Power & Schmidt, 2007) was specifically developed for people aged over 60, and contains items that directly question respondents’ experiences of ageing, such as ‘I am losing my physical
independence as I get older’. Such measures are therefore not well-suited to those who have not yet experienced these specific features of ageing. Measures that do not refer to an individual’s experiences of aging can also differ in terms of whether they contain self-referential items that relate to their expectations about their own future ageing (e.g. ‘I expect that as I get older I will become more forgetful’; Sarkisian, Steers, Hays & Mangione, (2005)), or non-self-referential items that relate to their attitudes to ageing or old age as a general construct (e.g. ‘Old age is a time to enjoy life’; Sarkisian, Hays, Berry & Mangione, (2002)). Broadly speaking, there are therefore three different constructs that may be captured in measures of attitudes to ageing, namely attitudes relating to: 1) an individual’s experiences of their own ageing, 2) their expectations about their own future ageing, or 3) ageing as a general construct. Researchers trying to select or interpret the results of individual measures need to consider the specific attitudinal construct that is being measured in each case.

Given the known longitudinal relationships between attitudes to ageing and health and wellbeing outcomes, reliable and valid assessments of attitudes to ageing among younger people are crucial for targeting, developing and evaluating interventions designed to promote healthy ageing. However, no review of psychometric properties of the tools available to assess attitudes to ageing in younger populations has yet been conducted. To address this gap, the aims of the present systematic review were therefore to: (a) identify and describe the range of tools that have been developed to assess attitudes to ageing in adults aged younger than 60 years old, and (b) assess the psychometric properties of these tools as measures of attitudes to ageing.

Methods

Search Strategy
Six electronic databases (PsycINFO, MEDLINE, CINAHL, EMBASE, Web of Science and Applied Social Sciences Index and Abstracts (ASSIA)) were systematically searched from inception to 1 May 2018. The search strategy consisted of three concept strands: (1) Psychometric Property; (2) Measurement Tool; and (3) Attitude to Ageing, which were combined with the AND operator. Each concept strand contained various synonyms, spelling variants, and truncations of the construct of interest, and relevant MeSH terms, where available, which were combined with OR operators. The search strategy was initially developed for use in PsycINFO, EMBASE and MEDLINE, and then adapted when necessary for use in the other databases. Appendix 1 of the supplementary online material provides an example of the search strategy that was used for PsycINFO. Backward searching of papers listed in the reference sections of articles identified for inclusion in the review was also performed.

Inclusion Criteria

The inclusion criteria were:

(i) The study included a measure of “attitudes to ageing”. In line with conceptual definitions of attitudes (Janeckova, Dragomorecka, Holmerova & Vankova, 2013; Kornadt & Rothermundt, 2015; Levy, 2003; Steverink, Westerhof, Bode & Dittmann-Kohli, 2001), the measure could relate to affective (e.g., fears, concerns, feelings, or worries about ageing), cognitive (e.g., experiences, perceptions, beliefs or and expectations related to growing older), and/or behavioral (e.g., activities or decisions connected with ageing) components.

(ii) An identified aim of the study was to assess one or more of the psychometric properties of the measure. This is in line with published recommendations for
performing systematic reviews of measurement properties (Terwee, de Vet, Prinsen & Mokkink, 2011a), which advise against including studies in which only indirect evidence about measurement properties of the target measures are reported, for instance assessing responsiveness of a measure as part of a randomized control trial (RCT) of an intervention study, due to problems in identifying and interpreting such evidence. If the attitudes to ageing scale was a subscale of a broader measure, then the article was only included if psychometric properties of the relevant subscale(s) were examined.

(iii) The majority of study participants, or the majority of a discrete subsample that was analyzed independently, were adults under 60 years of age. This was operationalized as either: at least 50% of the participants were aged 18 to 60 years, or the mean or median age of the sample was lower than 60 years. Studies that did not provide an age breakdown, but which described their sample using terms that implied that they were young adults (e.g. ‘students’ or ‘employees’ etc.) were included, whilst studies using terms that implied that the sample were older adults (e.g. “elderly” etc.) were excluded.

(iv) The measurement tool must be publicly available (either free to view; for purchase; or provided by the authors on request), so that it could be utilized in future studies.

(v) The study must have been published in a peer-reviewed journal. Grey literature articles were excluded. This is because non-peer reviewed literature in this field is subject to higher levels of bias due to political motivations (e.g., by those seeking to exaggerate the negative effects of ageing), and vested interests (e.g., business interests) in presenting certain measurement tools more or less favorably.
Study Selection

After removing duplicate entries, one member of the research team (FNMF) and one researcher external to the team each independently screened a subset of titles and abstracts, and then compared the decisions made. The initial process involved 300 of 5509 records (approximately 5%) of titles and abstracts screened by both researchers. This resulted in some discrepancy between the first and second researcher, with the first researcher (FNMF) retaining 15 articles, while the second researcher retained just nine. Through discussion, the researchers clarified the inclusion criteria, and agreed that more than nine of the articles should have been retained. Following this calibration exercise, another 300 records were independently screened by the same two researchers. This time both researchers retained the same 11 articles, indicating that the screening process was reliable. The remaining records were then screened by the first researcher only. A similar process was used when screening full text articles. This time, both researchers identified the same four articles from the first 25 (approximately 10%) of full texts as meeting the inclusion criteria. As agreement was so high, the remaining full text articles were screened by the first researcher only, with any ambiguities resolved through discussion with one or more other team members.

Data Extraction

Data relating to the characteristics of the scale; the nature of the attitudinal construct being assessed (i.e. whether items related to participants’ experiences of ageing; their future own ageing; ageing in general; and/or anything else); the country and sample in which it was assessed; and the assessed psychometric properties of the scale were extracted by one researcher
(FNMF or LJEB). For articles describing the original development of a scale, data relating to the process of scale development were also retrieved. For those articles that did not describe the original scale development, the original paper that described the scale development (regardless of the age group it was developed in) was retrieved, and data relating to the process of scale development was extracted. Data relating to the psychometric properties of the measures was extracted by FNMF, and then checked for accuracy by a second researcher (LJEB). Authors of papers were contacted, when necessary, to provide additional information not reported in the papers.

**Assessment of Study Quality**

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist by Terwee, Mokkink, Knol, Ostelo, Bouter and de Vet (2011b) was used to assess the methodological quality of the assessment of each psychometric property in each study. This checklist contains nine sets of items that relate to the following specific measurement properties: internal consistency (A), reliability (B), measurement error (C), content validity (D), structural validity (E), hypothesis testing (F), cross-cultural validity (G), criterion validity (H), and responsiveness (I). Boxes (A) to (C) are grouped into the reliability domain, whilst boxes (D) to (G) are grouped into the validity domain. An additional box of items is used to assess the requirements of studies that apply Item Response Theory (IRT) methods. The COSMIN Checklist is a modular tool, meaning that only the boxes relating to the specific psychometric properties assessed in a study need to be completed. Each item within a section is scored using a four-point scoring system (0 = poor, 1 = fair, 2 = good, and 3 = excellent). For example, when scoring the quality of studies assessing internal consistency, items relate to: the sample size used for the analysis (ranging from ratings of ‘excellent’ for studies with 100 or more participants to
‘poor’ for studies with fewer than 30 participants); the percentage of missing items; whether
details of how missing items were handled had been included; whether scale unidimensionality
had been assessed; whether internal consistency statistics had been calculated separately for each
unidimensional (sub)scale; the presence of any important flaws in the study design or
methodology; and whether appropriate statistical methods had been used to assess internal
consistency. A quality score for each box (i.e. each psychometric property assessed) is attained
by using a ‘worse score counts’ method i.e. taking the lowest rating of any item in a box (Terwee
et al., 2011b).

The COSMIN checklist does not yet include a rating box for assessing measurement
invariance. Therefore, following consultation with one of the authors of the COSMIN team,
(C.B.Terwee, personal communication, May 22, 2017), studies assessing measurement
invariance were assessed using items 1-3, and 12-15 (except for the items on translation) of the
cross-cultural validity box (Box G). Two members of the research team (FNMF and LJEB)
performed independent quality assessments of each paper, and then met to discuss and resolve
any discrepancies.

Best Evidence Synthesis: Level of Evidence

To summarize the evidence, data relating to each psychometric property of each
included instrument was first rated using checklists of criteria, adapted from Terwee et al.
(2007). This results in one of the following rating options being applied to each psychometric
property (e.g., internal consistency) assessed in each study: (+) = adequate, (-) = not adequate,
(?) = unclear, or (0) = no information available. For example, when assessing internal
consistency, a rating of ‘adequate’ (+) is given for a unidimensional (sub)scale that has
Cronbach’s alpha value of $\geq .70$, whereas a rating of ‘inadequate’ (-) is given for a (sub)scale that is either not unidimensional or has a Cronbach’s alpha value of $< .70$. As with the COSMIN tool, ratings of measurement invariance were made using the criteria for cross-cultural validity.

Individual ratings were then synthesized across all studies that rated a particular property of a particular tool using the levels of evidence criteria adapted from the Cochrane Back Review Group (van Tulder, Furlan, Bombardier, Bouter & Editorial Board, 2003). The levels of evidence criteria for overall quality of measurement properties have been used in various systematic reviews (e.g. Coombes, Wiseman, Lucas, Sangha & Murtagh, 2016; Dobson, Hinman, Hall, Terwee, Ross & Bennell, 2012). This method takes the number and methodological quality of the studies, and the consistency of results, into account to produce one of eight overall ratings of the level of evidence available for the psychometric property of each measure: strong, moderate, or limited evidence that a property of a scale is adequate or inadequate; conflicting evidence; or unknown. For example, a rating of strong evidence that a property was adequate would be given if there were consistent reports of adequacy either across multiple studies of ‘good’ methodological quality or in one study of ‘excellent’ methodological quality. Evidence summary and synthesis ratings were made by one researcher (FNMF), and then checked by a second researcher (LJEB), with any discrepancies resolved through discussion.

**Results**

**Selection of Studies**

The search of electronic databases yielded a total of 8351 records. After removing 2842 duplicates, a further 5253 articles were excluded through title and abstract screening. A further 257 articles were then excluded at the full-text review stage. One additional article (Ornelas,
Gasté, Jeanette and Judith, 2016) was identified through manual searching of the reference list of an included paper, resulting in 21 articles being included in the review (Figure 1). These 21 articles related to 10 different measures of attitudes to ageing (Table 1).
Figure 1 Flowchart showing the search and study selection process.
Characteristics of Measures

Table 1 provides an overview of the characteristics of the 10 measures included in the review. Full details of the characteristics of all these measures are provided in Appendix 2 of the supplementary online material.

All measures were developed between 1975 and 2014, and in the English language, except for the Personal Experience of Ageing (PEA: Steverink et al., 2001), which was developed in German. Five of the measures (AAQ: Laidlaw et al., (2007); Ageing Perception Questionnaire (APQ: Barker et al., 2007); Attitudes to Own Ageing (ATOA: a subscale of the Philadelphia Geriatric Center Morale Scale (PGCMS: Lawton, 1975); Expectations Regarding Ageing (ERA-38: Sarkisian et al., 2002); and ERA-12 (Sarkisian et al., 2005) were originally developed for assessing the attitudes of older people, and then subsequently applied to, or modified for use with adults younger than 60. The other five measures were developed for various age ranges, including younger adults. Specifically: the PEA (Steverink et al., 2001) was developed for adults aged 40-69 years; the Attitude- Aging- Visual Analogue Scales (At-Ageing-VAS: Ligon, Ehlman, Moriello, Russo & Miller, 2014) was developed for younger adults, aged from 18 to 42; whereas the Anxiety about Ageing Scale (AAS: Lasher & Faulkender, 1993); Reactions to Ageing Questionnaire (RAQ: Gething, 1994); and Personal Anxiety toward Ageing (PAA: Kafer, Rakowski, Lachman & Hickey, 1980) were all developed for a broad age range from young (18 years/twenties) to older adults (65 years +).

The number of items in each measure ranged from one (At-Ageing-VAS) to 38 (ERA-38). Seven scales contained subscales, with the numbers of subscales ranging from three (ERA-12, AAQ, PEA) to 10 (ERA-38). The items of five measures were developed from focus groups
with middle-aged or older adults (ERA-38; ERA-12; ATOA; PEA) or health care participants (RAQ). The items of three measures (AAS; PAA; and APQ) were developed from literature reviews, whilst the items of the AAQ were developed using both a literature review and focus groups. The single-item of the At-Ageing-VAS was developed by the researchers and later refined by experts in gerontology.

Most of the multi-item measures contained items that related to more than one attitudinal construct. Five of these contained items that related to participants’ experiences of ageing (AAQ; AAS; APQ; ATOA; PAA); seven to their future ageing (AAS; APQ; ERA-38; ERA-12; PAA; PEA; and RAQ); and six to ageing in general (AAQ; ATOA; ERA-38; ERA-12; PEA; and some versions of the RAQ). Two measures (the AAS and the 15-item version of the PAA) also included items that asked about attitudes to older people. The single-item At-Ageing-VAS, and some of the items in the PEA, asked participants about their ‘own ageing’, without specifying whether this was current and/or future ageing. Only one multi-item measure (the RAQ) was available in versions that had items and/or instructions relating to just one construct, namely participants’ future ageing.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Validation studies reviewed</th>
<th>No. of items (sub-scales)</th>
<th>Populations validated in</th>
<th>Languages validated in</th>
<th>Attitudinal construct(s) covered</th>
<th>Level of evidence for each psychometric property assessed for the full-scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attitudes to Ageing Questionnaire (AAQ)</strong></td>
<td>Brown et al., (2015)</td>
<td>24 items (3 sub-scales)</td>
<td>Middle aged general population aged 40-60 years</td>
<td>English</td>
<td>Experiences of Ageing; Ageing in General.</td>
<td>Structural validity (?); measurement invariance (?)</td>
</tr>
<tr>
<td><strong>Anxiety about Ageing Scale (AAS)</strong> (Lasher &amp; Faulkender, 1993).</td>
<td>Lasher &amp; Faulkender (1993); Gao (2012); Koukouli et al., (2013); Mir &amp; Mir (2014); Sargent-Cox et al., (2014); Ornelas et al., (2016)</td>
<td>Original has 20 items (4 sub-scales). Modified versions have 14-16 items.</td>
<td>Student samples ranging from 18-45 years; General population aged 20-97; Healthcare professionals aged 18-65.</td>
<td>English; Chinese; Greek; Persian; Spanish</td>
<td>Experiences of Ageing; Future own Ageing; Attitudes to Older People</td>
<td>Internal consistency (++); content validity (+); structural validity (++); cross-cultural validity (?); hypothesis testing (+); measurement invariance (-)</td>
</tr>
<tr>
<td><strong>Ageing Perceptions Questionnaire (APQ)</strong> (Barker et al., 2007)</td>
<td>Ingrand et al., (2012)</td>
<td>32 items (7 sub-scales)</td>
<td>Middle-aged and older general population aged 55-64.</td>
<td>French (also available in English)</td>
<td>Experiences of Ageing; Future own Ageing</td>
<td>Internal consistency (+++); Structural validity (?); cross-cultural validity (?); measurement invariance (+).</td>
</tr>
<tr>
<td>Scale Name</td>
<td>Authors/Year</td>
<td>Items/Version</td>
<td>Sample Details</td>
<td>Language(s)</td>
<td>Subscales/Options</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Attitudes Toward Aging VAS</strong> (At-Aging-VAS)</td>
<td>Ligon et al., (2014)</td>
<td>1 item</td>
<td>Students aged 18-42.</td>
<td>English</td>
<td>Own Ageing Reliability (+); content validity (+); hypothesis testing (--);</td>
<td></td>
</tr>
<tr>
<td><strong>Attitude Toward Own Ageing</strong> (ATOA)</td>
<td>Miche et al., (2014); Jung &amp; Siedlecki (2018)</td>
<td>Original has 5 items. One modified version has 4 items.</td>
<td>Middle-aged general population samples (mean ages ranging from 43.8 (SD=0.9) to 55.3 (SD=1.1)).</td>
<td>German (also available in English)</td>
<td>Experiences of Ageing; Ageing in General (5-item version only). Internal consistency (+); structural validity (?); hypothesis testing (+); measurement invariance (±).</td>
<td></td>
</tr>
<tr>
<td><strong>Expectations Regarding Ageing -38 (ERA-38)</strong></td>
<td>Sparks et al, (2013)</td>
<td>Original has 38 items (10 sub-scales). Modified version has 22 items (6 sub-scales)</td>
<td>Middle-aged and older general population aged 45-74.</td>
<td>English</td>
<td>Future own Ageing; Ageing in General. Internal consistency (+); structural validity (+)</td>
<td></td>
</tr>
<tr>
<td><strong>Expectations Regarding Ageing -12 (ERA-12)</strong></td>
<td>Joshi et al., (2010); Park &amp; Kweon (2014)</td>
<td>12 items (3 sub-scales)</td>
<td>Middle-aged and older general population samples ranging from 40-75.</td>
<td>English; Korean</td>
<td>Future own Ageing; Ageing in General. Internal consistency (+++); content validity (?); structural validity (++; cross-cultural validity (?); hypothesis testing (++).</td>
<td></td>
</tr>
<tr>
<td><strong>Personal Anxiety toward Ageing</strong> (PAA): subscale of the Ageing Opinion Scale (AOS)</td>
<td>Kafer et al.(1980); Lynch (2000)</td>
<td>Original has 15 items. Modified version has 6 items.</td>
<td>General population samples aged 18-64; Students (mean age 21.9</td>
<td>English</td>
<td>Experiences of Ageing; Future own Ageing; Attitudes to Own Ageing</td>
<td>Internal consistency (?); structural validity (?) measurement invariance (±)</td>
</tr>
<tr>
<td><strong>Personal Experience of Ageing (PEA)</strong> (Steverink et al., 2001)</td>
<td><strong>Reactions to Ageing Questionnaire (RAQ)</strong> (Gething, 1994).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Experience of Ageing (PEA) (Steverink et al., 2001)</td>
<td>Reactions to Ageing Questionnaire (RAQ) (Gething, 1994).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 items (3 sub-scales)</td>
<td>Original has 27 items (6 sub-scales). One modified version has 25 items.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-aged and older general population aged 40-85.</td>
<td>Middle-aged and general population aged 44-55 and 18-60. Health professional samples aged 15-69.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German (also available in English)</td>
<td>English; Swedish; Malay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Future own Ageing;</td>
<td>Future own Ageing; Ageing in General (2004 version only).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal consistency (+); structural validity (+).</td>
<td>Internal consistency (+++); structural validity (+); cross-cultural validity (?); hypothesis testing (?).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+++ OR --- Strong evidence that property is adequate/inadequate. i.e. Consistent findings in multiple studies of ‘good’ methodological quality OR in one study of ‘excellent’ methodological quality;
++ OR -- Moderate evidence that property is adequate/inadequate. i.e. Consistent findings in multiple studies of ‘fair’ methodological quality OR in one study of ‘good’ methodological quality;
+ OR - limited evidence that property is adequate/inadequate. i.e. One study of ‘fair’ methodological quality;
± Conflicting evidence. i.e. Conflicting findings
? Unknown; i.e. Only studies of ‘poor’ methodological quality OR lack of relevant information reported.
Details of the Validation Studies

Summary details of the populations and languages in which each measure has been validated, and any revisions that have been made to the measures, are presented in Table 1. Full details of these are presented in Appendix 3 of the supplementary online material.

The five scales that were originally developed for use in older people (AAQ; APQ; ATOA; ERA-38; ERA-12) were validated in younger samples to which they had been developed for. The other five scales (AAS; At-Ageing-VAS; PAA; PEA; RAQ) were validated in the same or similar age groups to which they were developed. Nine of the validation studies resulted in slightly modified versions of the original scales being produced with, for instance, items revised, deleted or moved to different subscales.

The measure that had been validated most frequently in younger populations was the AAS, which was validated in six studies (Gao, 2012; Koukouli, Pattakao-Parasyri & Kalaitzaki, 2013; Lasher & Faulkender, 1993; Mir & Mir, 2014; Ornelas et al., 2016; Sargent-Cox, Rippon & Burns, 2014), followed by the RAQ ($n = 4$) (Gething, 1994; Gething et al., 2004; Netz, Guthrie, Garamszegi & Dennerstein, 2001; Faudzi, Armitage, Bryant & Brown, 2018). Two studies validated the ERA-12 (Joshi, Malhotra, Lim, Ostbye & Wong, 2010; Park & Kweon, 2014); the PAA (Kafer et al., 1980; Lynch, 2000); and ATOA (Miche, Elsasser, Schilling & Wahl, 2014; Jung & Siedlecki, 2018). The other five measures were validated in just one study each (e.g., ERA-38; AAQ; APQ; PEA; AT-Ageing-VAS).

All articles were written in English, except those by Ornelas et al. (2016) and Park and Kweon (2014), which were written in Spanish and Korean, respectively. Most of the measures were validated in their original language: German for the PEA, and English (in Australian,
British, American, Canadian, and Singaporean populations) for the remainder. Translated versions of some measures were also validated. Authors of eight of these studies performed the translations themselves. These were for the AAS (which was translated into Chinese, Greek & Persian); RAQ (Swedish, Malay); ERA-12 (Korean), APQ (French); and ATOA (German), although the translation process was not described for the latter scale (Miche et al., 2014; Jung & Siedlecki, 2018). One study (Ornelas et al., 2016) used an existing Spanish version of the AAS by Rivera-Ladesma, Lena, Rangel and Sanchez-Sosa (2007) in a Mexican population. Quality ratings of the translation procedures for this study (i.e. items 4-11 cross-validity box of COSMIN) were therefore made on the basis of information from the Rivera-Ladesma et al., (2007) paper.

**Measurement Properties and Methodology Quality Assessments**

Details of which properties were assessed for each scale are presented in Table 1. Full details of the assessment ratings given to each study that assessed each property are presented in Appendix 4 of the supplementary online material.

The most commonly assessed properties were structural validity (n=18) and internal consistency (n = 17), followed by cross-cultural validity (n = 8, although seven of these described the translation process only); measurement invariance (n = 7); hypothesis testing (n = 6); content validity (n = 3); and test-retest reliability (n = 1). None of the measures included in our review had had all its measurement properties assessed. Quality ratings varied from poor to excellent, with ‘fair’ being the most frequent rating across all domains assessed, except for cross-cultural validity, for which seven out of eight studies assessed were rated as being ‘poor’ quality.

**Data Synthesis**
The synthesised results relating to the psychometric properties of the ten measures are summarized in Table 1. Full details of these syntheses (including the synthesised results of the subscales of each measure) are presented in Appendix 5 of the supplementary online material.

Attitudes to Ageing Questionnaire (AAQ)

The AAQ was originally designed to assess the experience and attitudes of older adults with regards to the ageing process (Laidlaw et al., 2007), but was used in a middle-aged sample in Brown et al. (2015). However, there was insufficient evidence to draw conclusions about the structural validity and measurement invariance properties for the AAQ in this sample.

Anxiety about Ageing (AAS)

The AAS was designed to assess anxieties about ageing as distinct from other influences on anxiety (e.g., death, see Lasher & Faulkender, 1993). There is moderate evidence for adequacy of internal consistency and structural validity using the total score of the AAS. However, only one of the four subscales (the ‘Fear of Old People (FOP) subscale) has the same moderate level of evidence for internal consistency, with conflicting evidence for the other three subscales. Limited evidence of adequacy was found for the full scale and all four subscales for hypothesis testing, whereas limited evidence of inadequacy of measurement invariance was seen for three of the four subscales only. Conclusions about the cross-cultural validity of the translated versions of the scale or subscales could not be made due to the poor quality of the studies assessing this.

Ageing Perception Questionnaire (APQ)

The APQ was designed to assess self-perception of ageing (Barker et al., 2007). Strong evidence of high internal consistency was found for five of the seven subscales of the APQ, with strong
evidence of inadequacy for the other two. There was limited evidence to show that the total APQ score was invariant across age groups. However, there was insufficient evidence to draw conclusions about assessment of the structural or cross-cultural validity of the total score of the APQ.

Attitude- Aging- Visual Analogue Scales (At-Ageing-VAS)

The At-Ageing-VAS is a single-item measure specifically designed to capture potential changes in attitudes toward one’s own ageing process (Ligon et al., 2014). Limited evidence for the adequacy of test-retest reliability and content validity was found for this scale. However, moderate evidence of inadequacy was found for hypothesis testing.

Attitudes toward Own Ageing (ATOA)

The ATOA, a subscale of PGCMS, was developed to assess self-perception of ageing (Lawton, 1975). Limited evidence of adequacy was found for the internal consistency and hypothesis testing of this subscale, while evidence relating to the adequacy of measurement invariance was to be conflicting. There was insufficient evidence to draw conclusions about assessment of the structural validity of the ATOA subscale.

Expectations Regarding Ageing-38 (ERA-38)

The ERA-38 was originally designed to measure older adults’ expectations regarding ageing (Sarkisian et al., 2002), but was validated here in a study involving participants aged over 45 years (Sparks, Meisner & Young, 2013). Limited evidence of adequate internal consistency was found for the total score and all subscales of the ERA-38 (Table 4). A similar level of evidence
was also found for structural validity of the total score of ERA-38 (and for both of its subscales: Ageing Self-expectation (ASE) & General Ageing Expectation (GAE)).

Expectations Regarding Ageing-12 (ERA-12)

The ERA-12 is a short form of the ERA-38 (Sarkisian et al., 2005). Strong evidence for adequate internal consistency was found for the full scale and two of the three subscales of the ERA-12. Moderate evidence for adequacy of its structural validity and hypothesis testing was also found for the full-scale, with more mixed results relating to hypothesis testing using the subscales. There was insufficient evidence to draw conclusions about content and cross-cultural validity.

Personal Anxiety toward Ageing (PAA)

The PAA subscale relates to one’s anxiety, uneasiness, fear, or dread concerning aging (Kafer et al., 1981). There was insufficient evidence available for its internal consistency and structural validity; while conflicting evidence was found for its measurement invariance.

Personal Experience of Ageing (PEA)

The PEA was designed to measure the personal experience of ageing (Steverink et al., 2001). Limited evidence was found for the adequacy of the structural validity of the full-scale, and for the internal consistency of all three subscales.

Reactions to Ageing Questionnaire (RAQ)

The RAQ was designed to assess how individuals anticipate their own personal ageing in future (Gething, 1994). Strong evidence for adequate internal consistency across the full scale, and two of the subscales of the original version of the RAQ, was found. Moderate evidence for adequate internal consistency of four subscales of the revised version of the RAQ was also found (Table
4). However, differences in the number and content of the factor structures reported and assessed across studies meant that it was not possible to synthesise data relating to the subscales across all studies. Conflicting evidence was reported for structural validity of the whole scale, while insufficient evidence was available to draw conclusions about cross-cultural validity (Gething et al., 2004 version) and hypothesis testing (Gething, 1994 version; Faudzi et al., 2018) of the RAQ.

**Discussion**

The aims of the present systematic review were to identify and evaluate the psychometric properties of measures of attitudes to ageing that have been validated in younger adult samples. Twenty-one studies, describing ten measures were identified. These ranged from single-item to multi-scale measures, which were suited to a range of different age groups, measured several attitudinal constructs, and had been translated into ten different languages. However, whilst assessments of structural validity and internal consistency were available for most studies, other psychometric properties were assessed less frequently. The quality of the studies assessing psychometric properties was also rated as fair or poor in many cases, which limits the strength of the conclusions that can be drawn about the psychometric properties of each measure.

The review revealed considerable diversity in terms of the nature of the attitudinal constructs covered, and the populations to which the measures have been targeted and validated in. One key dimension on which the tools differed was the age group for which they had originally been developed. Five measures identified in the present review (AAQ, APQ, ATOA, ERA-12, ERA-38) were initially developed for evaluating attitudes to ageing in people aged over 60, but have subsequently been administered to middle aged (40 years and older), but not younger, adults. For example, for one measure (the ERA-38), several items had been
removed leaving those 22 items that were deemed suitable for middle age participants (Sparks et al., 2013). Another scale (PEA) had been directly developed for middle-aged and older adults. Given that the items in some of these measures (AAQ, APQ, ATOA) required participants to reflect on their own ageing experiences, it is unclear how well these measures would perform in people younger than 40. In contrast, four measures (the AAS; At-Ageing-VAS; PAA and RAQ) had been initially developed for adults as young as 18 years old. Despite this, two of these measures (AAS and PAA) still contained at least some items that required participants to reflect on their own ageing experiences, highlighting the importance of considering the nature of the attitudinal construct that is assessed by each measure, as well as the ages of the populations for whom it has been developed for, or validated in.

Another key dimension on which the tools differed was in the number of languages/cultures in which the measures had been evaluated. Seven measures have been validated in Western countries, and just three (AAS, RAQ & ERA-12) in Eastern countries. This is important given that broad cultural differences between Eastern and Western countries are believed to correspond to differences in the content and valence of attitudes to ageing (Lockenhoff, Lee, Buckner, Moreira, Martinez & Sun, 2015). In particular, the historical philosophical and ecological determinants of many Eastern cultures means that stable, long-term and familial relationships are more likely to be favoured over the development of new relations with people from outside of the ‘in-group’ (Voronov & Singer, 2002), and that values such as filial piety and respect for older people are more commonly practiced, than in Western cultures (Lin & Bryant, 2009). Indeed, one of the few studies that adapted a Western measure of attitudes to ageing to an Eastern (Malaysian) population identified a new sub-scale containing items relating to family relationships and religion (Faudzi et al, 2018) that had not been shown when
the scale was originally developed in a Western culture (Gething, 1994). Such findings highlight the need to ensure that measures have been carefully adapted to, or validated in, particular cultural and linguistic groups to ensure that they fully and accurately capture the attitudes held in these populations.

With regard to the second aim of the present study, the measures that received positive ratings for the greatest number of psychometric domains were the ERA-12 and the AAS, although some mixed or negative results were also found for some of the psychometric properties of these scales or their composite subscales. Whilst the ERA-12 has only been validated in middle-aged populations, none of its items explicitly ask participants to reflect on their own experiences of ageing, which suggests that it would also be suitable for use with younger participants. In contrast, whilst the AAS has been validated in a wider range of younger adults, a couple of its items actually refer to experiences of ageing (e.g. ‘When I look in the mirror, it bothers me to see how my looks have changed with age’). These items may therefore be less well-suited to, or at least, may be interpreted differently by, younger participants for whom age-related changes are not yet salient: perhaps accounting for the inadequate measurement invariance identified for this scale. On the basis of the current evidence, the ERA-12 may therefore be the best option for researchers wishing to assess attitudes to ageing in a wider range of younger adults, whilst the AAS might be better-suited to studies with middle-aged and older adults. Another questionnaire that may be particularly worthy of consideration is the RAQ, which, as well as showing some evidence of adequacy for some psychometric properties, is the only multi-item measure available in versions that tap into a single attitudinal construct (attitudes towards future own ageing), and so may be particularly useful for researchers who wish to isolate this construct.
The strength of conclusions relating to the psychometric properties of the measures reviewed was limited by the low quality of many of the studies used to assess them. A common reason for poor quality ratings being given was due to either the percentage of missing data not being reported, or no explanation of how missing items were handled being given. Other common reasons for studies being rated as lower quality were: not reporting the internal consistency for each subscale independently (for assessments of internal consistency); the selection of items using the target populations being inadequately performed (for assessments of content validity); only performing exploratory factor analysis, where confirmatory factor analysis would have been more appropriate (for assessment of structural validity); no information being available on the psychometric properties of the comparator instrument (for assessment of hypothesis testing); and the samples being not similar for characteristics, such as educational background and age groups (for assessment of measurement invariance).

Other common methodological issues related specifically to measures that had been translated and/or validated for use in a different language or cultural group. Seven of these studies were rated as being of ‘poor’, and one study rated as being of ‘fair’, quality in terms of their descriptions of the translation and cross-cultural validation procedures that were adopted. The most common reason for the “poor” rating was that the translated instrument was not pre-tested in the target population before use to check the cultural significance and interpretation of the translation. Another reason for cross-cultural assessment being rated as lower quality was due to a lack of clarity as to whether translators worked independently. Researchers conducting future assessments of the psychometric properties of attitudes to ageing measures should therefore address these issues in order to improve the strength of evidence available.
Whilst the present review comprises a systematic search of the literature that aimed to assess the psychometric properties of measures of attitudes to ageing in younger populations, it should be noted that not all studies reporting relevant metrics will have been included. This is because, in line with consensus recommendations for performing systematic reviews of measurement properties (Prinsen et al., 2018; Terwee, de Vet, Prinsen & Mokkink, 2011a), only studies that listed the assessment of one or more psychometric properties as an explicit aim were eligible for inclusion. Additional, indirect, evidence about the psychometric properties of relevant measures may therefore be available from other studies in which a psychometric property of a scale may have been checked as an adjunct to a separate research question. Data collected and reported this way are more difficult to interpret (Terwee, de Vet, Prinsen & Mokkink, 2011a), and also cannot be comprehensively gathered though a systematic search of the literature (Mokkink et al., 2017). Thus, whilst the data reported in this study cannot be considered an exhaustive picture of the psychometric properties of the available measures, they are arguably likely to provide the clearest, and most objective, summary.

Conclusion

This is the first study to systematically identify and evaluate measures of attitudes to ageing that have been validated in younger adult populations. From the studies reviewed here, the ERA-12 and the AAS were the measures that received positive ratings for the greatest number of psychometric domains. Some measures, including the AAS, include items that asked participants to reflect on their own experiences of ageing, and so may be better suited to middle-aged and older adults. Other measures, including the ERA-12, only require participants to reflect on their attitudes to ageing in general, or to their future expectations of their own ageing, and so may be better suited a wider range of younger age groups. Future research should focus on developing or
adapting measures for a wider range of linguistic and cultural groups, and conducting high quality assessments of a wider range of psychometric properties.

**Acknowledgments:** The research was supported by the NIHR Manchester Biomedical Research Centre and the NIHR Greater Manchester Patient Safety Translational Research Centre, and a PhD studentship from the Malaysian Ministry of Education awarded to FNMF. We would also like to thank the three anonymous reviewers for the very useful comments they made on an earlier version of this manuscript.

**Declarations of interest:** None
References


Reactions to Ageing Questionnaire, Assessment. Advance online publication https://doi.org/10.1177/1073191118766400


### Appendix 1 Search strategy (using PsycInfo database)

<table>
<thead>
<tr>
<th>Group Term 1</th>
<th>Group Term 2</th>
<th>Group Term 3</th>
<th>Group Term 4</th>
<th>Group Term 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combine by ‘OR’</strong></td>
<td><strong>Combine by ‘OR’</strong></td>
<td><strong>Combine by ‘OR’</strong></td>
<td><strong>Combine by ‘OR’</strong></td>
<td><strong>Combine by ‘OR’</strong></td>
</tr>
<tr>
<td><strong>PSYCHOMETRIC</strong></td>
<td><strong>MEASURE (62)</strong></td>
<td><strong>ATTITUDE (64)</strong></td>
<td><strong>AGING (60)</strong></td>
<td><strong>ATTITUDE &amp; AGING CONCEPT (61)</strong></td>
</tr>
<tr>
<td>(63)</td>
<td>(65)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometr*</td>
<td>measure*</td>
<td>Attitude*</td>
<td>Ag?ng</td>
<td>Attitude* adj4 ag?ig</td>
</tr>
<tr>
<td>Rating Scales/Factor analysis/Test construction/ Test Reliability/Self-report/Foreign Language</td>
<td>attitude measures/ attitude measurement tool*</td>
<td>Perception*</td>
<td>Aging (attitudes toward)/ aging Perception* adj4 ag?ng</td>
<td></td>
</tr>
<tr>
<td></td>
<td>instrument* scale*</td>
<td>Perception/ Belief*</td>
<td></td>
<td>Belief* adj4 ag?ng</td>
</tr>
<tr>
<td></td>
<td>questionnaire* survey*</td>
<td>View*</td>
<td></td>
<td>View* adj4 ag?ng</td>
</tr>
<tr>
<td>Structure/ Test Validity/ Psychometrics/ Valid* Statistical validity/ Reliab* Statistical reliability/ interrater reliability/ test reliability/</td>
<td>inventor* assessment*</td>
<td>Anxiety/Test anxiety Opinion*</td>
<td></td>
<td>Perspective* adj4 ag?ng</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opinion*</td>
<td></td>
<td>Reaction* adj4 ag?ng</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Experience*</td>
<td></td>
<td>Opinion* adj4 ag?ng</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expect*</td>
<td></td>
<td>Experience* adj4 ag?ng</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Self-perception*” adj4 ag?ing</td>
<td></td>
<td>“Self perception*” adj4 ag?ing</td>
</tr>
</tbody>
</table>
Predict* Prediction/
“internal consistency”
“Cronbach* alpha*”
Correlat*
Statistical correlation/
“item* select*”
“item*reduct*”
“test-retest”
Translat*
Explor*
Statistical regression/
“split-half*”
Cross-cultur* adj3 adapt*
Culture* adj3 adapt*

attitudes
‘self-perception*’
Self-perception/ aging
‘self perception*’

‘Self-stereotyp*’
‘Self stereotyp*’
‘age-stereotyp*’
‘age stereotype*’

“Self-stereotyp*” adj4 ag?ing
“Self stereotyp*” adj4 ag?ing
“age-stereotyp*” adj4 ag?ing
“age stereotype*” adj4 ag?ing

Combine by ‘AND’
## Appendix 2: Full details of the Original Scales used in the Validation Studies

<table>
<thead>
<tr>
<th>Scale &amp; Authors</th>
<th>Country &amp; Language</th>
<th>Initial Process of Scale Development</th>
<th>Scale Structure</th>
<th>Response Option &amp; Scoring</th>
</tr>
</thead>
</table>
| **Attitudes to Ageing Questionnaire (AAQ)** (Laidlaw et al., 2007). | Worldwide (e.g: Scotland, Spain, UK, USA) (English) | - A set of items were generated from a literature review and developed through focus groups with 35 older adults (aged 62-95, \( M = 75 \))  
- Further focus groups were run in 15 centers in different cities around the world; involving older adults aged 60 – 80+ years to generate subject areas for inclusion in a measure.  
- A preliminary set of items and subject areas were reviewed in a Delphi exercise involving 15 centers to get feedback on content, translation issue, format and design of the scale.  
- From the feedback from focus group and Delphi exercise, 44 items were generated and administered in pilot study involving 1356 older adults aged over 60 years old. A series of EFAs and CFAs were conducted and some items were eliminated on the basis of their psychometric properties.  
- 33 items from the preliminary study and an additional five items (from the Center’s responses) were administered to 5566 older adults (aged 60-100 years) in a field trial involving 20 centers. CFA and IRT analyses were run, and some items removed. | 24 items; 3 subscales (i) Psychological Loss (8); (ii) Physical Change (8); and (iii) Psychological Growth (8). | 5 point scale from 1 (strongly disagree) to 5 (strongly agree). Higher total scores indicated more positive attitudes toward one’s own aging process. (Score range: 24 – 120) |
<p>| <strong>Anxiety about</strong>        | USA                | -84 items were derived from a literature review 20 items; 4 subscales                                                                                                                                                                                                                                                                                                                          |                 | 4-point scale from |</p>
<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Country</th>
<th>Language</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ageing Scale (AAS)</td>
<td>(Lasher &amp; Faulkender, 1993)</td>
<td>English</td>
<td>All items were administered to 312 participants, aged from under 25 to over 74, recruited from a university or a foster grandparent program. Items loading above .40 or showing good face validity were retained. (i) Fear of Old People (5); (ii) Psychological Concerns (5); (iii) Physical Appearance (5); &amp; (iv) Fear of Losses (5)</td>
</tr>
<tr>
<td>Ageing Perceptions Questionnaire (APQ)</td>
<td>Ireland and Northern Ireland</td>
<td>English</td>
<td>Older adult focus group participants were asked about their experience of ageing, along with questions based on an adapted version of the Self-Regulation Model (Leventhal, Halm, Horowitz, Leventhal &amp; Ozakinci, 2004). Psychometric properties of 35 items were examined in the first preliminary study (n= 129) and second preliminary study (n = 143) with older adults aged 65 and over, to determine which items should be retained. 32 items; 7 subscales (i) Timeline chronic (5); (ii) Timeline cyclical (5); (iii) Consequences positive (3); (iv) Consequences negative (5); (v) Emotional representations (5); (vi) Control positive (5); &amp; (vii) Control negative (4)</td>
</tr>
<tr>
<td>Attitudes Toward Aging VAS (At-Aging-VAS)</td>
<td>USA</td>
<td>English</td>
<td>The wording of the item and scale anchors was developed by the researchers, and later refined on the basis of feedback from six experts in ageing. One of two single-item visual analogue scales (the other assessing attitudes to older adults). Participants place an ‘X’ on scale of 0% to 100% anchored with the words ‘Negative’ and Positive’ (10 centimeter in length). The number closet to the ‘X’ was considered to be numeric representation of participants’ attitudes to own ageing.</td>
</tr>
</tbody>
</table>
| Attitude | USA | - The PGCMS was developed using 22 items The ATOA subscale contains 5 | Participants provide ‘Yes’
Toward Own Ageing (ATOA) subscale of the Philadelphia Geriatric Center Morale Scale (PGCMS) (Lawton, 1975)

from a morale scale developed by Lawton (1972, cited in Lawton, 1975). These items were validated against ratings given by staff about several hundred older residents’ homes for the elderly. Factor analysis yielded six factor solutions.

- The 22 items were completed by 828 older tenants and community residents (M= 72.6).
- A factor analysis of 17 items with the highest factor loadings from a previous analysis showed three factor solutions.

Expectations Regarding Ageing (ERA-38) (Sarkisian et al., 2002)

- 26 domains were identified from qualitative interviews and focus group with 38 older adults (M = 78 years), and were used to construct a survey.
- 94 items were administered to 58 older participants from senior centers in pilot study. Items with many missing responses or that had extreme distributions were eliminated.
- 56 items were then administered to 588 older adults. Items with high ceiling effects or that correlated less than .30; were redundant; or had lower correlation with all scales were removed.
- 38 items finally were administered to 429 older adults (M = 76, SD = 6.9) recruited from community residing

ERA-12 (Sarkisian et al., 2005) USA (English & Spain)

- 636 older adults aged from 65-100 (M = 77.5, SD = 6.5) recruited at senior centers

42 items; 10 subscales

38 items; 10 subscales

12 items, 3 subscales

4 point scale from 1 (definitely true) to 4 (definitely false).

Higher values indicate more positive expectations regarding ageing. (Score range: 38 – 152)
- 12 items from ERA-38 (Sarkisian et al., 2002) were selected using principle components analyses; examination of focus group rankings; internal consistency data; and correlations with related variables, to produce a shorter scale representing the same constructs as the ERA-38.

Health (4); & (iii) Cognitive Function (4)

- Higher values indicate more positive expectations on ageing (Score range: 12-48)

### Personal Anxiety toward Ageing (PAA):
- To create the AOS, a 203 item-pool was created from three relevant instruments identified through a literature review, then sorted into 15 categories by two graduate students.
- These content areas were used as a guide to generate 120 new items (Form 1). These were administered to 20 adults. Items with minimal variances were removed.
- 60 items (Form 2) were administered in second study to 100 students and 100 practitioners. Items that were factorially complex or had low correlations with the total score were removed. New items were constructed to replace those eliminated items.
- The revised 60 items (Form 3) were administered to 102 students ($M = 21.9, SD = 5.2$) and 118 practitioners ($M = 36.8, SD = 14.8$).
- Factorial analysis revealed four subscales with fifteen items each. One subscale was excluded due to low reliability (Adaptive Ability of Older People, $\alpha = .47$)

#### Personal Anxiety toward Ageing (PAA)
- The subscale that assesses attitudes to ageing contains 15 items.
- Higher scores indicate lower negative stereotypes, anxiety to personal ageing and high perceived social value (Score range: 15-75)

<table>
<thead>
<tr>
<th>Personal Anxiety toward Ageing (PAA): subscale of the Ageing Opinion Scale (AOS) (Kafer et al., 1980)</th>
<th>USA (English)</th>
<th>47 statements were formulated by the research</th>
<th>12 items, 3 subscales</th>
<th>4 point scale from 1 (strongly agrees) to 5 (strongly disagree). Higher scores indicate lower negative stereotypes, anxiety to personal ageing and high perceived social value (Score range: 15-75).</th>
</tr>
</thead>
</table>

- 5 point scale from 1 (strongly agrees) to 5 (strongly disagree). Higher scores indicate lower negative stereotypes, anxiety to personal ageing and high perceived social value (Score range: 15-75).
### Experience of Ageing (PEA) (German)

Steverink et al., 2001

- All 47 items were administered to 4034 Community dwelling participants. 1446 (36%) of participants were aged 40-54; 1475 (36.6%) aged 55-69. \( M = 60 \).
- After a series of EFA, items with factor loading below .4 were removed; 30 items were retained.
- Another factor analysis revealed five factors. The three factors with internal consistency values of at least .70 were retained.

(i) Physical decline (4); (ii) Continuous Growth (4); (iii) Social Loss (4)

Higher score indicated more positive experience of ageing. (Score range: 12 – 48)

### Reactions to Ageing Questionnaire (RAQ) (English)

Australia

Gething, 1994

- 35 items were first developed from the responses of 60 nurses who were asked to write open-ended responses about how they felt about their own ageing. These items were administered to 300 nurses and 150 members of the general public in a pilot study.
- Items with poor reliability, or that loaded inconsistently, were excluded. A shorter version with 30 items was then administered to 531 health professionals aged 20-69 in main study.
- Items that did not achieve high loading were removed.

27 items; 6 subscales

(i) Anxiety about the future (7); (ii) Physical Well-being (7); (iii) Psychological Well-being (5); (iv) Denial of Ageing (3); (v) Isolation (2); & (vi) Activity (3).

6-point scale from 1 (agree very much) to 6 (I disagree very much)

Higher scores indicate more positive attitude towards ageing. (Score range: 27 – 162)
## Appendix 3: Details and results of the validation studies identified in the review.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Country &amp; Scale Validated Version</th>
<th>Study Design</th>
<th>Sample Characteristics</th>
<th>Version of Measure Assessed</th>
<th>Modifications Made to Measure</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAQ</td>
<td>Brown et al., (2015) Australia (English)</td>
<td>Longitudinal</td>
<td>T1 (2003): 517; T2 (2013): 492 Midlife community dwelling participants. 40-60 years old ($M = 52.1$ years; $SD = 5.49$)</td>
<td>Original AAQ by Laidlaw et al., (2007).</td>
<td>No modifications.</td>
<td>NA</td>
<td>-Structural Validity: CFA showed the original model fitted poorly. For the model without correlated residuals ($\chi^2 = 1347$, CFI = .85, RMSEA = .098); and with correlated residuals ($\chi^2 = 1109$, CFI = .879, RMSEA = .088). ESEM solution with three factors; and inclusion of correlated residual improved model fit ($\chi^2 = 519$, CFI = .954, RMSEA = .060). No information about amount of variance accounted for presented. -Measurement Invariance: The three factor ESEM solution was fitted to both mid and older age groups, with the exception of item 13, which varied between groups.</td>
</tr>
<tr>
<td>AAS</td>
<td>Lasher &amp; Faulkender (1993) USA (English)</td>
<td>Cross-sectional</td>
<td>312 Undergraduate</td>
<td>Original AAS development study.</td>
<td>NA</td>
<td>Internal Consistency: Overall (.82)</td>
<td>-Content Validity: The initial 84 items were deemed ‘sufficient’ to capture the theoretical construct</td>
</tr>
</tbody>
</table>
students and adults from elderly foster grandparents program. (aged under 25 - over 74 years)

Cronbach’s alphas for subscales:
Fear of Old People (.78); Psychological Concerns (.74); Physical Appearance (.71); & Fear of Losses (.69).

based on a sorting exercise by three graduate psychology graduates.

-Structural Validity: Exploratory Factor Analysis (EFA) was conducted, and only items loading greater than .4 (and that were considered to have good face validity within the emerging factor structure) were retained. Twenty items (from four factors) were retained, which accounted for 50.60% of the total variance.

-Hypothesis Testing: The four subscales and total score correlated significantly with Self-Efficacy ($r = - .516$) and Facts on Aging Scale ($r = - .193$).

Gao (2012) Taiwan (Chinese) Cross-sectional
334 Students and volunteers
176 Young (19-32, $M = 20.6$ years) & 158 Elderly (50-84, $M = 62.9$ years)

Chinese translation (by author) of original AAS by Lasher & Faulkender (1993)

Scale modified during validation to 16 items across 4 subscales: Fear of Old People (4 items); Psychological Concerns (4 items); Physical Appearance (3 items); Fear of Losses (5 items).

Internal Consistency: (of 16-item scale)

CTT: Fear of Old People (.88); Psychological Concerns (.67); Physical Appearance (.68); & Fear of Losses (.69).

IRT: Fear of Old People (.86); Psychological Concerns (.71);

-Structural Validity: Confirmatory Factor Analysis (CFA) of 20-item version was performed. Two items were removed due to items loading below .60 (items 4 and 20, both from factor 3: physical appearance). Item 5 was modified from factor 2 (Psychological Concerns) to factor 4 (Fear of Losses). No information about amount of variance accounted for provided.

-Cross-cultural Validity: CFA showed that original structure not confirmed.

-Measurement Invariance: IRT of
18 items led to another two items were removed due to differential item functioning between age; young and old groups (items 3 and 17). Measurement invariance across genders was demonstrated.

Koukouli et al., (2013) Greece Cross-sectional 320 participants: 147 health care professionals 18-65 (Mdn = 36-40 years); 74 students 21-45 (Mdn = 21-25 years); 99 community residents 23-59 (Mdn = 31-35 years)

Greek translation (by authors) of the original AAS by Lasher & Faulkender (1993) No modifications.

Internal Consistency:
Overall (.79) Professional (.86), Student (.84), Community resident (.80)

Cronbach’s alphas for subscales:
Fear of Old People (.89); Psychological Concerns (.71); Physical Appearance (.66); & Fear of Losses (.67).

Structural Validity:
CFA of 20-item version was performed. CFA showed unacceptable fit (Model I). Error covariance was identified in item 5 and 17, thus were incorporated in Model II. Due to content overlapped in item 1 and 13 (Model III), the error covariance was also incorporated. Model III was the best fitting model (CFI = .91, RMSEA = .07). No information about amount of variance accounted for provided.

Cross-cultural Validity: Although there was overlapping content, internal consistency values were comparable to original version, and so no changes were made. CFA confirmed original four factor structure.

Mir & Mir (2014) Iran Cross-sectional 84 undergrad Persian translation (by No modifications.

Internal Consistency:

Structural Validity: EFA was conducted and produced the original
36 postgraduate students (21-32 years) authors) of original AAS by Lasher & Faulkender (1993) version

Overall (.76). Further test of internal consistency was conducted by excluding item 7 to see if any improvement is attained or not. Overall scale improved to .84 if removing item 7

Cronbach’s alphas for subscales: Fear of Old People (.81); Psychological Concerns (.71) (Improved from .31 if deleting item 7); Physical Appearance (.76); & Fear of Losses (.59).

-Structural Validity: CFA of 20-item version was performed and supported the original first order four-factor structure of the AAS. It was a better model ($\chi^2=491$, CFI = .911, RMSEA = .050) than second order model ($\chi^2=509$, CFI = .906, RMSEA = .051). No information about amount of variance accounted for presented.


Internal Consistency: Cronbach’s alphas for subscales: Fear of Old People (.80); Psychological Concerns (.80); Physical Appearance (.73); & Fear of Losses (.69).
20-97 ($M = 57.3$ years, $SD = 13.66$)

Measurement Invariance: AAS factor loadings were equal across males and females, except for Fear of Losses domain, where one item (item 17) cross-load onto Psychological concerns in female. AAS factor loadings in mid age showed adequate fit, but a poorer fitted in younger and older age group. For young adults, model fit was enhanced by correlating residual variance of item 8, 12, 14 and 15. For older age group, model fit was improved by cross-loading item 20 from Physical Appearance with Fear of losses. Importantly, invariance across age was only achieved when three items from the Fear of Losses scale (2, 8, and 14) were not constrained, suggesting that loadings for these items were not invariant.

-Structural Validity: CFA of 20-item version was performed. Model M1 is acceptable but not optimal (Subsample 1: CFI = .91, RMSEA = .064, Variance 57% /Subsample 2: CFI = .89, RMSEA = .073, Variance 59%). Therefore, 6 items were deleted in Model M1B (Items 2, 4,5,6,7 and 16). With the remaining 14 items, this model showed optimal result (Subsample 1: CFI = .98,
years ($M = 20.77$ years, $SD = 1.83$) & Fear of Old People (.86); Psychological Concerns (.70); Physical Appearance (.83); & Fear of Losses (.75).

RMSEA = .044, Variance 68%; Subsample 2: CFI = .95, RMSEA = .064, Variance 68%.

Cross-cultural Validity: CFA showed that the original factor structure was not optimal, but within an acceptable range.

APQ


Internal consistency:

Cronbach’s alphas for subscales:
Timeline chronic (.77); Timeline cyclical (.88); Consequences positive (.75); Consequences negative (.82); Emotional representations (.88); Control positive (.62); Control negative (.59).

- Structural Validity: CFA was performed and the seven subscales model produced an acceptable, but not exact, fit: RMSEA = .069; CFI = .95). No information about amount of variance accounted for presented.

- Cross-cultural Validity: Original factor structure confirmed.

- Measurement Invariance: A second model constraining the factor loadings to be equal in the two age groups lead to larger significant value than the first model, which assumed different parameters in the two age groups (<65) and (>65). Model with constrained factor loadings, although significant worse than one with different parameters, still showed a reasonable level of fit (RMSEA = .071; CFI = .94)

At-Ageing-VAS
Ligon et al., (2014) USA Cross-sectional 198 students (Both T1 & T2) 18-42 years ($M = 19.60$ years, $SD = 3.01$) Original Ageing-VAS development study. The anchors of the scale were changed from 0% - 100% to negative-positive. **Reliability:** ICC: 0.84, 95% CI (.780, 0.874) (Time interval: 1 week)

**Content Validity:** The scales and operational definitions were submitted to a panel of six content experts (five gerontology experts, one sociologist) for further input and feedback on the content of the scales. 5/6 agreed that VAS measured the attitudes toward own ageing construct. Reviewers reported concerns regarding the anchor, thus the anchors of the scale were changed.

**Hypothesis Testing** (with modified anchors): Positive significant relations between ATOA-VAS and AAS: Session 1 was .38, and .43 for session 2.

**Measurement Invariance:** CFA showed the 5 items loaded on one common factor in middle age samples and across all three measurement waves. Items 2 and 4 indicated increasing item difficulties in this age group, therefore showed evidence of lack of measurement invariance across age groups.

ATOA subscale (PGMCS)

Miche et al., (2014) German Longitudinal: 12 years. T1(92/93) T2(96/97) T3 (04/05) Middle age: T1= 501 ($M = 43.8$ years, $SD = 0.9$) T2 = 447 ($M = 47.6$ years, $SD = 0.9$) T3 = 408 German translation (by authors) of the 5 items of the original ATOA subscale of the PCGMS (Lawton, 1975) No modifications. NA

Middle aged (40-60 years old) = 2617 ($M = 51.5$, $SD = 5.65$).

Young old (61-74) = 2429 ($M = 67.97$ years, $SD = 3.83$).

Old-old (75+) = 1045 ($M = 79.62$ years, $SD = 3.37$).

German translation of the 5 items of the original ATOA subscale of the PCGMS (Lawton, 1975) drawn from the longitudinal German Ageing Survey of the German Centre of Gerontology.

Scale modified during validation to 4 items

**Internal consistency:**
Cronbach’s alphas for five items (in total sample): (.75).

- **Structural Validity:** CFA was performed on data for the total sample for a one-factor model with 5 items, but the model fit was below the satisfactory range ($\chi^2 = 342.30$, $df = 5$, RMSEA = 0.107, CFI = 0.94). The one-factor model with 4-items (excluding useful) showed improved fit model and was the best fitting model ($\chi^2 = 51.91$, $df = 2$, RMSEA = 0.065, CFI = 0.99). No information about amount of variance accounted for presented.

- **Hypothesis Testing:** Age was significantly correlated with ATOA ($r = -.15$); Life Satisfaction (LS) ($r = .10$); Positive Affect (PA) ($r = -.17$); and Negative Affect (NA) ($r = -.13$). Structural models assessing the relationships between the ATOA items and items assessing related constructs of LS, PA and NA showed significant and moderate positive correlations between the ATOA factor and LS ($r = .70$) and PA ($r = .64$) factors, and weak negative correlations between the ATOA and NA factors ($r = -.30$). ATOA items also continued to load significantly
on the ATOA factor when paths between ATOA variables and the LS, PA and NA variables were added to the model.

-Measurement Invariance: Configural invariance (structure across groups) ($\chi^2 = 49.19$, df = 6, RMSEA = 0.060, CFI = 0.99) and metric invariance (factor loading across group) ($\Delta\chi^2 = 12.00$, $\Delta$df = 6, $\Delta$CFI = 0.002) was observed across the three age groups. However, scalar invariance (item intercept across groups) was not observed across groups, suggesting that the observed intercepts are not invariant ($\Delta\chi^2 = 337.40$, $\Delta$df = 8 $\Delta$CFI = 0.081). Partial scalar invariance across age groups was attained by allowing the intercept of three items (pep, happy and worse) to vary ($\Delta\chi^2 = 10.27$, $\Delta$df = 2, $\Delta$CFI = 0.005). One aspects of structural invariance was also equivalent across age groups. 

Hypotheses Testing

Cross-cultural Validity Measurement Invariant

Sparks et al, (2013) Canada (English) Cross-sectional 167 Adults 45-74 $M = 59.46$ years; $SD$

Original ERA-38 by Sarkisian et al., (2002). Scale modified during validation from 38 items (Sarkisian et al, 2002) to 22 items with six subscales. These

<table>
<thead>
<tr>
<th>Internal Consistency:</th>
<th>Satisfaction/</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAE: (of 12-item scale) Overall (.83)</td>
<td>Cronbach’s alphas for subscales:</td>
</tr>
</tbody>
</table>

-Structural Validity: EFA was conducted separately for each domain. GAE- 12 items loaded on three distinct factors which explained 69.7% of the variance. ASE: 10 items produced three
six subscales are categorized into two domains:

(i) General ageing expectation (GAE) is regarding general aging expectations. (12 items) - Satisfaction/Contentment (4 items); Physical Function (5 items); Cognitive Function (3 items).

(ii) Ageing Self-expectation (ASE) is concerning ageing self-expectation: (10 items) - Functional Health (5 items); Social Health (3 items) & Sexual Function (2 items).

Contentment (.84); Physical Function (.84); Cognitive Function (.88).

ASE: (of 10-item scale) Overall (.77); Cronbach’s alphas for subscales: Functional Health (.80); Social Health (.79) & Sexual Function (.78) distinct factors which explain 67.3% of the total variance.
### ERA-12

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Design</th>
<th>Sample Size</th>
<th>Original Source</th>
<th>Modifications</th>
<th>Internal Consistency</th>
<th>Structural Validity</th>
<th>Hypothesis Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joshi et al., (2010)</td>
<td>Singapore (English)</td>
<td>Cross-sectional</td>
<td>981 Middle aged participants</td>
<td>Original ERA-12 by Sarkisian et al., (2005)</td>
<td>No modifications</td>
<td>Overall ERA-12 &amp; all subscales exceeded .7</td>
<td>EFA of 12-item version was performed; and three factors (physical health; mental health and cognitive function) explained 64% of the variance.</td>
<td>Total ERA-12 had a significant positive correlation with self-rated health (SRH, r = .13) and a negative correlation with Geriatric Depression Scale (GDS-5, r = -.25).</td>
</tr>
<tr>
<td>Park &amp; Kweon (2014)</td>
<td>Korea (Korean)</td>
<td>Cross-sectional</td>
<td>298 Middle-aged and elderly women</td>
<td>Korean translation (by authors) of original ERA-12 by Sarkisian et al., (2005)</td>
<td>No modifications</td>
<td>Overall (.81)</td>
<td>EFA was performed. Three original factors structure were found, explaining 56% overall variance.</td>
<td>Total ERA-12 (r = .33); Physical health (r =.13) and mental health (r =.39) subscales showed significant positive correlation with Rosenberg Self-Esteem Scale (RSES)</td>
</tr>
</tbody>
</table>

**Content validity**: Calculated and verified by one elderly nurse and three nursing professor. The Content Validity Index (CVI) was >.8

**Structural Validity**: EFA was performed. Three original factors structure were found, explaining 56% overall variance.

**Hypothesis Testing**: Total ERA-12 (r = .33); Physical health (r =.13) and mental health (r =.39) subscales showed significant positive correlation with Rosenberg Self-Esteem Scale (RSES)

**Cross-cultural Validity**: CFA was not applied and DIF was not assessed.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Country</th>
<th>Study Design</th>
<th>Sample Size</th>
<th>Sample Description</th>
<th>Measure</th>
<th>Internal Consistency</th>
<th>Structural Validity</th>
<th>Measurement Invariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kafer et al.,</td>
<td>USA</td>
<td>Cross-sectional</td>
<td>118</td>
<td>Practitioners $(M = 36.8$ years, $SD = 14.8$); and 102 students $(M = 21.9$ years, $SD = 5.2$)</td>
<td>PAA subscale from Kafer et al., (1980)</td>
<td>NA</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>(1980)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lynch (2000)</td>
<td>USA</td>
<td>Cross-sectional</td>
<td>979</td>
<td>Adults aged 18-39 $(M = 29$ years, $SD = 6.06$); 216 adults aged 40-49 years &amp; 225 adults aged 50-64 $(M = 56.32$ years, $SD = 4.66$)</td>
<td>Seven indicators of Personal Ageing Anxiety Scale drawn from Kafer et al., (1980)</td>
<td>Scale modified during validation to 6 items, which are viewed as a unidimensional construct of Personal Anxiety toward Ageing.</td>
<td>NA</td>
<td>-Structural Validity: CFA showed the first model with all 7 items loading onto ageing anxiety) was acceptable but not impressive ($X^2$s ranging from 103 – 130; RMSEAs ranging from .034 - .052), Model 2 (in which the error covariance between physical disability (items 5) and loss of cognitive ability to make decision (item 7) was added was better ($X^2$s ranged from 36 – 77). RMSEAs ranged from .019 - .021). Model 3 was better after removal of social loss (y6) ($X^2$ ranged from 20-36; RMSEA ranged from .01 - .023). No information about amount of variance accounted for presented. -Measurement Invariance: The remaining six items tap the same latent dimension between races</td>
</tr>
<tr>
<td>Study</td>
<td>Language</td>
<td>Type</td>
<td>Sample Size</td>
<td>Participants</td>
<td>Validation Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>-------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PEA</strong></td>
<td>German</td>
<td>Cross-sectional</td>
<td>4034</td>
<td>1446 (36%) of participants were aged 40-54 years; 1475 (36.6%) aged 55-69 years. (M = 60 years).</td>
<td>Internal consistency: Cronbach’s alphas for subscales: (i) Physical decline (.79); (ii) Continuous Growth (.78); (iii) Social Loss (.77)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steverink et al., (2001)</td>
<td></td>
<td>Community dwelling</td>
<td>Original PEA development study</td>
<td>NA</td>
<td>12 items (following two initial rounds of item reduction) were retained. The total amount of variance explained by three factors adds to 60.5%.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAQ</strong></td>
<td>English</td>
<td>Cross-sectional</td>
<td>531</td>
<td>Overall for 30 items (.81)</td>
<td>Internal consistency: Cronbach’s alphas for subscales (for 27 items): Anxiety about the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gething (1994)</td>
<td></td>
<td>Health professional</td>
<td>Original RAQ development study</td>
<td>NA</td>
<td>-Structural Validity: EFA was conducted on 30 items. Items that did not achieve high loading were excluded. Six factors with 27 items were derived, which accounted for 36.2% of overall variance. -Hypothesis Testing: RAQ showed significant relationship with Ageing Semantic Differential (r = -.26) and (White &amp; non-white); and sexes, but there are variations in the meaning of indicators across ages.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Future (.78); Physical Well-being (.72); Psychological Well-being (.62); Denial of ageing (.68); Isolation (.49) & Activity (.20)

the Facts about Ageing Quiz ($r=.13$).

**Gething et al.,(2004)**

Australi a, United Kingdom, Sweden (English & Swedish)

Cross-cultural Health professional translation (by authors); and original 27-item RAQ by Gething (1994)

Swedish Scale modified during validation to: Australia: 15 items with 3 subscales: Negativity about Growing Older (8 items); Perceived Personal Attributes (3 items); & Positive Aspect of Ageing (4 items).

UK: 16 items with 3 subscales: Negativity about Growing Older (7 items); Perceived Personal Attributes (3 items).

Sweden

(102: aged 15-69 years)

UK (218: aged 15-69 years)

Australia (147: aged 15-59 years)

Swedish translation (by authors); and original 27-item RAQ by Gething (1994)

Scale modified during validation to: Australia: 15 items with 3 subscales: Negativity about Growing Older (8 items); Perceived Personal Attributes (3 items); & Positive Aspect of Ageing (4 items).

**Internal consistency:** (of 15-item scale for Australia and 16 items for UK and Sweden)

Overall : Australia & Sweden ($\alpha = .87$); UK ($\alpha = .88$)

**Structural Validity:** EFA was conducted in each country. UK: 6 factors were produced that accounted for 57.1% of the variance. Australia: 8 factors explaining 65% overall variance. Sweden: 7 factors accounted for 58.8% variance. Only the first three factors from each country were homogeneous, and so examined in detail. The variances for the three factors were: UK (42.3%); AUS (44%) & SWE (41%).

**Cross-cultural Validity:** 12 items from original version were removed in Australia sample; 11 items were removed in UK and Sweden samples. CFA not applied and DIF not assessed.
Sweden: 16 items with 3 subscales: Negativity about Growing Older (5 items); Perceived Personal Attributes (3 items); & Positive Aspect of Ageing (8 items).

Netz et al., (2001) Australia (English) Cross-sectional 381 Middle aged-women 44-55 years Original 27-item RAQ by Gething (1994) Subscales content revised during validation with 5 subscales: Negative thoughts about the future (6 items); Social/Antisocial behavior in old age (3 items); Positive though about the future

**Internal consistency:** Overall (.91)

- **Structural Validity:** EFA was conducted, and five factors were found, accounted for 47.3% overall variance.
Faudzi et al. (2018). Malay Cross-sectional 911 adults aged from 18 – 60 years old. Malay translation (by authors); original 27-item RAQ by Gething (1994); and 8 additional items by authors based on open ended responses about what people think about ageing when they are 65 years old.

| Internal consistency: Overall for 25 items (.88) |
| Cronbach’s alphas for subscales: |
| Negative Thoughts about Growing Older (.89); Perceived Personal Attributes (.82); Positive Aspect of Ageing (.73); & Family and Religion (.84). |

Scale content revised and modified during validation to: 25 items with 4subscales: Negative Thoughts about Growing Older (11 items); Perceived Personal Attributes (5 items); Positive Aspect of Ageing (6 items); & Family and Religion (3 items).

-Structural Validity: EFA was conducted, and 3 subscales similar to Gething et al., (2004) and Netz et al., (2001) were found. One additional factor (Family & Religion) was also identified. The four factors accounted for 40.61% of variance overall.

-Cross-cultural Validity: Nine items from the original version (Gething, 1994) and one new item (from the authors) were removed. CFA was not applied, and DIF not assessed.

-Hypothesis Testing: There was a main effect of age on RAQ scores \( (p < .001, \eta^2 = .021) \), with higher scores in participants aged 50-60 compared to those aged 18-29 or 30-39 years. There was a main effect of gender \( (p = .041, \eta^2 = .005) \), with men reporting more positive attitudes than women.
### Appendix 4 Methodological quality rating for each property assessed by each study

<table>
<thead>
<tr>
<th>Study</th>
<th>Internal Consistency</th>
<th>Reliability</th>
<th>Content Validity</th>
<th>Structural Validity</th>
<th>Cross-cultural Validity</th>
<th>Hypotheses Testing</th>
<th>Measurement Invariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown et al.</td>
<td></td>
<td></td>
<td></td>
<td>Fair</td>
<td></td>
<td></td>
<td>Poor</td>
</tr>
<tr>
<td>AAS</td>
<td></td>
<td></td>
<td></td>
<td>Fair</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lasher &amp; Faulkender</td>
<td>Fair</td>
<td>Fair</td>
<td>Fair</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gao</td>
<td>Fair (IRT &amp; CTT)</td>
<td>Fair (IRT &amp; CTT)</td>
<td>Poor*</td>
<td></td>
<td></td>
<td></td>
<td>Poor</td>
</tr>
<tr>
<td>Koukouli et al.</td>
<td>Fair</td>
<td></td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mir &amp; Mir</td>
<td>Fair</td>
<td></td>
<td></td>
<td>Fair</td>
<td>Poor*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sargent-Cox et al.</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ornelas et al.</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td>Poor*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingrand et al.</td>
<td>Excellent</td>
<td>Excellent</td>
<td>Poor*</td>
<td></td>
<td></td>
<td></td>
<td>Fair</td>
</tr>
<tr>
<td>At-Ageing-VAS</td>
<td></td>
<td></td>
<td>Fair</td>
<td></td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ligon et al.</td>
<td>Fair</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATOA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michi et al.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jung &amp; Siedlecki</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERA-38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sparks et al.</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERA-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joshi et al.</td>
<td>Excellent</td>
<td></td>
<td>Good</td>
<td></td>
<td>Poor (SHR)/ Fair (GDS-5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park &amp; Kweon</td>
<td>Fair</td>
<td>Poor</td>
<td>Fair</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kafer et al.</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lynch</td>
<td></td>
<td></td>
<td>Excellent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>RAQ</td>
<td>RAQ</td>
<td>RAQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steverink et al.</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Studies that described a translated measure, but for which cross-cultural comparisons had not been performed. For these studies, only items 4 to 11 of the cross-cultural validity box were used to provide a rating of the quality of the translation process only.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rivera-Ladesma et al.</td>
<td>Excellent</td>
<td>Excellent</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gething et al.</td>
<td>Poor (AUS; UK &amp; SWE)</td>
<td>Fair (AUS; UK &amp; SWE)</td>
<td>Poor (UK &amp; SWE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netz et al.</td>
<td>Poor</td>
<td></td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faudzi et al.</td>
<td>Good</td>
<td></td>
<td>Fair *</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The translation procedures used to inform the quality rating for cross-cultural validity for this study came from Rivera-Ladesma et al. (2007).*
### Appendix 5 Quality of Measurement Properties per Questionnaire

<table>
<thead>
<tr>
<th>Measure</th>
<th>Internal Consistency</th>
<th>Reliability</th>
<th>Content Validity</th>
<th>Structural Validity</th>
<th>Cross-Cultural Validity</th>
<th>Hypothesis Testing</th>
<th>Measurement Invariance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AAQ – Total score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological Loss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AAS – Total score</strong></td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Fear Of Old People</td>
<td>++</td>
<td>±</td>
<td></td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Psychological Concern</td>
<td>±</td>
<td>±</td>
<td></td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>Physical Appearance</td>
<td>±</td>
<td>±</td>
<td></td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Fear of Losses</td>
<td>±</td>
<td>±</td>
<td></td>
<td>?</td>
<td>+</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td><strong>APQ – Total score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Timeline chronic</td>
<td>+++</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timeline cyclical</td>
<td>+++</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consequences positive</td>
<td>+++</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consequences negative</td>
<td>+++</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional representations</td>
<td>+++</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control positive</td>
<td>- -</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control negative</td>
<td>- -</td>
<td></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>At-Ageing-VAS – Total score</strong></td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- -</td>
</tr>
<tr>
<td><strong>ATOA (PGMCS) – Total score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+ ±</td>
</tr>
<tr>
<td><strong>ERA-38 – Total Score</strong></td>
<td>+ (GAE &amp; ASE)</td>
<td></td>
<td></td>
<td>+ (GAE &amp; ASE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction/ Contentment (GAE)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Function(GAE)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Function(GAE)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Health (ASE)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Health (ASE)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual Function (ASE)</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ERA-12 – Total score</strong></td>
<td>+++</td>
<td>?</td>
<td>++</td>
<td>?</td>
<td>++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Health</td>
<td>+++</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental Health</td>
<td>±</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Function</td>
<td>+++</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PAA (AOS) – Total score</strong></td>
<td>?</td>
<td>?</td>
<td>±</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PEA – Total score</strong></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical decline</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous Growth</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Loss</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAQ – Total score</strong></td>
<td>+++</td>
<td>±</td>
<td>?</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety about the Future (Gething, 1994)</td>
<td>+++</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Wellbeing (Gething, 1994)</td>
<td>+++</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological Wellbeing (Gething, 1994)</td>
<td>- -</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denial of Ageing (Gething, 1994)</td>
<td>- -</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolation (Gething, 1994)</td>
<td>- -</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity (Gething, 1994)</td>
<td>- -</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negativity about Growing Older (Netz, 2001; Gething, 2004; Faudzi, 2018)</td>
<td>++</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Personal Attributes</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antisocial Behavior (Netz, 2001; Gething, 2004; Faudzi, 2018)</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive Aspect of Ageing (Netz, 2001; Gething, 2004; Faudzi, 2018)</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family &amp; Religion (Faudzi, 2018)</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear of Becoming Frail and Senile (Netz, 01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear of Death (Netz, 01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+++ OR --- Strong evidence that property is adequate/inadequate. i.e. Consistent findings in multiple studies of ‘good’ methodological quality OR in one study of ‘excellent’ methodological quality;

++ OR -- Moderate evidence that property is adequate/inadequate. i.e. Consistent findings in multiple studies of ‘fair’ methodological quality OR in one study of ‘good’ methodological quality;

+ OR - limited evidence that property is adequate/inadequate. i.e. One study of ‘fair’ methodological quality;

± Conflicting evidence. i.e. Conflicting findings

? Unknown; i.e. Only studies of ‘poor’ methodological quality OR lack of relevant information reported.