Community Dental Health (2016) **33,** 274–280
Received 30 September 2015; Accepted 1 February 2016

doi:10.1922/CDH_3868Flynn07

Psychometric properties of the English version of the Oral Health Literacy Adults Questionnaire - OHL-AQ

P.M. Flynn[1], M.T. John[2], A. Naik[2], N. Kohli[3], J.J. VanWormer[4] and K.Self[1]

[1]Department of Primary Dental Care, University of Minnesota School of Dentistry, USA; [2]Department of Diagnostic and Biological Sciences, University of Minnesota School of Dentistry, USA; [3]Department of Educational Psychology, College of Human Development, University of Minnesota, USA; [4]Center for Clinical Epidemiology & Population Health, Marshfield Clinic Research Foundation, Marshfield Clinic, Wisconsin, USA

Objective: To test the psychometric properties of the Oral Health Literacy Adult Questionnaire (OHL-AQ) in English. The OHL-AQ was designed to test functional oral health literacy in general populations and was initially validated in Iran. ***Methods:*** The instrument was administered to 405 adult subjects (mean age 45 (SD 16) years and 67% female) attending the 2014 Minnesota State Fair. The OHL-AQ is composed of 17 items measuring four conceptual dimensions: reading comprehension, numeracy, listening, and decision-making. Participants selected the best answer for written or verbally administered items and entered answers on an electronic tablet. Item responses for each individual were combined into a summary score (range 0–17) with higher scores indicating better oral health literacy. Score dimensionality, reliability, and validity were investigated. ***Results:*** For dimensionality, both exploratory factor analysis and a parallel analysis yielded evidence for scale unidimensionality. Reliability was sufficient indicated by a Cronbach's alpha >0.74. Validity of scores was supported by "small" and "medium" effect sizes for construct validity. "Small" effect sizes were observed for global oral health self-report, OHIP-5 scores, treatment urgency, and having a regular dentist. "Medium" effect sizes were seen for presence of dentures, number of natural teeth present, and educational level. ***Conclusions:*** Dimensionality, reliability and validity of the English version of the OHL-AQ in a general adult English-speaking population is supported, providing sufficient psychometric properties in an important target population of the instrument.

Key words: *health literacy, psychometrics, factor analysis, functional oral health literacy, instrument validation, OHL-AQ*

Introduction

Oral health literacy (OHL) has been identified as a domain of health literacy and a determinant of health (NIDCR, 2005). Defined as "*the degree to which individuals have the capacity to obtain, process, and understand basic oral health information and services needed to make appropriate health decisions and act on them*," measuring OHL is seen as a necessary initial step toward improving oral health and reducing oral health disparities (ADA, 2009). Multiple instruments have been developed to measure OHL with the most widely used instruments quantifying reading and numeracy skills (Dickson-Swift *et al*., 2014). While a deficit of these skills is associated with poorer general and oral health, contemporary researchers have identified additional domains as important attributes of functional OHL. These include the ability to communicate with health care providers and decision-making skills (Dickson-Swift *et al*., 2014). One instrument designed to assess the functional domains of OHL is the Oral Health Literacy Adult Questionnaire (OHL-AQ). Specifically, the OHL-AQ includes items to measure reading comprehension, numeracy, listening (communication skills), and decision-making (Naghibi Sistani *et al*., 2014). These dimensions are important because a broad range of skills are needed to achieve high levels of OHL.

While other OHL instruments were developed and tested in research or clinical settings, the OHL-AQ was originally developed to measure OHL in a public health setting among the general population. Survey administration settings using the same or similar behavioral items have generated different prevalence estimates. In addition to the data collection setting, differences are also attributed to different target populations, data collection methodology, type of consent, and timing of data collection (Kann *et al*., 2002). Participant attention can be more challenging in a general population as compared to clinical research settings where more time may be available for responding (Rhodes *et al*., 2003). Of particular significance to oral health literacy assessment, patients in clinical settings have already shown their ability to navigate the health care system, which may indicate a sample not representative of the wider population. Because successful public health interventions require that baseline measures be gathered in similar settings, it is important to use instruments that are well suited to the environment and population. The initial OHL-AQ was designed in Iran and administered in the Persian language with promising initial psychometric properties (Naghibi Sistani *et al*., 2014). Though validation findings were published in English, psychometric properties of the OHL-AQ have not yet been validated in an English-speaking population. The purpose of this study was to analyze the psychometric properties of OHL-AQ questionnaire in an English-speaking, general population of adults.

Correspondence to: Priscilla M. Flynn, 9-372A Moos Tower, 515 Delaware Street SE, Minneapolis, MN 55455, USA.
Email: flynn.priscilla@umn.edu

Methods

The original Persian language version of the instrument consisted of 21 questions based on an item pool extracted from existing international instruments (Naghibi Sistani *et al*., 2014). Following content and face validity procedures, the instrument was reduced to 17 items. Subsequent testing for internal consistency and stability of the questionnaire yielded satisfactory properties. Validity was established by known-groups comparisons, and the instrument discriminated well among subgroups by levels of education and tooth-brushing frequency (Naghibi Sistani *et al*., 2014).

The final instrument used in the current study was translated into English by the developing authors (Naghibi Sisanti *et al*., 2013; 2014), but was not back-translated nor piloted in an English-speaking population prior to publication (personal communication, Dr M.M. Naghibi Sistanti, 31.12.2015). Slight language modifications were made for the American English-speaking sample. Questions are grouped into four sections: reading comprehension, numeracy, listening and decision-making. Reading comprehension (reading and knowledge skills) consists of three questions with six missing words or phrases. Respondents chose from a list of five multiple choice responses to complete each sentence. Numeracy (reading, writing and calculation skills) was assessed by reading two sample prescriptions, each followed by answering two questions: one free-text and one multiple choice. Listening (listening, reading, writing, calculation, and communication skills) was assessed by a study coordinator reading aloud sample post-extraction instructions, followed by one free-text response and one multiple choice question. Appropriate decision-making (reading, comprehension, and decision making skills) consisted of five multiple choice questions. Correct answers were scored, 1, and incorrect answers, 0, with a summed OHL-AQ score potentially ranging from 0 to 17 points. Based on the original authors' criteria, OHL-AQ scores were categorized into one of three categories: inadequate OHL (0-9), marginal (10-11), and adequate (12-17) (Naghibi Sistani *et al*., 2014).

Self-reported oral health information, clinical findings, and saliva samples were collected from a convenience sample of adult attendees of the 2014 Minnesota State Fair which attracted about 1,824,000 attendees or ~20% of all Minnesota residents. The University of Minnesota has a research facility for data collection on the Fair site where this study's data were collected in one of several booths in the building over six days of the 12 day fair. Individuals self-selected to interact with a study coordinator (SC) who invited participation and determined study eligibility. Inclusion criteria were: aged 18 or older, conversational competence in English, and not physically or cognitively impaired. Participants provided oral consent and the SC reviewed survey instructions and use of electronic tablets for data entry. Participants self-administered the survey until reaching the questions requiring oral administration. The SC then read aloud the "listening" section, and participants completed the remaining questions. An oral screening using the Basic Screening Survey (ASTDD, 2015) followed completion of the survey.

The University of Minnesota Institutional Review Board approved the study.

The correlation matrix of the questionnaire items was assessed to determine the extent of co-variance between constructs that represent all the possible relationships between individual responses from the instrument (Pett *et al*., 2003). Because item scoring was dichotomous, tetrachoric correlations were calculated. Resulting correlation coefficients of 0.3 or greater, meaning the items share about 10% or more of their variance, were used as a threshold to define notable correlations. Identifying these notable correlations in the data allowed for highly correlated variables to be grouped, which is useful for assessing the data's tendency to separate into different dimensions.

An item analysis was performed (Allen and Yen, 2001) with the proportion of correct response items compared in terms of their ability to characterize subjects along the item continuum (item 'difficulty'). Item 'discrimination' was measured using the corrected item-total correlations. This refers to the extent to which the questionnaire item differentiates between subjects with different levels of the trait. This was computed as the correlation between each item and the summed responses of the rest of the scale items.

Dimensionality was assessed by using exploratory factor analysis (EFA) and subjected a tetrachoric correlation matrix to EFA using the iterated principal factor method (Woods, 2002). Dimensionality was determined by considering the ratio of the first-to-second eigenvalue (Hattie, 1985), Cattell's scree plot (1966), and Horn's parallel analysis (1965). It has been suggested that a ratio of first-to-second eigenvalues >4 is evidence of unidimensionality (Reeve *et al*., 2007). Cattell's method plots the eigenvalues in decreasing order and retains as many factors as there are eigenvalues above the elbow of the plot. Horn's parallel analysis modifies Cattell's scree plot by comparing the observed eigenvalues to random simulated eigenvalues, retaining as many factors as the number of observed eigenvalues that exceed the simulated eigenvalues.

Cronbach's alpha was computed as a measure of the internal consistency among items. Known-groups validity was assessed by comparing OHL-AQ scores in groups of participants known to differ in their oral health (OH). Several hypotheses were investigated:

- Subjects with poorer self-reported oral health, measured by a global oral health indicator, should have lower OHL. Global oral health was assessed with the question "How would you describe the condition of your mouth and teeth --- including false teeth or dentures?" with response categories excellent, very good, good, moderate, and poor.
- Subjects with more impaired oral health-related quality of life (OHRQoL) should have lower OHL. Oral health quality of life was measured with the five-item Oral Health Impact Profile (OHIP-5) (John *et al*., 2006).
- Subjects with dentures should have lower OHL. Subjects with upper and with lower dentures were assessed separately using the Basic Screening Survey (ASTDD, 2015).
- Subjects with fewer teeth should have lower OHL. The "number of natural teeth present" was dichotomized as ≥20 teeth or <20 teeth (Sheiham *et al*., 2002).

- Treatment urgency assessed by a dental health professional should also be related to OHL, with higher treatment urgency observed in those with lower OHL. Treatment urgency was categorized as 'no obvious problem', 'early care' and 'urgent care'. Because of the small number of subjects early and urgent care were grouped together as 'treatment need'. Group differences were expressed as effect sizes.

Effect sizes of d=0.2 are considered to be small, 0.5 medium and 0.8 large (Cohen, 1988). For this computation data were dichotomised into 'excellent', 'very good' and 'good' for comparison with 'fair' and 'poor'. Likewise, OHIP-5 scores were grouped as 'better' (1st and 2nd quintiles) versus 'worse'. OHRQoL (3rd, 4th and 5th quintiles) and education level was grouped into 'college graduates' (postgraduate, some postgraduate and college graduate) versus 'non college graduates' ('trade/vocational/technical training', 'some college credit', 'high school graduates', and 'some high school').

All analyses were performed using the statistical software package STATA v13.1 and the probability of a type I error was set at the 0.05 α-level. For validity analyses, items with information missing were Global oral health indicator, n=2; OHIP scores, n=2; Presence of denture, n=7; Treatment urgency, n=9; Number of natural teeth present, n=8; Regular dentist/, n=3; and for Education level, n=2.

Results

Surveys were completed by 405 adults with a mean age of 45.0 (SD 15.7) years. Most were white (91%), non-Hispanic (89%), female (67%), and had at least a college education (72%). Many (86%) participants had 'adequate' OHL with OHL-AQ scores between 12-17, and the remaining participants presented equal proportions of 'inadequate' and 'marginal' OHL (OHL-AQ scores between 0-9, and 10-11, respectively).

The correlation matrix scores for the OHL-AQ items showed substantial, but variable correlations (Table 1). All 17 items correlated moderately (≥0.3) in 86 of 136 possible correlations. As over 60% of the items shared at least 10% of their information with each other, these results were interpreted as evidence that the OHL-AQ items shared common underlying factors. Average inter-item correlation was 0.18.

The proportion of correct responses for all items ranged from 22.2-97.8% (see Table 2). The largest proportion of correct responses was for a reading comprehension question (Brushing …at least twice a day [answer]…can prevent tooth decay). The smallest proportion of correct answers was for the free text response in the 'listening' domain. Item-rest correlations were between 0.21 and 0.50 indicating a substantial correlation between each item and the construct as a whole. Table 2 includes recoding definitions used subsequently in Table 3 to identify each survey question.

The eigenvalues of the tetrachoric correlation matrix suggested the presence of a dominant general factor. The first eigenvalue was 6.9 and the second eigenvalue was 1.1, yielding a first-to-second eigenvalue ratio of 6.2. Cattell's scree plot presented a steep drop from factor one to factor two, also supporting a strong general factor. Horn's parallel analysis also suggested a one-factor solution (Figure 1).

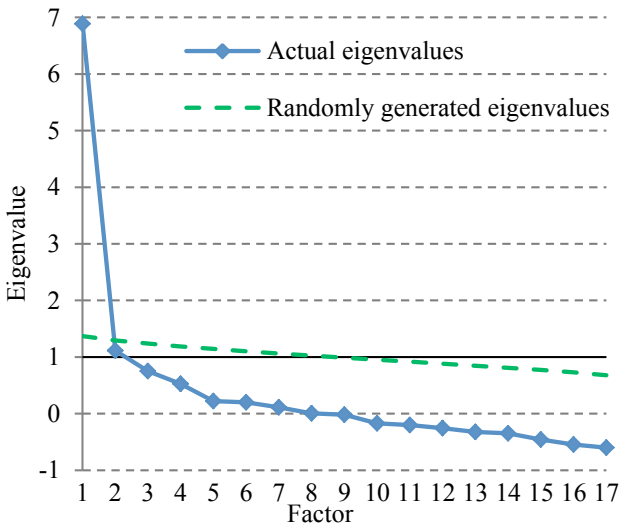

Figure 1. *Parallel analysis: plot of actual and randomly generated eigenvalues*

Table 1. Correlation matrix for the OHL-AQ Instrument

	Reading comprehension							*Numeracy*			*Listening*		*Decision-making*			
	Q1	*Q2*	*Q3*	*Q4*	*Q5*	*Q6*	*Q7*	*Q8*	*Q9*	*Q10*	*Q11*	*Q12*	*Q13*	*Q14*	*Q15*	*Q16*
	Link	*Fluoride*	*Brushing*	*Sugar*	*Teeth*	*Eruption*	*Tablet*	*Stop*	*Swallow*	*Eat*	*Gauze*	*Hot food*	*Eating*	*Stain*	*Pain*	*Legal*
Q2	**0.40**															
Q3	**0.44**	**0.61**														
Q4	**0.35**	**0.74**	**0.65**													
Q5	**0.40**	**0.58**	**0.57**	**0.69**												
Q6	0.22	0.18	0.11	0.28	**0.41**											
Q7	0.06	**0.36**	**0.47**	**0.41**	**0.32**	0.09										
Q8	0.13	**0.49**	**0.42**	**0.61**	**0.40**	0.26	**0.52**									
Q9	0.49	**0.68**	**0.66**	**0.71**	**0.50**	0.26	**0.49**	**0.74**								
Q10	0.07	0.21	**0.77**	**0.30**	0.23	0.10	**0.31**	**0.36**	**0.40**							
Q11	0.12	0.25	**0.74**	**0.39**	0.11	0.10	0.15	0.15	0.23	**0.70**						
Q12	-0.01	0.24	0.28	0.29	0.16	0.20	0.13	0.23	**0.49**	0.18	0.09					
Q13	0.24	**0.39**	**0.44**	**0.56**	**0.54**	**0.44**	**0.30**	**0.44**	**0.52**	0.20	0.26	**0.38**				
Q14	0.24	**0.33**	**0.39**	**0.59**	**0.39**	0.26	**0.34**	**0.43**	**0.60**	0.15	0.07	**0.46**	**0.64**			
Q15	0.24	**0.50**	**0.51**	**0.56**	**0.37**	**0.33**	0.19	0.28	**0.48**	**0.30**	**0.44**	**0.41**	**0.63**	**0.55**		
Q16	0.26	**0.53**	**0.53**	**0.65**	**0.48**	0.29	**0.36**	**0.44**	**0.57**	**0.35**	**0.34**	0.26	**0.62**	**0.54**	**0.74**	
Q17	0.12	**0.44**	**0.47**	**0.34**	**0.38**	0.23	0.27	0.24	**0.43**	0.23	0.24	0.20	**0.38**	**0.44**	**0.49**	**0.49**

Correlations ≥0.3 bolded; Refer to Table 2 for more detailed question wording

Table 2. Item analysis of OHL-AQ instrument questions

Section	*Item wording summary and recoding descriptors*		*Percentage of correct responses*	*Item-rest correlation*
Reading Comprehension	Q1, Link between oral and other diseases	(Link)	62.2	0.21
	Q2, Brushing with fluoride	(Fluoride)	93.0	0.40
	Q3, Brushing at least twice a day	(Brushing)	97.8	0.39
	Q4, Avoid sugar to prevent decay	(Sugar)	94.0	0.48
	Q5, Number of permanent teeth	(Teeth)	89.4	0.40
	Q6, Get the first tooth at 6 years old	(Eruption)	50.6	0.25
Numeracy	Q7, Time to take the tablet	(Tablet)	84.9	0.27
	Q8, Stop taking the medication	(Stop)	92.8	0.33
	Q9, Swallow the mouthrinse	(Swallow)	97.3	0.40
	Q10, Eat or drink	(Eat)	38.3	0.27
Listening	Q11, Take the gauze out of your mouth	(Gauze)	22.2	0.24
	Q12, Can you eat hot food	(Hot food)	89.0	0.21
Decision Making	Q13, Bleeding occurs after brushing	(Bleeding)	90.1	0.44
	Q14, Pain when swallowing	(Pain)	89.9	0.39
	Q15, Remove stains and calculus	(Stain)	88.6	0.46
	Q16, "I exonerate my dentist..."	(Legal)	84.9	0.50
	Q17, History of drug allergy	(Allergy)	75.0	0.34

Table 3. One-factor solution for 17 OHL-AQ items using recoding descriptors

Item	*Factor 1*	*Uniqueness*
Q1 Link	0.39	0.85
Q2 Fluoride	0.72	0.49
Q3 Brushing	0.81	0.35
Q4 Sugar	0.85	0.28
Q5 Teeth	0.67	0.56
Q6 Eruption	0.37	0.86
Q7 Tablet	0.49	0.76
Q8 Stop	0.63	0.60
Q9 Swallow	0.85	0.28
Q10 Eat	0.48	0.77
Q11 Gauze	0.44	0.81
Q12 Hot food	0.41	0.83
Q13 Eating	0.71	0.5
Q14 Stain	0.66	0.57
Q15 Pain	0.71	0.50
Q16 Legal	0.77	0.41
Q17 Allergy	0.54	0.70

See Table 2 for more detailed question wording

Based on the results of the parallel analysis, one factor was extracted. As displayed in Table 3, all OHL-AQ items had substantial loadings on this factor ranging from 0.37 to 0.85.

Internal consistency was considerable with a Cronbach's alpha of 0.74 with a lower limit of the 95% confidence interval of 0.70.

Almost all dose-response relationships between self-reported OHRQoL and OHL-AQ scores (Table 4) were statistically significant. While all relationships were observed in the expected direction, only the p-value for differences between OHL and treatment urgency was not statistically significant. The presence of dentures and treatment urgency presented a considerable influence on OHL-AQ scores in the hypothesized direction.

As predicted, higher OHL-AQ scores were also observed in participants with >20 natural teeth as compared to subjects with fewer natural teeth or in subjects who had a regular dentist compared to subjects without a regular dentist. Furthermore, educational level had an influence on OHL-AQ scores. Overall, even when some categories did not present the exact pattern predicted, e.g., high school graduates had slightly lower OHL scores than subjects with only some high school education, a trend for better OHL scores with higher levels of education was observed. When group differences were expressed in a standardized metric, "small" effect sizes were observed for self-reported OHRQoL, OHIP-5 scores, treatment urgency, and having a regular dentist. "Medium" effect sizes were seen for presence of dentures, number of natural teeth present, and educational level.

Discussion

This study provides dimensionality, reliability, and validity findings for scores of the English-language version of the OHL-AQ. The instrument is conceptually appealing and had promising basic psychometric properties in its original language. The instrument evidently has similar properties in English-speaking subjects, making it suitable for future use in this and similar populations.

Dimensionality is a psychometric property not explored in the original OHL-AQ or other OHL instrument analyses but is a fundamental property of instrument scores because it determines the number of scores required to characterize the construct adequately. This study's findings provide preliminary evidence that OHL measured with the OHL-AQ has a strong general factor. This is in contrast to the aforementioned theory of OHL. The construct is believed to be multidimensional, i.e., it supposedly has several important attributes within the OHL umbrella construct. The original OHL-AQ authors suggested that four components exist: reading comprehension, numeracy,

Table 4: Group Comparisons (Construct Validity) of OHL-AQ scores

Variables	*Nº of subjects*	*OHL-AQ 17 mean (SD)*	p¶	*Effect size, Cohen's d (95% CI)*
Global oral health indicator [a]	403		0.04	0.32 (0.05, 0.58)
Excellent	44	14.0 (1.7)		
Very good	129	13.7 (1.9)		
Good	163	13.3 (2.8)		
Fair	55	12.6 (3.7)		
Poor	12	13.3 (1.8)		
OHIP5 scores [a]	403		<0.001	0.31 (0.11, 0.51)
1 quintile (best OHRQoL)	117	13.8 (2.1)		
2	49	14.0 (1.3)		
3	105	13.5 (2.4)		
4	74	12.9 (3.2)		
5 quintile (worst OHRQoL)	58	12.6 (3.0)		
Presence of denture [b]	398		0.02	0.75 (0.00, 1.50)
No	391	13.5 (2.4)		
Yes	7	11.7 (3.4)		
Treatment urgency [b]	396		0.10	0.24 (-0.12, 0.60) [c]
No obvious problem	364	13.6 (2.4)		
Early care	31	13.0 (2.2)		
Urgent care	1	15.0 (0)		
Number of natural teeth present [b]	397		0.03	0.62 (-0.04,1.08)
More than 20 teeth	388	13.6 (2.3)		
Less than 20 teeth	9	12.0 (2.3)		
Regular dentist/dental home [b]	402		<0.001	0.43 (0.13, 0.73) *
Yes	349	13.6 (2.4)		
No	50	12.5 (3.0)		
Don't know	3	10.0 (6.2)		
Educational level [a]	403		<0.001	0.60 (0.38, 0.82)
Postgraduate degree	114	13.8 (2.3)		
Some postgraduate work	32	13.8 (2.0)		
College graduate	145	13.9 (1.9)		
Trade/vocational/tech. training	33	12.5 (3.8)		
Some college credit	46	12.8 (2.6)		
High school graduates	25	11.4 (3.6)		
Some high school	9	11.8 (3.6)		

Spearman rank correlation coefficient; ¶ T test or Anova test; [a] Spearman Correlation coefficient; [b] Point Biserial Corelation coefficient; *Excluding don't know category; [c] Excluding 'Urgent care' group because of too few subjects

listening and decision-making. Other authors suggest these and additional components such as communicating with health care providers and ability to navigate the health care system (Dickson-Swift *et al.*, 2014). While consensus among the research community regarding the exact OHL dimensions has not been established, the current results indicate that although multiple dimensions most likely exist, they are substantially correlated. The end result is that OHL can be summarized adequately with one total score.

While direct comparisons of this study's results with those of other OHL questionnaires are impractical because they did not perform dimensionality analyses, results can be compared to the widely used construct of oral health-related quality of life (OHRQoL). Here, theory also supports multidimensionality of the construct and several studies point to the existence of dimensions although these are poorly understood (John *et al*., 2014b). Specifically, in a large study using the Oral Health Impact Profile to determine the dimensions of OHRQoL (John *et al*., 2014b), four dimensions (oral functions, orofacial pain, orofacial appearance, psychosocial impact) were found with different factor analytic approaches (John *et al*., 2014a; c). Notably, the OHRQoL construct could also be characterized with the total score of the general OHRQoL factor. It seems that the OHRQoL construct and OHL – two of oral health's major latent concepts - share similarities, in particular their dimensionality. As with OHRQoL, this study's OHL findings suggest that this multidimensional construct can also be characterized by an overall score and that possible dimensions are not too different so that they can be summarized.

Following the determination of score dimensionality, reliability findings can be interpreted. Internal consistency was used as a measure of reliability. Satisfactory reliability (α=0.74) was observed consistent with the original Persian version (α=0.72) (Naghibi Sistani *et al*., 2014). Other OHL instruments such as the TOFHLiD (Gong *et al*., 2007) had lower (α=0.63); similar, Comprehensive Measure of Oral Health Knowledge (α=0.74) (Maeck *et*

al., 2010); or higher reliability REALMD 20 item version (α=0.86) (Gironda *et al*., 2013), REALMD 84 item version (α=0.96) (Atchison *et al*., 2010). However, the varying results may reflect the heterogeneity of instrument length as longer instruments tend to have higher internal consistency. Cronbach's alpha is a function of the inter-item correlation and therefore, in this study, a low inter-item correlation of 0.17 translated into a respectable Cronbach's alpha of 0.74. According to the Spearman-Brown Formula, a theoretical 84-item version of the study's instrument, would give an alpha of 0.93. This result is similar to the 84-item REALMD and may indicate that internal consistency among OHL instruments does not differ substantially when taking the number of items into account. Conversely, the findings and the low inter-item correlation among OHL instruments in general could be important to consider when determining the components of the construct.

This study's dimensionality finding is in line with earlier studies investigating OHL validity. Previous authors often reported a questionnaire summary score suggesting that they considered the OHL construct as sufficiently unidimensional to be adequately described with one score (Lee *et al*., 2007; Richman *et al*., 2007; Stucky *et al*., 2011). The small and moderate effect size estimates also support the construct validity of OHL-AQ scores. The results can be compared to a medium effect size of 0.34 for HeLD-14 (Jones *et al*., 2014) and 0.32 for HeLD-29 (Jones *et al*., 2015) for a comparison of global oral health indicator and OHL-AQ score.

Associations between oral health measures and OHL-AQ scores varied as compared to other instruments reporting similar relationships in the literature. A correlation coefficient of 0.16 was observed for OHIP-5 scores in this study (not presented in results.) Correlation results between the REALD-30 (Lee *et al*., 2007) and the OHIP-14 showed a similar relationship (r=0.14). Two other studies showed stronger correlations with the OHIP-14. The REALD-99 (Richman *et al*., 2007) reported an absolute value of 0.74. This result seems high in light of a correlation of 0.90 (Larsson *et al*., 2014; Shueb, 2014) for two instruments that measure the same construct (14-item version and 5 item version of OHIP as measures of OHRQoL) and also share some items. A lower correlation was reported for the TOFHLiD (Gong *et al.,* 2007) and OHIP-14 (r= 0.50). A correlation of this magnitude can be compared to correlations between two different instruments that measure the same construct. For example, a global self-reported oral health indicator, measuring perceived oral health, and OHIP-14 or OHIP-5, also characterizing perceived oral health, correlated with 0.53 (Larsson *et al*., 2014) or 0.48 in a Japanese prosthodontic patient population (Baba *et al*., 2008).

On calculating correlation coefficients to express the relationship between OHL scores and validity variables (findings not presented in the results section), the correlation of r=0.09 between perceived oral health and OHL-AQ was similar to the findings from REALMD (Atchison *et al*, 2010) (r=0.07), but substantially lower than both the REALD 30-item (Lee *et al*., 2007) (r=0.35) and 99-item versions (Richman *et al.*, 2007) (r = 0.61). The magnitude of the REALD-99 is surprising because it reaches levels of validity coefficients usually observed between two instruments measuring the same construct.

As with some other studies, participants with less education scored lower on the OHL-AQ. Indeed, this study found a difference between educational levels and the OHL-AQ. The REALMD (Atichison *et al*., 2010), REALMD 20 (Gironda *et al*., 2013) and the Comprehensive Measure or Oral Health Knowledge (CMOHK) (Macek *et al*., 2010) also found a significant relationship between educational attainment and oral health literacy levels. In contrast, the associations of Hong Kong Rapid Estimate of Adult Literacy in Dentistry (HKREALD-30) (Wong *et al*., 2012) and Oral Health Literacy Instrument (OHLI) (Sabbahi *et al*., 2009) with educational level were not statistically significant. This variation in the magnitude of this association implies that educational level alone is not clearly understood, and is not an adequate conceptual proxy for oral health literacy.

The sample size in this study was relatively large and likely included broader segments of the population as compared to prior OHL studies restricted to specialized dental patients during routine clinic visits. The sample, compared to the whole state of Minnesota, was less racially but more ethnically diverse (91% vs 85% White, and 89% vs 96% non-Hispanic), and more highly educated (98% vs 72% with a high school education or greater). More participants were female compared to Minnesota's general population (67% vs 50%) and were older (45 vs 37 years of age) (US Census Bureau, 2015). While a randomly selected sample would have been preferable, representativeness is not required for psychometric analyses. Limitations included the self-reported nature of some covariates. Future research should replicate and extend validation activities of the OHL-AQ in more diverse populations, in particular, populations with expected low OHL, and with subjective as well as a larger number of objective markers of oral health. Test-retest reliability is an important part of instrument development, a psychometric property that was not evaluated in this study.

Conclusion

Perhaps this study's most important finding, subject to further confirmation in other populations, is that OHL can be validly gauged with the single overall score from the OHL-AQ. The lower than expected correlations with some other important oral health indicators provides stimulating insight about the construct of OHL, perhaps indicating that the OHL-AQ measures a rather distinct aspect of OHL when compared to other tools.

References

Allen, M. and Yen, M. (2001): *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.

American Dental Association (2009): American Dental Association Council on Access, Prevention and Interprofessional Relations. In: *Health Literacy in Dentistry Action Plan 2010–2015*. Chicago: ADA.

Association of State & Territorial Dental Directors, ASTDD (2015): *ASTDD Basic Screening Surveys*. http://www.astdd.org/basic-screening-survey-tool

Atchison, K., Gironda, M., Messadi, D., Der-Martirosian, C. (2010): Screening for oral health literacy in an urban dental clinic. *Journal of Public Health Dentistry* **70**, 269-275.

Baba, K., Inukai, M. and John, M. (2008): Feasibility of oral health-related quality of life assessment in prosthodontic patients using abbreviated Oral Health Impact Profile questionnaires. *Journal of Oral Rehabilitation* **35**, 224-228.

Cattell, R. (1966): The scree test for the number of factors. *Multivariate Behavioral Research* **1**, 245-276.

Cohen, J. (1988): *Statistical Power Analysis for the Behavioral Sciences*. L. Hillsdale, NJ: Eribaum Associates.

Dickson-Swift, V., Kenny, A., Farmer, J., Gussy, M. and Larkins, S. (2014): Measuring oral health literacy: a scoping review of existing tools. *BMC Oral Health* **14**, 148.

Gironda, M., Der-Martirosian, C., Messadi, D., Holtzman, J. and Atchison, K. (2013): A brief 20-item dental/medical health literacy screen (REALMD-20). *Journal of Public Health Dentistry* **73**, 50-55.

Gong, D., Lee, J., Rozier, R., Pahel, B., Richman, J. and Vann, W. (2007): Development and testing of the test of functional health literacy in dentistry (TOFHLiD). *Journal of Public Health Dentistry* **67**, 105-112.

Hattie, J. (1985): Methodology review: assessing unidimensionality of tests and ltenls. *Applied Psychological Measurement* **9**, 139-164.

Horn, J.L. (1965): A rationale and test for the number of factors in factor analysis. *Psychometrika* **30**, 179-185.

John, M.T., Feuerstahler, L., Waller, N., Baba, K., Larsson, P., Čelebić, A., Kende, D., Rener-Sitar, K. and Reissmann, D.R. (2014a): Confirmatory factor analysis of the Oral Health Impact Profile. *Journal of Oral Rehabilitation* **41**, 644-652.

John, M.T., Miglioretti, D., LeResche, L., Koepsell, T., Hujoel, P. and Micheelis, W. (2006): German short forms of the oral health impact profile. *Community Dentistry and Oral Epidemiology* **34**, 277-288.

John, M.T., Reissmann, D.R., Feuerstahler, L., Waller, N., Baba, K. and Larsson, P. (2014b): Dimensions of Oral Health-Related Quality of Life Project–overview and studied population. *Journal of Prosthodontic Research* **1**, 26-34.

John, M.T., Reissmann, D.R., Feuerstahler, L., Waller, N., Baba, K., Larsson, P., Szabo, G. and Rener-Sitar, K. (2014c): Exploratory factor analysis of the Oral Health Impact Profile. *Journal of Oral Rehabilitation* **41**, 635-643.

Jones, K., Brennan, D., Parker, E and Jamieson, L. (2015): Development of a short-form Health Literacy Dental Scale (HeLD-14). *Community Dentistry and Oral Epidemiology* **43**, 143-151.

Jones, K., Parker, E., Mills, H., Brennan, D. and Jamieson, L. (2014): Development and psychometric validation of a Health Literacy in Dentistry scale (HeLD). *Community Dental Health* **31**, 37-43.

Kann, L., Brener, N., Warren, C., Collins, J. and Giovino, G. (2002): An assessment of the effect of data collection setting on the prevalence of health risk behaviors among adolescents. *Journal of Adolescent Health* **31**, 327-335.

Larsson, P., John, M., Hakeberg, M., Nilner, K. and List, T. (2014): General population norms of the Swedish short forms of Oral Health Impact Profile. *Journal of Oral Rehabilitation* **41**, 275-281.

Lee, J., Rozier, R., Lee, S., Bender, D and Ruiz, R. (2007): Development of a Word Recognition Instrument to Test Health Literacy in Dentistry: The REALD-30–A Brief Communication. *Journal of Public Health Dentistry* **67**, 94-98.

Macek, M., Haynes, D., Wells, W., Bauer-Leffler, S., Cotten, P. and Parker, R. (2010): Measuring conceptual health knowledge in the context of oral health literacy: preliminary results. *Journal of Public Health Dentistry* **70**, 197-204.

National Institute of Dental and Craniofacial Research (2005): The invisible barrier: literacy and its relationship with oral health. *Journal of Public Health Dentistry* **65**, 313-321.

Naghibi Sistani, M.M., Yazdani, R., Virtanen, J., Pakdaman, A. and Murtomaa, H. (2013): Determinants of oral health: does oral health literacy matter? *ISRN Dentistry* **13**, 249591.

Naghibi Sistani, M.M., Montazeri, A., Yazdani, R. and Murtomaa, H. (2014): New oral health literacy instrument for public health: development and pilot testing. *Journal of Investigative and Clinical Dentistry* **5**, 313-321.

Pett, M., Lackey, N. and Sullivan, J. (2003): *Making Sense of Factor Analysis*. Thousand Oaks, CA: Sage.

Reeve, B., Hays, R., Bjorner, J., Cook, K., Crane, P., Teresi, J.A., Thissen, D.; Revicki, D.A., Weiss, D.J., Hambleton, R.K., Liu, H., Gershon, R., Reise, S.P., Lai, J. and Cella, D. on behalf of the PROMIS Cooperative Group (2007): Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care* **45**, S22-31.

Rhodes, S., Bowie, D. and Hergenrather, K. (2003): Collecting behavioural data using the world wide web: considerations for researchers. *Journal of Epidemiology and Community Health* **57**, 68-73.

Richman, J., Lee, J., Rozier, R., Gong, D., Pahel, B. and Vann, W. (2007): Evaluation of a Word Recognition Instrument to Test Health Literacy in Dentistry: The REALD-99. *Journal of Public Health Dentistry* **67**, 99-104.

Sabbahi, D., Lawrence, H., Limeback, H and Rootman, I. (2009): Development and evaluation of an oral health literacy instrument for adults. *Community Dentistry and Oral Epidemiology* **37**, 451-462.

Sheiham, A., Steele, J., Marcenes, W., Finch, S. and Walls, A. (2002): The relationship between oral health status and body mass index among older people: a national survey of older people in Great Britain. *British Dental Journal* **192**, 703-706.

Shueb, S. (2014): *Comparing the oral health related quality of life in four orofacial pain conditions*. University of Minnesota Digital Conservancy. http://hdl.handle.net/11299/170834.

Stucky, B., Lee, J., Lee, S. and Rozier, R. (2011): Development of the two-stage rapid estimate of adult literacy in dentistry. *Community Dentistry and Oral Epidemiology* **39**, 474-480.

US Census Bureau. Acessed 14/01/2016. www.census.gov.

Wong, H., Bridges, S., Yiu, C., McGrath, C., Au, T. and Parthasarathy, D. (2012): Development and validation of the Hong Kong Rapid Estimate of Adult Literacy in Dentistry. *Journal of Investigative Clinical Dentistry* **3**, 118-127.

Woods, C.M. (2002): Factor analysis of scales composed of binary items: illustration with the Maudsley Obsessional Compulsive Inventory. *Journal of Psychopathology and Behavioral Assessment* **24**, 215–216.