An Evaluation of the Factor Structure of the Self-Assessed Wisdom Scale (SAWS) and the Creation of the SAWS-15 as a Short Measure for Personal Wisdom

Journal:	*International Psychogeriatrics*
Manuscript ID	IPG-08-20-476.R2
Manuscript Type:	Original Research Article
Date Submitted by the Author:	10-Dec-2020
Complete List of Authors:	Leeman, Tirlas; University of Southern Queensland Knight, Bob; University of Southern Queensland, Fein, Erich; University of Southern Queensland Winterbotham, Sonya; University of Southern Queensland Webster, Jeffrey; Langara College
Keywords:	Aging, Health aging, Psychogeriatrics
Abstract:	Objectives: Although wisdom is a desirable life span developmental goal, researchers have often lacked brief and reliable construct measures. We examined whether an abbreviated set of items could be empirically derived from the popular 40-item five factor Self-Assessed Wisdom Scale (SAWS). Design: Survey data from 709 respondents were randomly split into two and analyzed using confirmatory factor analysis (CFA). Setting: The survey was conducted online in Australia. Participants: The total sample consisted of 709 participants (Mage = 35.67 years; age range = 15–92 years) of whom 22% were male, and 78% female. Measurement: The study analyzed the 40 item Self-Assessed Wisdom Scale. Resultsw: Sample 1 showed the 40-item SAWS did not fit the data. Exploratory factor analysis (EFA) on Sample 2 offered an alternative model based in a 15-item, five factor solution with the latent variables Reminiscence/Reflection, Humor, Emotional Regulation, Experience, and Openness. This model, which replicates the factor structure of the original 40-item SAWS with a short form of 15 items, was then confirmed on Sample 1 using a CFA that produced acceptable fit and measurement invariance across age groups. Conclusions: We suggest the abbreviated SAWS-15 can be useful as a measure of individual differences in wisdom, and we highlight areas for future research.

An Evaluation of the Factor Structure of the Self-Assessed Wisdom Scale (SAWS) and the Creation of the SAWS-15 as a Short Measure for Personal Wisdom

Trilas M. Leeman, Bob G. Knight, Erich C. Fein, and Sonya Winterbotham

University of Southern Queensland, Toowoomba, Australia

Jeffrey Dean Webster

Langara College, Vancouver, Canada

Address correspondence to: Trilas M. Leeman, University of Southern Queensland, West Street, Darling Heights, Queensland, Australia, 4350. Email: Trilas.Leeman@usq.edu.au; Telephone: +61746312638; Bob.Knight@usq.edu.au; Erich.Fein@usq.edu.au; Sonya.Winterbotham@usq.edu.au

Biographical data

Trilas M. Leeman, BPharm, Psychologist, is a PhD Candidate at the University of Southern Queensland. After working extensively as a Pharmacist in the United Kingdom, Africa, Papua New Guinea, and Australia; her research now focuses on the relationship between wisdom, intelligence, age, and gender.

Bob G. Knight, PhD is Professor of Psychology and Counselling at the University of Southern Queensland and Honorary Professor of Psychology at the University of Queensland. Prior to moving to Australia, he was at the Andrus Gerontology Center at the University of Southern California.

Erich C. Fein, PhD, completed his doctorate in psychology at the Ohio State University with a focus on research methods and individual differences. His research focuses on motivation and performance, individual differences, leadership, and health outcomes. He supervises numerous PhD students.

Ms Sonya Winterbotham PhD is a lecturer in the School of Psychology and Counselling, and Research Assistant for the Institute of Resilient Regions, both at the University of Southern Queensland.

Jeffrey Dean Webster, PhD, works within a lifespan developmental perspective to investigate issues of reminiscence functions, time perspective, and wisdom. He has developed questionnaire assessments to measure all three constructs over the span of 35 years of teaching in the Psychology Department of Langara College.

Abstract

Objectives: Although wisdom is a desirable life span developmental goal, researchers have often lacked brief and reliable construct measures. We examined whether an abbreviated set of items could be empirically derived from the popular 40-item five factor Self-Assessed Wisdom Scale (SAWS).

Design: Survey data from 709 respondents were randomly split into two and analyzed using confirmatory factor analysis (CFA).

Setting: The survey was conducted online in Australia.

Participants: The total sample consisted of 709 participants (M_{age} = 35.67 years; age range = 15–92 years) of whom 22% were male, and 78% female.

Measurement: The study analyzed the 40 item Self-Assessed Wisdom Scale.

Results: Sample 1 showed the traditional five factor structure for the 40-item SAWS did not fit the data. Exploratory factor analysis (EFA) on Sample 2 offered an alternative model based in a 15-item, five factor solution with the latent variables Reminiscence/Reflection, Humor, Emotional Regulation, Experience, and Openness. This model, which replicates the factor structure of the original 40-item SAWS with a short form of 15 items, was then confirmed on Sample 1 using a CFA that produced acceptable fit and measurement invariance across age groups.

Conclusions: We suggest the abbreviated SAWS-15 can be useful as a measure of individual differences in wisdom, and we highlight areas for future research.

Keywords: SAWS, Wisdom, SAWS-15, measurement, reliability, age group, invariance

Introduction

Wisdom is a key aspect of positive mental health and is thought to develop over the course of the adult lifespan. It is multifaceted concept which has proved difficult to define and measure (Glück et al., 2013). A main measurement challenge concerns which elements of wisdom can be captured with self-report instruments (Webster, 2003, 2019). Wisdom researchers theorize the construct can be indirectly gauged by assessing agreed upon attributes of the wise, although these measures depend on some consensus of what the basic components of wisdom are (see Bangen, Meeks, & Jeste, 2013; Grossmann & Kung, 2019; Meeks & Jeste, 2009). With self-report measurement, measures generally fall into one of two broad categories: "Personal wisdom," wisdom gained from self-reflection and personal life experiences, and "general wisdom," which is concerned with problem solving applied to general life experiences (Glück et al., 2013). Most self-report measures measure personal wisdom (Glück et al., 2013).

Given that researchers often want to measure a wide range of psychological constructs, they are frequently hampered by lengthy measures (Thomas, Bangen, Ardelt, & Jeste, 2017). Shorter measures are easier and faster to administer and can be beneficial for researching vulnerable population groups who could fatigue easily due to medical conditions or groups with shorter attention spans (Thomas et al., 2017). In addition, longitudinal research typically involves large sets of variables so there is pressure to minimize number of items per scale.

Although there is scarcity of brief wisdom measures, some attempts have been made to generate such measures. For example, there is a brief version of the Three-Dimensional Wisdom Scale-12 (3D-WS-12) from Ardelt's (2003) 39-item 3D-WS measure (Thomas et al., 2017). There is also a brief wisdom screening scale by Glück and colleagues (2013) reflecting a broad conception of wisdom, drawing items from the 3D-WS, the Adult Self-

Transcendence Inventory (ASTI; Levenson, Jennings, Aldwin, & Shiraishi, 2005), and the Self-Assessed Wisdom Scale (SAWS; Webster, 2007). Webster (2019) notes that convergent validity between the 3D-WS and the SAWS is relatively weak, suggesting they may be measuring different facets of wisdom. The SAWS has stronger subscale reliability compared to other measures (Glück et al., 2013) and has a distinct theoretical background (described further below) and a brief English version of the SAWS could be beneficial in stimulating future wisdom research. Recently, Fung, Chow, and Cheung (2020) constructed a brief 9-item Chinese wisdom measure from SAWS items for studying wisdom in older adults, however, their measure did not preserve the five factor structure of the SAWS. Furthermore, since the scale is unidimensional, its utility might be limited as the measure cannot be used by other researchers to study different facets of the wisdom construct.

Development of the SAWS

Webster (2003) developed the SAWS after an extensive search of the literature, originally developing a 30-item tool which he later expanded into its current 40 items. Webster (2007) defined wisdom as, 'the c*ompetence* in*, intention* to, and *application* of *critical* life experiences to facilitate the *optimal development* of *self* and *others'* (p. 164; italics by original author). The SAWS included five components which integrate to typify characteristics of those who are wise, that is: Reminiscence/Reflection, Humor, Emotional Regulation, Critical Life Experiences, and Openness.

During the SAWS development, Webster (2003, 2007) used heterogeneous samples encompassing the adult lifespan including diverse ethnicities. The SAWS has reported some high-quality psychometric characteristics with very good reliabilities for some sub-scales. Furthermore, Glück et al. (2013), demonstrated convergent validity between the SAWS and performance measures such as the Berlin Wisdom Paradigm and with other self-report wisdom tools (Taylor, Bates, & Webster, 2011). Moreover, the SAWS predicts several

theoretically relevant wisdom constructs such as ego-integrity and meaning in life (Webster, 2010; Webster, Weststrate, Ferrari, Munroe, & Pierce, 2018) as well as wisdom nominees (Krafcik, 2015), and generally shows good to excellent psychometric properties with translated scales when used in cross-cultural research (e.g., Alquraan, Alshraideh, & Bsharah, 2010).

Finally, whilst self-report indicators have often been associated with social desirability, evidence from Taylor et al. (2011), and Thomas et al. (2019) indicated that the SAWS had non-significant association with social desirability. When Brienza, Kung, Santos, Bobocel, and Grossmann (2018) compared four self-report measures of wisdom, the SAWS displayed the least impression management bias (r =.22).

Thus, the SAWS is an appropriate base for the development of a short measure of wisdom. Smith, McCarthy, and Anderson (2000) indicated using an insufficiently validated parent scale often lead to error in the development of its abbreviated form. Therefore, we chose to use the SAWS as the foundation of a short measure of wisdom, and accordingly, the overall objective of this research is to re-examine the SAWS five dimensions for possible construction of a brief measure using a sample that is inclusive of adolescents and adults.

Measurement Model Issues

The analyses used to develop the five factor structure of the SAWS have been challenged in the literature. Webster (2007) conducted a principal components analysis (PCA) on the scale items setting the number of components to five a priori based on a search of the literature. The PCA findings supported the five factors with a good alpha reported for the total scale (α = .90). However, reliability estimates produced with PCA are inherently unstable (Flora & Flake, 2017; Tabachnick & Fidell, 2019).

We note that Webster (2007) replicated the PCA model in a CFA using the same sample. Nonetheless, a factor structure derived from PCA will almost always fit well in a

CFA using the same sample as the technique capitalizes on chance factors in the data (Flora & Flake, 2017). Of note is that in PCA the components are based on shared, unique, and error variances whereas in exploratory factor analysis (EFA) the unique and error variances are estimated and factored out and are not used to create the factors (Tabachnick & Fidell, 2019). PCA therefore, benefits from the error variance in the matrix of loadings to estimate components, which EFA eliminates. Also, during the CFA, Webster entered the SAWS subscales as manifest variables rather than as latent indicators.

In CFA the error terms in the model are specifically estimated. When factors (i.e., latent variables) are entered as manifest variables, the error unique to individual items is aggregated into a single error term, which may obscure problems in measurement related to error variance or that the items may not be functioning well together. Webster (2007, p. 175) noted his CFA methodology used factors as manifest variables rather than as latent indicators was not ideal, reminding us, 'The CFA must be interpreted cautiously given the number of parameters to number of participants ratio and some weak and non-supportive fit indices results (e.g., significant χ^2 value)". Indeed, Webster's CFA results of the SAWS prompted wisdom researchers Greene and Brown (2009) to suggest, "The instrument requires revision before being used as a measure of wisdom" (p. 292).

Potential Age Differences

Wisdom is generally thought to increase from adolescence through adulthood and on into older adulthood, although the trajectory is not consistent in the literature or across measures. The SAWS generally shows a weak positive relationship to age when samples only include adolescents to younger adults. This is in accord with studies by Pasupathi, Staudinger, and Baltes (2001) who found wisdom-related knowledge and judgement increased in a sample of adolescents and young adults aged 14–37 years, with no further increases after 25 years. However, the SAWS demonstrates a weak curvilinear relationship

with the apex at midlife when middle aged and older adults are included (e.g., Webster, Westerhof, & Bohlmeijer, 2014). This latter finding is consistent with that of other contemporary scholars (e.g., Ardelt, Pridgen, & Nutter-Pridgen, 2018; Sternberg, 2005; Thomas et al., 2017). The indication is that wisdom requires time, difficult life experiences, self-reflection, and learning from one's experiences for wisdom to actualize (Ardelt et al., 2018; Webster et al., 2014). Yet, despite older individuals' distinct advantage in greater life experiences, older age is often accompanied by increases in dogmatism and mental rigidity which are known to hinder wisdom development (Meacham, 1990). Furthermore, compared to older persons, midlife individuals, show greater openness to experiences giving them an advantage in wisdom (Webster et al., 2014).

Summary

In summary, a primary goal of our research was to reevaluate the five factor structure of the SAWS in order to examine whether an abbreviated set of items could be empirically derived. We started with a CFA instead of EFA because Webster's (2003, 2007) research suggests that the 40-item SAWS is a valid and reliable scale. Since our purpose was to generate a short form of the SAWS, an initial CFA implies that we are confirming a pre-established structure. However, a convergent EFA conducted with an independent sample would provide additional support for the existing structure and items that could be selected for a short version of the SAWS. In line with these objectives the following research questions were proposed:

Question 1: Will the five-dimensional factor structure of the full SAWS replicate using a CFA?

Question 2: If the answer to Question 1 is that the five factor structure does not replicate, will an EFA of the SAWS items in a separate sample produce a five-factor solution?

Question 3: Will a brief form of the SAWS will show acceptable fit and measurement invariance using a CFA and significant mean wisdom differences across age groups from adolescence into the older adult years? Measurement invariance by age is key to accurate understanding of the trajectory of wisdom across the adult life span.

Method

Study Design

This study used a cross-sectional survey design in an Australian setting. Items measuring wisdom were self-assessed. All the measurements were incorporated into an electronic survey format.

Participants

The total sample consisted of 709 participants (M_{age} = 35.67 years; age range = 15–92 years) of whom 22% were male, and 78% female. According to Erikson's (1959) psychosocial stages, the sample included: Adolescents (15-18 years, n = 81), young adults (19-40 years, n = 396), middle age (41-65 years, n = 190), and older adults (66-92 years, n = 42). Of the respondents 80.10% were White Australians, 3.40% Aboriginal and Pacific Islanders, 10.40% reported other ethnicities, while 6.10% did not specify their ethnicity. English as a first language was spoken by 93.80% of the sample. Participants were well educated (M = 14.79 years; range = 2–36 years), with good self-reported health (M = 7.53; range = 1–10). Data were randomly split into N = 356, and N = 353 subsamples for the CFA, and EFA. There were no significant differences in these demographic variables between the two groups (see Appendix A).

Measures

Background variables. Participants were asked their age, gender, educational level, ethnicity, occupation, the country in which they were born, whether they were retired or not, and the main language spoken at home. They were also asked to rate their health on a scale

of 1-10 (1 = poor, 10 = excellent). Other measures were administered, but the SAWS was the relevant measure for this study.

Wisdom. Wisdom was assessed with Webster's (2007) 40-item SAWS, with its five dimensions of: Reminiscence/Reflection, Humor, Emotional Regulation, Experience, and Openness. Each subscale is made up of eight positively phrased statements and responses are rated on a 6-point Likert-type scale (1 = *strongly disagree* to 6 = *strongly agree*). Raw scores are summed up to produce a total wisdom score. The SAWS is unrelated to age, or gender (Moberg, 2008; Webster, 2007). PCA supported a five factor structure (Taylor et al., 2011; Webster, 2007). Webster reported high reliability for the total scale ($\alpha = .90$), with a two-week test-retest reliability of ($\alpha = .84$).

Procedure

Participants were recruited via newspaper advertisements, community groups, and word of mouth at the School of Psychology courses at a demographically diverse Australian university. Data were collected according to the university's Human Research Ethics Committee protocol for online surveys. Student participants received credits depending on rules of the class in which they were enrolled. Non-student respondents were required to be 18 years or older, but psychology university students could be as young as 15. Participation was voluntary and anonymous. Participants were informed that the survey would take about 30-45 minutes to complete and could be accessed individually at home via the provided link. Informed consent was assumed when participants navigated to the survey after reading the information page which contained the purpose of the study including the rights of participants.

Statistical Analyses

IBM SPSS version 26 software program was utilized for data screening and the EFA. Analysis of Moment Structures (AMOS) version 26 software, using the robust Maximum

Likelihood (ML) parameter estimation method, was employed for the CFA. This method involves performing 1) Bootstrapping, 2) Bias Corrected Confidence Intervals, and 3) Bootstrapped ML. We note that in large samples, variables with statistically significant skewness often do not deviate enough from normality to make a meaningful difference in the analysis (Tabachnick & Fidell, 2019). Also, when variables are not expected to be normally distributed, Hoyle (1995) recommended using the ML estimation method when conducting CFAs, as ML has been found to be robust to violations of multivariate normality.

In accord with Byrne (2010) identification was achieved through specification of our model to account for a unique solution to the model parameters. Specifically, we fixed the initial loading on each latent variable and allowed the other parameters to be estimated. We then calculated the information available from the observed variance-covariance matrix and ensured this information was sufficient to estimate the unknown parameters, thus ensuring we had an over identified model. Our a priori calculations were then compared to the output, which confirmed the model was sufficiently identified when our CFA was conducted.

Hu and Bentler's (1999) criteria for acceptable model fit were followed. We reported several fit indices. The Comparative Fit Index (CFI), the Goodness of Fit Index (GFI), and the Tucker-Lewis Index (TLI), where values > .95 indicate well-fitting models. The Standardized Root Mean-square Residual (SRMR; < .08), and the Root Mean Square Error of Approximation (RMSEA; < .06 indicates good fit, .06–.08 satisfactory fit, and >.10 poor models). The χ^2 was reported, but it is not considered an adequate estimate of model fit, as it is sensitive to large samples (Kahn, 2006). The χ^2 / df ratio which considers sample size is reported; with values of < 3 considered acceptable (Byrne, 2016).

Data Screening

Data screening revealed one out of range score which was corrected manually. Analysis of missing data indicated Little's (1988) Missing Completely At Random (MCAR) test was not significant ($p = .12$) supporting that missing data patterns were random. The missing values represented 2.54% of the data. Mean replacement was used to impute missing values, with item level imputation in the SAWS, and the original sample of 709 was preserved. We created two subsamples of $N = 356$ for Sample 1 and $N = 353$ for Sample 2 based on random assignment of cases to samples. Both samples matched the a priori sample size estimate of 300 cases suggested as common "rules of thumb" for CFA and EFA procedures (Comfrey & Lee, 1992; Worthington & Whittaker, 2006). The subsamples of $N = 356$ for Sample 1 and $N = 353$ for Sample 2 were therefore deemed adequate.

Results

CFA findings. A CFA was performed on the first subsample ($N = 356$) to attempt replication of the five factor structure. The factors were allowed to covary but we did not use covariation of the error terms. The model fit indices were $\chi^2 (730) = 2133.83$, $p < .01$, $\chi^2/df = 2.92$, CFI = .72, GFI = .74, TLI = .70, SRMR = .08, RMSEA = .07, 90% CI [0.07, 0.08]. Although the RMSEA was marginal, the hypothesized model was a poor fit given the very low values of CFI, GFI, and TLI. To answer Question 2 an EFA was conducted on Sample 2 ($N = 353$) to see what factor structure emerged empirically. We changed samples to reduce the potential influence of unique error variance within the first sample when extracting the EFA factors. Thus, we used the second sample for an EFA comparison against the CFA results from the first sample.

EFA findings. Different methods can be used to specify the number of factors to retain during factor analysis, although the five factor model developed by Webster was not a good fit with the full 40 items, we wanted to maintain that conceptual structure with a smaller

set of items if possible. Also, the scree plot suggested five factors which converged with Webster's (2003, 2007) theoretical concepts for the SAWS factors.

The EFA was performed utilizing the Maximum Likelihood extraction method with Promax rotation (permitting correlated factors) on the 40-item SAWS. In accord with Webster's factor structure of the SAWS (2003) we fixed the number of factors to be extracted at five. Correlations between factors all exceed .32 (see Appendix B) which suggests there is sufficient overlap in variance among the factors to support an oblique rather than orthogonal rotation (Tabachnick & Fidell, 2019). Since the correlations are well below .70, the indication is that each facet is assessing a different aspect of the wisdom construct.

The Kaiser-Meyer-Olkin (KMO) value was .87, and all KMO values for individual items were greater than .74, exceeding the recommended value of .50 (Field, 2017). Bartlett's test of sphericity reached statistical significance, supporting the factorability of the correlation matrix. Initial analysis was run to obtain eigenvalues for each factor in the dataset. Howard's (2016) rule that variables should: (a) 'load onto their primary factor above 0.40, (b) load onto alternate factors below 0.30, and (c) demonstrate a difference of 0.20 between their primary and alternative loading' (p. 55) was used. The five factors in combination explained 45.10% of the common variance.

Factor 1 was composed of eight items (3, 8, 13, 18, 23, 28, 33, 38) from the SAWS Reminiscence/Reflection factor. Factor 2 was composed of seven items (4, 9, 14, 24, 29, 34, 39) from the Humor factor. Factor 3 was composed of five items (12, 17, 22, 27, 37) from the Emotional Regulation factor. Factor 4 included five items from the Experience factor (1, 6, 16, 21, 26), while Factor 5 included four items from the Openness factor (5, 15, 25, 35), with one item (36) from the Experience factor. Table 1 shows the factor loadings after rotation.

Insert Table 1 about here

There is no precise rule for the retention of items within an EFA (Meyers, Gamst, & Guarino, 2006). In general items should load as highly as possible with no significant loadings on multiple factors. However, the reliability of items within a factor structure seems to be significantly compromised when loadings are below .30, with items loading around .40 being marginal (Gorsuch, 1983; Stevens, 2002). Loadings above .50 are desirable and are considered good for the replicability and utility of scales (Comfrey & Lee, 1992; Meyers et al., 2006). Based on this guideline, we noted that there were at least three items loading above .45 for all extracted factors. We observed that Webster (2003, 2007) had an equal number of items for each factor. In addition, DeVellis (2017) posited that questionnaires should possess the most parsimonious, that is, the simplest structure.

Accordingly, to retain Webster's (2007) scale structure of an equal number of items for each subscale, to not privilege any subscale in the total score by having unequal number of items per subscale and to maximize the internal consistency and parsimony, we selected the three highest loading items on each of the five factors to comprise a 15-item version of the SAWS. Thus, we had equal number of items per factor, with a loading of at least .50 for each item.

The Cronbach alphas for this 15-item five factor SAWS were: Reminiscence/Reflection (α = .74), Humor (α = .72), Emotional Regulation (α = .85), Experience (α = .73), and Openness (α = .56), and α = .80 for the total SAWS-15 scale. Prior to combining the scores into a single wisdom score we evaluated a higher order construct of wisdom (see Appendix C). The fit indices for this higher order measurement model were good (χ^2 (85) = 325.92, $p < .001$, χ^2/df = 3.83, CFI = .92, GFI = .94, TLI = .91, SRMR = .05, RMSEA = .06, 90% CI [0.06, 0.07]).

SAWS-15 five factor CFA. A CFA was performed on Sample 1 ($N = 356$) to assess fit for the alternative five factor 15-item SAWS model derived from Sample 2. The items included in the CFA served as manifest indicators of Reminiscence/Reflection (items 8, 13, 23), Humor (items 4, 14, 24), Emotional Regulation (items 12, 17, 27), Experience (items 1, 6, 26), and Openness (items 5, 15, 35). We used the variance-covariance matrices where the factors covaried, but we did not covary the error terms. Results showed a good fit to the data in this independent sample χ^2 (80) = 217.05, $p < .001$, χ^2/df =2.71, CFI = .91, GFI = .93, TLI = .89, SRMR = .06, RMSEA = .07, 90% CI [0.06, 0.08]. The model gave a good fit for the data and explained 67.76% of the common variance.

Multigroup measurement invariance. The hypothesis that the SAWS-15 is measurement invariant was tested. Measurement invariance assesses whether different age groups interpret a measure in a conceptually similar manner; a necessary step before assessing mean differences (Byrne, 2016). The sample was subdivided using Erikson's (1959) psychosocial stages, as each stage accomplishes a different developmental task. The groups were: Adolescents (15-18 years, $n = 81$), young adults (19-40 years, $n = 396$), middle age (41-65 years, $n = 190$), and older adults (66-92 years, $n = 42$). We compared the fit of progressively restrictive models. Employing CFA modeling we focused our tests of invariance across age groups as recommended by Byrne (2016) with respect to (1) factor loadings, (2) factor loadings and intercepts, and (3) factor loadings, intercepts, and error variances.

Model 1 tested the overall fit of the baseline unconstrained (configural) model across the four age groups. The fit indices χ^2 (320) = 651.33, $p < .001$, CFI = .90, SRMR = .07, RMSEA = .04, 90% CI [0.03, 0.04] supported configural invariance, indicating suitability for metric invariance testing (Byrne, 2016). Model 2 tested metric invariance by constraining the factor loadings across age groups to be equal. Findings indicated acceptable data-model fit χ^2

(350) = 681.19, p <.001, CFI = .90, SRMR = .08, RMSEA = .04, 90% CI [0.03, 0.04]. Due to the sensitivity of χ^2 to sample size and non-normality, Cheung and Rensvold (2002) recommend a ΔCFI $\leq$ .01 between two nested models would support measurement invariance. The non-significant $\Delta\chi^2$ (30) = 29.86, p = .473, and the ΔCFI < .01 both supported metric equivalence at path coefficients level. Model 3 tested scalar invariance by constraining item intercepts to be equivalent across age groups. Results indicated acceptable model fit χ^2 (395) = 734.39, p <.001, CFI = .89, SRMR = .09, RMSEA = .04, 90% CI [0.03, 0.04]. The $\Delta\chi^2$ (75) = 83.06, p = .245 and the ΔCFI = < .01 between Model 2 and 3 indicated differences in factor variances and covariances are not due to age group-based differences.

Mean SAWS-15 differences. A one-way between groups analysis of variance (ANOVA) examined the impact of the four age groups on wisdom, measured by SAWS-15. There was a statistically significant difference at the p < .05 level in wisdom scores for the four age groups, F(3, 705) = 11.25, p < .001 with medium effect size, η^2 = .05 (Cohen, 1988). Post-hoc group comparisons using Hochberg's GT2 (homogeneity of variance with unequal sample sizes; Field, 2017) indicated the mean score for adolescents (M = 65.42, SD = 8.98) was significantly lower than all other groups. Young adults (M = 69.12, SD = 9.19) differed significantly from the middle age (M = 72.04, SD = 8.53), but both did not differ significantly from the older adults (M = 71.09, SD = 9.21).

Discussion

This study tested Webster's (2007) SAWS five factor structure of Reminiscence/Reflection, Humor, Emotional Regulation, Experience, and Openness. For a multidimensional construct like wisdom, subscales need to exhibit well defined, replicable factor structures that agrees with the underlying theory (Byrne, 2016).

The question of whether the 40-item five factor SAWS is a reliable measure of wisdom that would replicate in the current research through a CFA was answered in the

negative. A possible explanation for the current study's lack of concordance with Webster's (2007) CFA findings is the different analytic strategies used. In the current study, the SAWS five dimensional structure with its 40 items was submitted to a CFA. Webster used the SAWS five factors or latent variables of Reminiscence/Reflection, Humor, Emotional Regulation, Experience, and Openness as manifest variables rather than as latent variables. Although Webster argued that such a methodology was justified to simplify model parameters, he conceded, "This is a suboptimal strategy" (p. 171).

The earlier cross validation of the PCA findings with a CFA, using the same data (Webster, 2007), is also an issue in testing the validity of a factor structure. A factor structure derived from a PCA or EFA will almost always fit very well in a CFA when using the same sample because this procedure capitalizes on any chance factors present in the dataset (Flora & Flake, 2017). Having a large sample, we were able to split the data randomly into two independent samples so that the EFA could be conducted on an independent sample, and then validated with replication in the other sample using CFA. To our knowledge, this is the first report of such replication using SAWS items.

Exploratory factor analysis. Since the SAWS five factor structure was not supported by the CFA, we used EFA to find where the 40 items of the SAWS were fitting. Results showed that the SAWS EFA did not map onto the theoretical model given eight of the 40 items (20%) cross-loaded on other factors with loadings ≥ .30, and 11 items (27.50%) did not load on any factor. However, analysis also found that a subset of items from the SAWS using the original five factor structure fit the data and were a good fit in the replication sample.

The question of whether the derived brief scale (SAWS-15) would show measurement invariance and significant mean wisdom differences across age groups was answered in the affirmative. The relationship between wisdom and age showed that wisdom was greater in successive age groups from adolescents to the middle aged with older adults between and not

different from young and middle aged adults. Although Moberg (2008) and Webster (2007) found no relationship between age and wisdom with the 40-item SAWS more recently, other scholars (e.g., Ardelt et al., 2018; Thomas et al., 2017; Webster et al. 2014), reported a curvilinear relationship between wisdom and age with the peak at midlife. Despite most societies associating wisdom with older age (Assmann, 1994), the wisdom-age relationship is complex. This could be because growing old does not automatically confer individuals the ability to self-reflect and integrate one's life experiences in a wisdom fostering manner; qualities which wisdom scholars and lay persons agree are necessary for wisdom to actualize (Glück & Bluck, 2013).

Limitations

While our sample has the advantage of being large enough to support these analyses, the respondents were mostly female, less than 45 years of age, and the large majority were White Australians. Generalizability is thus potentially limited. Nevertheless, the use of a sample with a wide age range has the potential to help in explicating the wisdom-age relationship. Future research with other populations and with population-based sampling rather than convenience samples would strengthen the case of the usability of this brief version of the SAWS.

Self-report measures can limit validity due to social desirability responses by respondents (Shaughnessy, Zechmeister, & Zechmeister, 2012). An inclusion of a social desirability measure in future research would clarify whether respondents are giving socially desirable responses rather than accurately reporting their attributes, thus avoiding a common pitfall of self-report measures. This is important considering that Taylor et al. (2011) and Thomas et al. (2019) reported that the SAWS measure is free of biased responding, while other researchers (e.g., Brienza et al., 2018) have reported otherwise.

Finally, the internal consistency of the Openness subscale at α = .56 could be viewed as a limitation, although a coefficient in this range is acceptable during measurement development and when using scales with a low number of items (DeVellis, 2017). We suggest that because reliability estimates using the alpha coefficient is highly biased against scales with three items or less (Carmines & Zeller, 1979), in future research other types of reliability estimates should be used, such as retest reliability (Carmines & Zeller, 1979; DeVellis, 2017).

Conclusion

The present research explored the validation of the five factor structure of the SAWS and confirmed the five factor structure for a shorter version of the scale. To the best of our knowledge, this is the first time a brief five factor structure has been advanced and replicated with an independent sample.

A key finding was that a brief version of the SAWS with a five factor structure was a good fit to the data and was replicable in an independent sample and was measurement invariant with respect to age and so can be used to explore the trajectory of wisdom development across adulthood and into older ages. In respect to future research, we suggest that range content and criterion-related validation research be conducted as well as further reliability estimates using a range of methods. While clearly more research on this revised version is needed, this brief version can be seen as a step toward making one of the most commonly used measures easier to use in large scale research programs and in clinical and educational settings.

Conflict of Interest: None

Description of Author Roles:

T. Leeman analysis of data analyses and writing.

Bob Knight Study design and writing

Erich Fein supervised data analyses and contributed to writing

Sonya Winterbotham directed data collection and data cleaning

Jeffry Webster consulted on data analysis and writing.

References

Alquraan, M., Alshraideh, M., & Bsharah, M. (2010). Psychometric properties and differential item function (DIF) analyses of Jordanian version of self-assessed wisdom scale (SAWS-Jo*). International Journal of Applied Educational Studies, 9*(1), 52–66.

Ardelt, M. (2003). Empirical assessment of a three-dimensional wisdom scale. *Research on Aging, 25,* 275–324. doi:10.1177/0164027503025003004

Ardelt, M., Pridgen, S., & Nutter-Pridgen, K. L. (2018). The relation between age and three-dimensional wisdom: Variations by wisdom dimensions and education. *The Journals of Gerontology: Series B*, *73*, 1339–1349. doi:10.1093/geronb/gbx182

Assmann, A. (1994). Wholesome knowledge: Concepts of wisdom in a historical and cross-cultural perspective. In D. L. Featherman, R. M. Lerner, & M. Perlmuter (Eds.), *Lifespan development and behaviour (*pp. 187-224). Hillsdale, NJ: Erlbaum.

Bangen, K. J., Meeks, T. W., & Jeste, D. V. (2013). Defining and assessing wisdom: A review of the literature. *American Journal of Geriatric Psychiatry, 21,* 1254–1266. doi:10.1016/j.jagp.2012.11.020

Brienza, J. P., Kung, F. Y. H., Santos, H. C., Bobocel, D. R., & Grossmann, I. (2018). Wisdom, bias, and balance: Toward a process-sensitive measurement of wisdom-related cognition. *Journal of Personality and Social Psychology, 115*(6), 1093–1126. doi:10.1037/pspp0000171

Byrne, B. M. (2016). *Structural equation modelling with Amos: Basic concepts, applications, and programming* (3rd Ed.). New York, NY: Routledge.

Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. London: Sage.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255. doi:10.1207/S15328007SEM0902_5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Comfrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage.

Erikson, E. H. (1959). *Identity and the life cycle*. New York, NY: Norton.

Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th ed.). London: SAGE.

Flora, D. B., & Flake, J. K. (2017). The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Canadian Journal of Behavioural Science, 49,* 78–88. doi:10.1037/cbs0000069

Fung, S., Chow, E. O., & Cheung, C. (2020). Development and validation of a brief self-assessed wisdom scale. *BMC Geriatrics, 20*(54). Retrieved from: https://bmcgeriatr.biomedcentral.com/articles/10.1186/s12877-020-1456-9

Glück, J., & Bluck, S. (2013). The MORE life experience model: A theory of the development of personal wisdom. In M. Ferrari, & N. M. Weststrate (Eds.), *The scientific study of personal wisdom* (pp. 75–97). Dordrecht, Netherlands: Springer.

Glück, J., König, S., Naschenweng, K., Redzanowski, U., Dorner, L., Straßer, I., & Wiedermann, W. (2013). How to measure wisdom: Content, reliability, and validity of five measures. *Frontiers in Psychology, 4*, 1–13. doi:10.3389/fpsyg.2013.00405

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, MI: Erlbaum.

Greene, J. A., & Brown, S. C. (2009). The Wisdom Development Scale: Further validity investigations. *International Journal of Aging and Human Development, 68*, 289–320. doi:10.2190/AG.68.4.b

Grossmann, I., & Kung, F. Y. H. (2019). Wisdom and culture. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (2nd ed., pp. 343–364). New York, NY: Guilford Press.

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human–Computer Interaction*, *32*, 51–62. doi:10.1080/10447318.2015.1087664

Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 1–15). Thousand Oaks, CA: Sage.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1)*,* 1–55. doi:10.1080/10705519909540118

Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist, 34,* 684–718. doi:10.1177/0011000006286347

Krafcik, D. (2015) Words from the wise: Exploring the lives, qualities, and opinions of exemplars of the wise. *Integral Review, 11,* 7-35.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198–1202. doi:10.2307/220157

Meacham, J. (1990). The loss of wisdom. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 181–211). Cambridge, UK: Cambridge University Press.

Meeks, T. W., & Jeste, D. V. (2009). Neurobiology of wisdom: A literature overview. *Archives of General Psychiatry, 66*, 355–365. doi.10.1001/archgenpsychiatry.2009.8

Meyers, L.S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: Sage.

Moberg, D. (2008). Mentoring and practical wisdom: Are mentors wiser or just more politically skilled? *Journal of Business Ethics, 83*, 835–843. doi:10.1007/s10551-008-9668-5

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers, 32,* 396-402. doi:10.3758/BF03200807

Pasupathi, M., Staudinger, U. M., & Baltes, P. B. (2001). Seeds of wisdom: Adolescents' knowledge and judgment about difficult life problems. *Developmental Psychology, 37,* 351–361. doi:10.1037/0012-1649.37.3.351

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2012). *Research methods in psychology* (9th. ed.). New York, NY: McGraw–Hill.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*, 102–111. doi:10.1037/1040-3590.12.1.102

Sternberg, R. J. (2005). Older but not wiser? The relationship between age and wisdom. *Ageing International, 30*(1), 5–26. doi:10.100/BF02681005

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Harlow, Essex: England.

Taylor, M., Bates, G., & Webster, J. D. (2011). Comparing the psychometric properties of two measures of wisdom: Predicting forgiveness and psychological well-being with the Self-Assessed Wisdom Scale (SAWS) and the Three-Dimensional Wisdom Scale

(3D-WS). *Experimental Aging Research, 37*, 129–141. doi:10.1080/0361073X.2011.554508

Thomas, M. L., Bangen, K. J., Ardelt, M., & Jeste, D. V. (2017). Development of a 12-item abbreviated Three-Dimensional Wisdom Scale (3D–WS–12): Item Selection and Psychometric Properties. *Assessment, 24,* 71–82. doi:10.1177/1073191115595714

Thomas, M. L., Bangen, K. J., Palmer, B. W., Martin, A. S., Avanzino, J. A., Depp, C. A., … Jeste, D. V. (2019). A new scale for assessing wisdom based on common domains and a neurobiological model: The San Diego Wisdom Scale (SD–WISE). *Journal of Psychiatric Research, 108,* 40–47. doi:10.1016/jpsychires.2017.09.005

Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development, 10,* 13–22. doi:10.1023/A:1020782619051

Webster, J. D. (2007). Measuring the character strength of wisdom. *The International Journal of Aging & Human Development, 65,* 163–183. doi:10.2190/AG.65.2.d

Webster, J. D. (2010). Wisdom and positive psychosocial values in young adulthood. *Journal of Adult Development, 17*(2), 70–80. doi:10.1007/s10804-009-9081-z

Webster, J. D. (2019). Self-report wisdom measures: Strengths, limitations, and future directions. In R. J. Sternberg, & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 297 - 320). Cambridge, UK: Cambridge University Press.

Webster, J. D., Westerhof, G. J., & Bohlmeijer, E. T. (2014). Wisdom and mental health across the lifespan. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 69,* 209–218. doi:10.1093/geronb/gbs121

Webster, J. D., Weststrate, N. M., Ferrari, M., Munroe, M., & Pierce, T. W. (2018). Wisdom and meaning in emerging adulthood. *Emerging Adulthood, 6*(2), 118–136. doi:10.1177/2167696817707662

Worthington, R., & Whittaker, T. (2006). Scale development research: A content analysis and recommendations for best practices. *Counseling Psychologist*, *34*, 806–838. doi:10.1177/0011000006288127

Table 1: EFA for 40 SAWS items

Items	Factors and Loadings				
	Reminiscence/ Reflection	Humor	Emotional Regulation	Experience	Openness
Critical Life Experiences					
1. I have overcome many painful events in my life.	-.006	.021	-.039	**.720**	-.127
6. I have had to make important life decisions.	-.084	-.083	.150	**.701**	.052
11. I have dealt with a great many different kinds of people during my lifetime.	-.023	.095	.122	.330	.157
16. I have experienced many moral dilemmas.	-.020	.008	-.057	**.524**	.200
21. I have seen much of the of the negative side of life (e.g., dishonesty, hypocrisy).	.170	-.016	-.041	**.519**	-.035
26. I. have lived through many difficult life transitions.	.008	-.053	-.010	**.869**	-.048
31. I've personally discovered that "you can't always tell a book from its cover".	.042	.038	.053	.073	.108
36. I've learned valuable life lessons from others.	.087	.090	-.109	.102	**.486**
Emotional Regulation					
2. It is easy for me to adjust my emotions to the situation at hand.	-.154	.315	.312	-.037	.096
7. Emotions do not overwhelm me when I make personal decision.	-.335	.264	.179	-.067	.180
12. I am "tuned" in to my own emotions.	.017	.047	**.742**	.074	-.135
17. I am very good at reading my emotional states.	-.078	-.092	**.902**	.013	.017

Items	Factors and Loadings				
	Reminiscence/ Reflection	Humor	Emotional Regulation	Experience	Openness
22. I can freely express my emotions without feeling like I might lose control.	.088	.290	**.416**	-.109	-.077
27. I am good at identifying subtle emotions within myself.	.055	-.153	**.857**	-.009	-.019
32. I can regulate my emotions when the situation calls for it.	-.144	.261	.299	-.034	.133
37. It seems I have a talent for reading other people's emotions	.148	-.064	**.566**	.015	.069
Reminiscence/Reflection					
3. I often think about connections between my past and present.	**.596**	.016	-.122	.110	-.064
8. I often think about my personal past.	**.672**	.006	-.168	.109	-.098
13. I reminisce quite frequently.	**.717**	-.037	.052	-.131	-.135
18. Reviewing my past helps me gain perspective on current concerns.	**.481**	.103	.160	.073	.104
23. I often recall earlier times in my life to see how I've changed since then.	**.685**	-.095	.181	-.039	.037
28. Recalling my earlier days helps me gain insight into important life matters.	**.599**	.027	.034	.120	.158
33. I often find memories of my past can be important coping resources.	**.638**	.116	-.019	-.112	.168

Items	Factors and Loadings				
	Reminiscence/ Reflection	Humor	Emotional Regulation	Experience	Openness
38. Reliving past accomplishments in memory increases my confidence for today.	**.511**	.106	.144	-.065	.042
Openness					
5. I like to read books which challenge me to think differently about issues.	-.058	-.035	.050	.112	**.465**
10. I enjoy listening to a variety of musical styles besides my favorite kind.	.117	.142	-.067	-.011	.226
15. I enjoy sampling a wide variety of different ethnic foods.	-.126	.029	.004	.042	**.576**
20. I often look for new things to try.	-.038	.194	.062	.109	.385
25. Controversial works of art play an important and valuable role in society.	.091	.111	-.023	-.058	**.442**
30. I like being around persons whose views are strongly different from mine.	.031	.224	-.201	-.029	.227
35. I'm very curious about other religious and/or philosophical belief systems.	.061	-.212	.022	-.153	**.797**
40. I've often wondered about life and what lies beyond.	.302	-.074	-.002	.029	.156
Humor					
4. I can chuckle at personal embarrassments.	-.044	**.733**	.004	.059	-.174

Items	Factors and Loadings				
	Reminiscence/ Reflection	Humor	Emotional Regulation	Experience	Openness
9. There can be amusing elements even in very difficult life situations.	-.016	**.499**	-.100	-.089	.139
14. I try and find humorous side when coping with a major life transition.	-.039	**.728**	-.011	.090	-.071
19. I am easily aroused to laughter.	.181	.338	.115	-.024	-.018
24. At this point in my life, I find it easy to laugh at my mistakes.	.058	**.707**	.123	.013	-.149
29. I often use humor to put others at ease.	.090	**.545**	-.041	-.028	-.003
34. Now I find that I can really appreciate life's little ironies.	.036	**.419**	-.002	.157	.154
39. I can make fun of myself to comfort others.	.021	**.615**	-.134	-.128	.083

Note: Factor loadings >.40 are in boldface.