DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 15 | Issue 1 Article 22

5-2016

Validity of the Persian Blog Attitude Questionnaire: An Evidence-Based Approach

Vahid Aryadoust
National University of Singapore, vahid.aryadoust@nus.edu.sg

Zahra Shahsavar
Shiraz University of Medical Science, shahsavarzahra@gmail.com

Validity of the Persian Blog Attitude Questionnaire: An Evidence-Based Approach

Cover Page Footnote

1. For consistency and to avoid confusions, the term dimension is used across PCA, CFA, and RSM analyese.

Journal of Modern Applied Statistical Methods
May 2016, Vol. 15, No. 1, 417-451.

ISSN 1538 − 9472

Validity of the Persian Blog Attitude Questionnaire: An Evidence-Based Approach

Vahid Aryadoust
National University of Singapore
Singapore

Zahra Shahsavar
Shiraz University of Medical Sciences
Shiraz, Iran

Validity evidence is provided for a Persian blog attitude questionnaire (P-BAQ). P-BAQ was administered to 565 Iranians and factor analysis and rating scale model identified affective, behavioral, and perseverance, and confidence dimensions underlying the data. P-BAQ's validity argument was supported by the theoretical and psychometric evidence, although adding a few items to the instrument would improve its construct representativeness.

Keywords: blog, confirmatory factor analysis, Persian blog attitude questionnaire, principal component analysis, Rasch-Andrich rating scale model, validity argument

Introduction

Advanced technology has resulted in a wide range of Web 2.0 tools such as blogs that can be used for various purposes such as disseminating information and launching discussions (Wang & Woo, 2010). Blogs have recently been adapted into educational settings; they have facilitated information access and interaction among students and allowed them to collaborate on group projects, edit or add to page contents, and discuss their projects with their peers in a friendly and less formal environment (Liaw, Huang, Chen, 2007).

Recently, researchers have shown an increased interest in measuring students' attitudes toward technology (Chen, Shih, & Liu, 2013). Blog attitude refers to users' inclination to use blogs for various purposes such as disseminating

*Footnote: For consistency and to avoid confusions, the term dimension is used across PCA, CFA, and RSM analyese.

Vahid Aryadoust is Lecturer at the Centre for English Language Communication of the National University of Singapore. Email him at: vahid.aryadoust@nus.edu.sg. Zahra Shahsavar is Assistant Professor at Shiraz University of Medical Sciences. Email her at: shahsavarzahra@gmail.com.

information and introducing educational programs. Much of the interest in studying blog attitude is due to the reported influence of attitude on students' educational performance (Ayres, 2002). Studies show that in educational settings where blogs are used, students who harbor a positive attitude toward blogs would find learning attractive, efficient, and practical (Fageeh, 2011; Chu & Kamal, 2008). Positive attitudes toward blogs help students adapt themselves to a new set of learning and teaching techniques that technology offers, thereby allaying their stress and discomfort (Al-Tamimi & Shuib, 2009; McIntosh, 2009). By contrast, students with negative attitudes toward blogs would be less willing to involve themselves in learning because they might view technology as a threat to their learning (Shahsavar & Tan, 2011) or a threat to their well-established and traditional techniques (Pektas & Erkip, 2006).

To assess learners' attitudes toward technology, researchers have employed questionnaires. Substantial research has focused on examining learners' attitudes toward computers by using popular computer attitude scales (e.g., DeVellis, 2003; Loyd & Gressard, 1984; Selwyn, 1997; Shaft, Sharfman, & Wu, 2004). However, relative to computers, little research has been performed on blog attitude (e.g., Blackstone, Spiri, & Naganuma, 2007; Harwood, 2010). One of the limitations of the blog attitude questionnaires is that they are often written in English or are not clearly translated into the respondents' mother tongue. According to Harkness (2008), the inaccuracy of the translated questionnaire may adversely affect respondents' comprehension and consequently undermine the reliability of the data. Instruments should be presented in respondents' own mother tongue to enhance the quality and precision of data.

In addition, most blog attitude questionnaires present no solid evidence of psychometric quality and no explicitly articulated validity argument (see Kane, 2013, for a discussion on the importance of these requirements). To develop psychometrically valid questionnaires, an item pool containing items from available questionnaires should be initially developed. Items adapted from instruments written in a foreign language should be translated into the respondents' mother tongue. To ascertain construct equivalence—i.e. that the items in English and the respondents' mother tongue tap the same construct—the precision of the translated items should be verified by experts (Dörnyei & Taguchi, 2010).

The intent of the present study is develop and examine the psychometric features of the Persian Blog Attitude Questionnaire (P-BAQ). Developing P-BAQ is primarily motivated by the lack of validated Persian blog attitude questionnaires and a need for further research into Iranian students' blog attitude.

This lack can make it difficult for researchers in Iran (and other Farsi/Persian speaking countries) to perform research on blog attitude and obtain reliable and reproducible results. Although the need for validation and rigorous data analysis has been identified in attitude measurement research (e.g., Bangert, 2009; Morse, Gullekson, Morris, & Popovich, 2011; Zhang, 2007), previous research tends to overlook the importance of validation as an argument. The researchers survey the relevant literature and use factor analysis and Rasch-Andrich rating scale model (RSM) to examine the psychometric quality of P-BAQ. Subsequently, the evidence gleaned from the psychometric analyses is used to build a coherent validity argument for the instrument.

Literature Review

Attitude

Pickens (2005) defines attitude as learners' way of evaluating objects (or people) positively or negatively. There is consensus among researchers on the effect of students' attitude toward technology tools on their motivation and learning (Vandewaetere & Desmet, 2009). However, there has yet to be an agreement on the structure of attitude as a measureable construct; this lack of consensus has resulted in the development of multidimensional and unidimensional models of attitude in technology attitude research (Franzoi, 2003).

Researchers endorsing attitude multidimensionality propose several underlying dimensions, three commonly stated of which are affective, behavioral, and cognitive dimensions (Eagly & Chaiken, 1993; McLeod, 2009; Hogg & Vaughan, 2011; Mantle-Bromley, 1995; Mantle-Bromley & Miller, 1991; Zimbardo & Leippe, 1991). McLeod (2009) argues that these three dimensions would constitute a basic model for attitude, suggesting that attitude might emerge as a three- dimensional construct. The affective dimension refers to individuals' feelings toward people and using objects in their life. For example, students might express their dislike for course blogs, which results from their negative attitudes toward it. The behavioral dimension refers to individuals' aim to use objects or act in a particular situation. For example, students who dislike course blogs may be less participatory and attentive in class than the students who find blogging enjoyable. The cognitive dimension engages individuals' knowledge and beliefs about objects or people. For example, the students who disapprove of using course blogs may believe that technology is of no practical use for learning and teaching (McLeod, 2009).

Some researchers who endorse attitude's multidimensionality have excluded the cognitive dimension, arguing that cognition is a determinant of affective and behavioral dimensions rather than an attitudinal component. For example, Min (1998, p. 23) defines attitude as "an evaluative response to the environment, ideas, objects, and other people" which is predicted by individuals' beliefs, knowledge, and perceptions. Similarly, Mantle-Bromley (1995) argues that attitude is composed of individuals' emotions such as detest, fear, and love as well as behaviors. Selwyn's (1997) application of factor analysis yields empirical evidence backing Min's (1998) and Mantle-Bromley's (1995) conceptualizations of attitude, eliminating the cognitive dimension due to its poor loading coefficients. Tsai, Lin, and Tsai's (2001) and Kay's (1993) studies also yield highly similar results to Silwyn's study, casting doubt on the validity of the postulated cognitive dimension.

Contrary to the studies supporting the multidimensional structure of attitude, some researchers argue that the posited attitude dimensions may be statistically inseparable or unidimensional. Daud (1995) applied factor analysis on a pool of items measuring language teachers' and students' attitude toward computers, where a single attitude factor emerged. Smith, Caputi, and Rawstorne (2000) also found that attitude's posited dimensions were inseparable, arguing that the items that elicit students' attitudinal and "psychological reactions" might not load on distinct factors. Although this line of research is relatively underdeveloped, it is important to compare the fit of unidimensional and multidimensional structures to the data to ascertain the psychometric validity of the scale (Teo, 2012).

Finally, a number of educators have used the basic model of attitude to develop computer and internet attitude questionnaires (e.g., Jones & Clarke, 1994; Liaw, 2002; Fančovičová & Prokop, 2008; Sam, Othman, & Nordin, 2005). However, there is still no blog attitude questionnaire—specifically for Persian (and Farsi) speakers—with a reliable psychometric structure and validity argument. The present study seeks to develop a Persian blog attitude questionnaire (P-BAQ); examine its psychometric features; and lay out a validity argument for it.

Building P-BAQ's Validity Argument

For the validation of psychometric instruments, Kane (2006, 2013) proposed a framework with two main components: i) the claimed uses of the instrument and interpretation of its scores and ii) the supporting evidence. The advantage of this framework is that research findings can be organized into a "cohesive treatment"

for validation of a psychometric instrument (Aryadoust, 2013, p. 12). In greater detail, this framework consists of inferences made from the data, the evidence supporting these inferences, any assumptions being made, and the justifications for these assumptions (Aryadoust, Mehran, & Alizadeh, 2014). Hence, for the purpose of this study, we assert:

(a) *The domain definition inference* infers that Iranian tertiary students have a multicomponential attitude toward blogs that include an affective component (i.e., positive or negative feelings about the use of blogs for education) and a behavioral component (i.e., a preference for using blogs for education). This inference requires the following assumptions: i) we can determine the multiple components of attitude toward blogs and ii) we can measure these components. As the literature survey indicated that attitude toward blogs has multiple components, the first assumption is supported. The pool of attitude items identified from the literature supports the second assumption. Further discussion of the literature survey supporting the content-representativeness of the items can be found under *Attitude*.

(b) *The translation inference* occurs during the translation of the items into Persian and infers that the items have been clearly translated and do not contain any ambiguous language. This inference requires the following assumptions: i) the items can be translated and ii) the translated items can be unambiguously understood by Persian speakers. The use of a panel of expert translators to consult on the clarity and accuracy of the translations supports these assumptions.

(c) *The evaluation inference* infers that participants' responses are consistent and thus produce reliable data. This inference assumes that the scoring categories (i.e., a four-category Likert scale) is suitable. This assumption is warranted because a four-category Likert scale prevents respondents from only endorsing the mid-points of the scale ("flat-lining"). The psychometric qualities of the Likert scale are supported by the monotonic increments of its response categories and the fit statistics for each response category. This is supported by the RSM results, as presented below.

(d) *The generalization inference* infers that observed scores can be generalized to the universe of scores. This inference assumes that performance on the survey represents performance on the universe of

items. This assumption is warranted by psychometric measurements and the reliability statistics estimated by RSM.

(e) *The explanation inference* infers a relationship between the observed scores and the underlying construct. This inference requires the following assumptions: i) P-BAQ has good construct representativeness, ii) the sub-components of P-BAQ are highly correlated, and iii) there are no construct-irrelevant factors. These assumptions are warranted because the structure of P-BAQ is supported by theory, P-BAQ's sub-components are related and can be distinguished from each other, and the psychometric qualities of P-BAQ are reliable with no construct-irrelevant factors, which can be confirmed by dimensionality analysis. Psychometric analysis further supports this inference; specifically, principal component analysis (PCA) supports the multidimensional structure, confirmatory factor analysis (CFA) supports a multidimensional structure in a different sample, and RSM analysis supports the unidimensionality of P-BAQ's sub-components and shows there are no gaps in the item-person map (Aryadoust, 2013). (For consistency, the term dimension is used across PCA, CFA, and RSM analyses.)

As a one-to-one correspondence between the data analysis techniques and postulations is not possible (Aryadoust, 2013; Kane, 2013), multiple psychometric analyses are performed to obtain data supporting the inferences described above, e.g., RSM results can be used to support the generalization and explanation inferences while CFA results can be used to support the explanation inference.

Methodology

Participants

A sample of 565 Iranian tertiary-level students aged between 17 and 30 (M = 21.98; SD = 2.47) completed P-BAQ in late 2013. Among these students, 254 (44.6%) were male and 298 (53.1%) were female (missing gender information: 13; 2.3%), thereby yielding a balanced gender distribution. Of these, 361 reported their field of study, as follows: Dentistry: 84 (15%); Medicine, Hygiene, and Speech Therapy: 68 (12%); English Translation: 25 (4.5%); Computer Engineering: 117 (20.9%); and Industrial Engineering: 67 (11.9%). All students indicated their familiarity with blogs; had personal computers and home

Internet access; and had used blogs and/or one of the Web 2.0 tools such as email and Facebook.

The sample was randomly divided into two sub-sets containing 282 and 283 participants for psychometric analysis (see below).

Developing P-BAQ

Before attempting to develop their own tool, existing technology attitude questionnaires were surveyed, such as those developed by Blackstone, Spiri, and Naganuma (2007), Harwood (2010), and Shahsavar and Tan (2012). This resulted in an initial pool of 100 English language items, which was reduced after ambiguous/imprecise items were removed (see Aryadoust et al., 2014, for more details of such a process).

This refined item pool was then submitted to a panel of five experts – two with PhD in applied linguistics or educational/instructional technology and three PhD candidates in applied linguistics or English. The experts were identified as having expert knowledge in incorporating new technology (blogs, social media) in education, as well as practical experience developing questionnaires about online learning. Experts were selected with this background to optimize construct-representativeness. Each expert performed an independent review of item pool, resulting in 29 items identified by at least three of the experts.

These 29 items were then translated into Persian by the authors. The translations were submitted to experienced translators who reviewed the clarity, naturalness, and appropriateness (Farahzad, 1992) of the translation. In the end, a 29-item P-BAQ was formalized that evaluates Iranian blog users' attitudes using a four-point response scale: *strongly disagree (1)*, *disagree (2)*, *agree (3)*, or *strongly agree (4)* (see the Appendix).

Data Analysis

Principal Component Analysis (PCA)

The univariate normality of items estimated by skewness and kurtosis indices were initially examined—values falling between -1 and +1 indicate normality. To determine the constituent structure of P-BAQ, multiple PCA were performed on the first sub-set (n = 280) on IBM SPSS Statistics for Windows, Version 21.0 (IBM Corp., 2012).

Because Items 1, 4, and 19 displayed cross-loading patterns, they were eliminated one at a time. To achieve the optimal results, different rotation

methods were performed including oblique rotation methods which allow the components to correlate (i.e., direct oblimin and promax) and orthogonal rotation which does not permit the components to correlate (i.e., equimax, quartimax, and varimax).

The Kaiser criteria (where eigenvalues greater than unity represent independent components) of the various models were examined to identify what model resulted in the best solution. To determine the number of retainable components, the scree plot, which represents the number of substantive components, was examined (Kline, 2006). The communality (h^2), sum of the squared loadings per item, eigenvalues per component, or variance explained by each component were also estimated, where appropriate. Lastly, Cronbach's alpha indices were computed for each test component. Cronbach's alpha can be used to evaluate the internal consistency of test items, with a value greater than 0.70 indicating high consistency, 0.50 to 0.70 indicating moderate consistency, and a value below 0.50 indicating low consistency.

Confirmatory Factor Analysis (CFA)

AMOS (version 21) was used to perform a two-stage CFA on the second subset (n = 281) in order to test the fit of the optimal solution yielded in the PCA analysis. Initially independent CFA measurement models were generated for each component emerging in PCA. Each measurement model included a latent variable representing the sub-component measured by items alongside the items tapping that sub-component. Next, the complete CFA model was tested comprising the correlated sub-components and related items. Maximum Likelihood (ML) method of parameter estimation and multiple fit criteria to evaluate the fit of the model were employed as follows:

(a) Chi-square test (χ^2): An index representing the difference between the observed and implied covariance or correlation matrices. Non-significant χ^2values suggest good fit, although large samples can inflate this index.

(b) Normed χ^2 (χ^2 / df): The ratio of χ^2 to the degrees of freedom (df). This ratio is small in well-fitting models (preferably below 3).

(c) Two incremental indices: Non-Normed Fit Indices (NNFI) and Comparative Fit Indices (CFI). Both indices compare the postulated model to a baseline model that assumes that measures are not

correlated. Indices of 0.90 or above were chosen as indicators of satisfactory fit.

(d) Root Mean Square Error of Approximation (RMSEA): This measure is used to adjust the significance level for chi-square tests when using large samples; a low RMSEA value is preferred.

To measure the effect of the constructs or dimensions on participants' performance, standardized regression estimates were used which quantify the increase in the standard deviation of the items if the standard deviation of the construct increases by one unit. Similarly, non-standard regression coefficients indicate the magnitude of increase in the items when the construct increases by one unit. The critical ratios (CRs) were estimated for each regression estimate by dividing them by their standard deviation. CRs greater than 1.96 are statistically significant ($p < 0.05$).

Rasch-Andrich Rating Scale Model (RSM)

RSM analysis (Andrich, 1978) were performed using the *WINSTEPS* computer package, Version 3.75 (Linacre, 2013a). Separate RSM analyses were performed on each subscale (component) and examined the hierarchy of item and person measures in each subscale. Reliability and separation estimations were examined, adherence to unidimensionality, and psychometric features of response categories in each subscale.

The fit of the data to the model was estimated by calculating the infit and outfit mean square (MNSQ) statistics. Infit MNSQ is sensitive to perturbations near the ability level of individuals or difficulty level (endorsability) of items. Outfit MNSQ is sensitive to outliers, meaning that if a participant with a low trait level endorses a high response category on a difficult item, she is an outlier inflating the outfit MNSQ (Brodersen, Thorsen, & Kreiner, 2007). Inliers are weighted so as to decrease this sensitivity and balance out the effect of outliers. MNSQ values should be equal to unity, though a narrow deviation from unity, that is 0.6 – 1.4, is acceptable (Bond & Fox, 2007). If MNSQ indices are lower than 0.6, the item does not provide much information and is redundant. Conversely, if the value is greater than 1.4, the item is inconsistent with the rest of the items and is likely contaminated by construct-irrelevant variance (Linacre, 2013b).

To investigate adherence to unidimensionality, a principal component analysis of Rasch linearized residuals (PCAR) was carried out. This analysis is

performed after the RSM dimension is extracted from the data and the researcher intends to examine potentially substantive structures in the residuals. If an identified structure has substance (eigenvalues > 2.5), it may represent a dimension beside the RSM dimension (Linacre, 2013b). Variance explained was assessed by the RSM dimension with the components which emerged in the PCAR.

The response category features, thresholds, and category measures of the affective dimension were examined. At threshold points, the adjacent response category curves intersect and participants with ability measures equal to or greater than the threshold begin to have a higher probability to select the higher category. Thresholds should increase monotonically from a lower to a higher index. Monotonic increase should ideally be larger than 1.4 but smaller than 5 logits (Bond & Fox, 2007).

Results

Descriptive Statistics of Data

Table 1 presents the descriptive statistics of the items including mean scores, standard deviation, skewness, and kurtosis. Items 29 and 19 had the lowest and highest mean values of 1.82 and 3.21, respectively. The skewness and kurtosis indices fall between -1 and +1, indicating that the sample size had a normal distribution.

Principal Component Analysis (PCA)

Multiple PCA on the first subset ($n = 282$) were performed. Due to space constraints, only the results of the optimal model in this section are reported. After conducting several PCA, Items 1, 4, 19, and 27 were deleted due to their cross-loading patterns in several models. The 26 items left loaded on three components which was called *affective* dimension, *behavioral* dimension, and *perseverance and confidence* dimension (correlation coefficients: .535, .329, and .440). To estimate this model, promax with Kaiser normalization rotation was used, which outperformed other rotation techniques. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.888, indicating that the variance in the data was caused primarily by the underlying construct and that PCA can be conducted.

Table 1. Descriptive Statistics of the Items (n = 565)

Item	Mean	Standard deviation	Skewness	Kurtosis
1	2.97	.869	-0.750	0.084
2	2.96	.738	-0.595	0.489
3	1.87	.683	0.506	0.387
4	2.94	.779	-0.665	0.401
5	2.81	.766	-0.380	-0.061
6	2.15	.795	0.218	-0.495
7	2.17	.766	0.130	-0.495
8	1.96	.683	0.579	0.835
9	2.81	.719	-0.410	0.202
10	2.73	.766	-0.327	-0.136
11	2.83	.743	-0.296	-0.114
12	2.08	.750	0.429	0.058
13	2.10	.745	0.402	0.038
14	2.05	.682	0.414	0.421
15	2.88	.735	-0.474	0.250
16	2.84	.702	-0.547	0.559
17	2.04	.787	0.478	-0.089
18	2.18	.791	0.278	-0.337
19	3.21	.763	-0.959	0.988
20	2.45	.737	0.328	-0.216
21	1.96	.717	0.526	0.360
22	2.20	.831	0.355	-0.366
23	2.03	.729	0.475	0.246
24	2.71	.755	-0.454	0.042
25	2.62	.783	-0.167	-0.355
26	2.66	.749	-0.270	-0.161
27	1.97	.740	0.555	0.295
28	2.71	.719	-0.224	-0.102
29	1.82	.774	0.942	0.930

This finding was further supported by the significant value of the Bartlett's test of sphericity which was 0.00 ($p < 0.05$), indicating that the correlation matrix was not an identity matrix—a matrix where all diagonal values are unity and off-diagonal indices are zero.

Table 2 presents the three-componential PCA model. Loading coefficients and communality values of all items (except Item 3) are greater than 0.300 indicating that the amount of variance explained by the components is significantly high. For example, Items 3, 6, 7, 8, 12, 13, 14, 17, 18, 21, 23, 26, and 29 loaded on the affective dimension with loading coefficients ranging from .458 to .884. Cronbach's alpha coefficients of the affective and behavioral dimensions are greater than .800, indicating significantly high internal consistency; perseverance and confidence dimension has a moderate Cronbach's alpha coefficient of .581.

Table 2. Rotated Factor Matrix Comprising Three Components Generated in PCA (*n* = 282)

Item	Affective dimension	Behavioral dimension	Perseverance & confidence	Communality (h^2)
21	.886			.614
23	.757			.611
17	.749			.557
6	.629			.402
13	.608			.541
29	.599			.509
18	.579			.449
14	.568			.575
8	.563			.508
3	.534			.283
7	.472			.512
12	.459			.571
26	.566			.601
15		.777		.576
9		.749		.546
2		.746		.564
11		.744		.553
16		.506		.512
5		.482		.463
10		.481		.467
24			.737	.480
28			.627	.501
20			.512	.368
25			.452	.316
22			.378	.498
Cronbach's Alpha	.829	.858	.581	NA
Eigenvalues	8.080	1.896	1.397	NA
% Variance explained	32.32%	7.58%	5.59%	NA

*Note. Items 1, 4, 19, and 27 due to their cross-loading patterns in several models.

Confirmatory Factor Analysis (CFA) with AMOS

CFA was performed on the second subset (*n* = 283) to confirm the optimal PCA model across the second subset. The fit statistics of the measurement and full CFA models are presented in Table 3. The measurement models including the affective dimension, behavioral dimension, and perseverance and confidence dimension have good fit to the data. For example, the affective dimension's fit statistics are, as follows: $\chi^2 = 165.14$; $\chi^2/df = 2.54$; NNFI = 0.945; CFI = 0.960; and RMSEA = 0.052.

Table 3. Fit Statistics of Measurement and Three-Dimensional CFA Models (n = 283)

Model	χ^2	df	χ^2/df	NNFI	CFI	RMSEA	RMSEA boundaries
Affective dimension	165.14*	65	2.54	0.945	0.960	0.052	0.043–0.062
Behavioral dimension	13.612	14	.972	1.000	1.000	0.000	0.000–0.040
Perseverance & confidence dimension	306.65	15	.553	1.007	1.023	0.000	0.000–0.042
Unidimensional CFA model	994.11*	275	3.61	0.840	0.865	0.068	0.064–0.073
Three-dimensional CFA model	749.99*	272	2.75	0.912	0.910	0.056	0.051–0.061

*Note. * $p < 0.05$. n = 281.

Next, a unidimensional model was tested where all items were regressed on a general attitude dimension. The model had a poor fit to the data (χ^2 = 994.11; χ^2/df = 3.61; NNFI = 0.840; CFI = 0.865; and RMSEA = 0.068) and accordingly the fit of a three-dimensional CFA model composed of separate affective, behavioral, and perseverance and confidence dimensions was assessed, which fitted the data well (χ^2 = 749.99; χ^2/df = 2.75; NNFI = 0.912; CFI = 0.910; and RMSEA = 0.056). The χ^2 value of this model is significant at $p < 0.05$, which can be attributed to the sample size as well the complexity of the model as indicated by the degrees of freedom (df = 272).

Table 4 displays standard and non-standard regression estimates, standard error of measurement, CRs, and their p-values of the three-dimensional CFA model. All items have significantly high regression estimates as testified by their CRs and p-values. For example, Item 3 measures the affective dimension with a non-standard regression estimate of 0.750 ($p < 0.001$; CR = 9.738), suggesting that if participants' affective dimension increases by one unit, their performance on that item will increase by 0.75 units.

Table 4. Item Statistics of the Three-Dimensional CFA (*n* = 283)

Item	Dimension	Non-standardized estimate	Standardized estimate	SE	CR	*p*-value
3	Affective	.750	.470	.077	9.738	***
6	Affective	.986	.531	.091	10.797	***
7	Affective	1.067	.597	.090	11.798	***
8	Affective	.937	.587	.080	11.670	***
12	Affective	1.257	.719	.093	13.529	***
13	Affective	1.219	.699	.092	13.284	***
14	Affective	1.178	.742	.085	13.817	***
17	Affective	1.225	.667	.095	12.830	***
18	Affective	1.086	.587	.093	11.653	***
21	Affective	1.000	.597	Constrained for parameter estimation		
23	Affective	1.176	.691	.089	13.147	***
26	Affective	1.018	.582	.088	11.593	***
29	Affective	1.203	.665	.094	12.826	***
2	Behavioral	.956	.667	.066	14.495	***
5	Behavioral	.972	.654	.069	14.167	***
9	Behavioral	.920	.660	.064	14.307	***
10	Behavioral	.989	.666	.069	14.404	***
11	Behavioral	.987	.684	.067	14.822	***
15	Behavioral	1.000	.699	Constrained for parameter estimation		
16	Behavioral	.942	.691	.063	14.961	***
20	PersConf	1.287	.463	.192	6.702	***
22	PersConf	2.143	.684	.280	7.646	***
24	PersConf	1.000	.351	Constrained for parameter estimation		
25	PersConf	1.380	.468	.205	6.726	***
28	PersConf	1.488	.550	.208	7.160	***

*Note. PersConf = Perseverance & confidence dimension. CR = Critical ratio. SE = standard error of measurement. *** $p < 00.1$.

Finally, Figure 1 displays the three-dimensional CFA model with correlation coefficients (bidirectional arrows connecting the larger circles on the left side of the model) between the three dimensions, which indicate they are highly related. The figure also displays standardized regression coefficients which are just above the one headed arrows moving from dimensions (large circles) to items, represented by rectangles. Errors of measurement are represented as small circles on the right side of the figure.

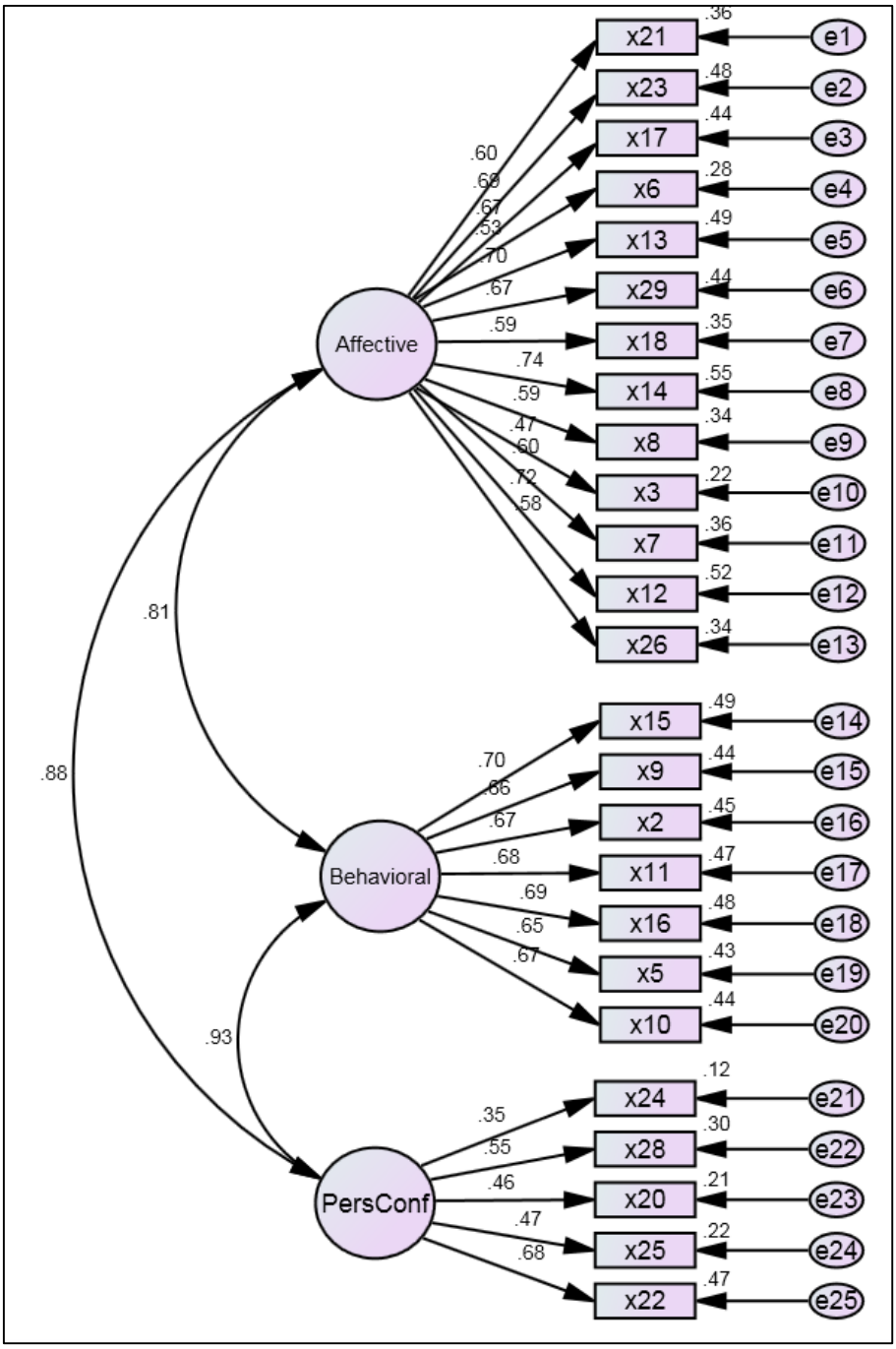

Figure 1. Graphic representation of the three-dimensional CFA model. (PersConf = perseverance and confidence dimension)

Rasch-Andrich Rating Scale Model (RSM)

Affective Dimension

The RSM was initially applied to the affective dimension consisting of 13 items. Person ability (or trait levels) ranged between -4.27 and 2.34 logits (M = -0.96; SD = 1.24) and item endorsability measures ranged between -1.46 and 0.74 logits (M = 0.00; SD = 0.52). Figure 2 presents an item person map with Andrich thresholds, which is a more precise map for polytomous data than a map without the thresholds. The left column represents respondents and the right column demonstrates the items located at their Andrich thresholds—points above which respondents begin to have a higher probability to choose higher categories. For example, Item 26's (x26) first threshold represented as 0.2 is the lowest of all and any respondent who has a higher ability (affective dimension measure) would have a higher probability to choose scoring category 3 on the item, whereas respondents below that point would most likely choose category 1. There is a gap between 0 and -1 logits, although there are not many respondents with trait levels corresponding to it and there are sufficient thresholds distinguishing respondents on the map.

Rasch model item and person reliability indices for this dimension were .98 (separation = 6.86) and .80 (separation = 1.94), respectively, indicating that if the items are administered to another group with similar features from the same population, there is high confidence that item and respondent measures (or locations) are reproducible (Linacre, 2013b).

Table 5 presents fit statistics, item endorsability measures, and raw scores for affective dimension items. The most lowly endorsed item (most difficult to endorse) is Item 13 (measure = 0.74), and most highly endorsed item is Item 12 (measure = -1.46). Infit MNSQ values range between 0.749 (Item 23) and 1.073 (Item 6) and outfit MNSQ values range between 0.719 (Item 23) and 1.080 (Item 6). This provides evidence that the data has not been contaminated by construct-irrelevant variance and high and low performers responded to the items as expected by the RSM.

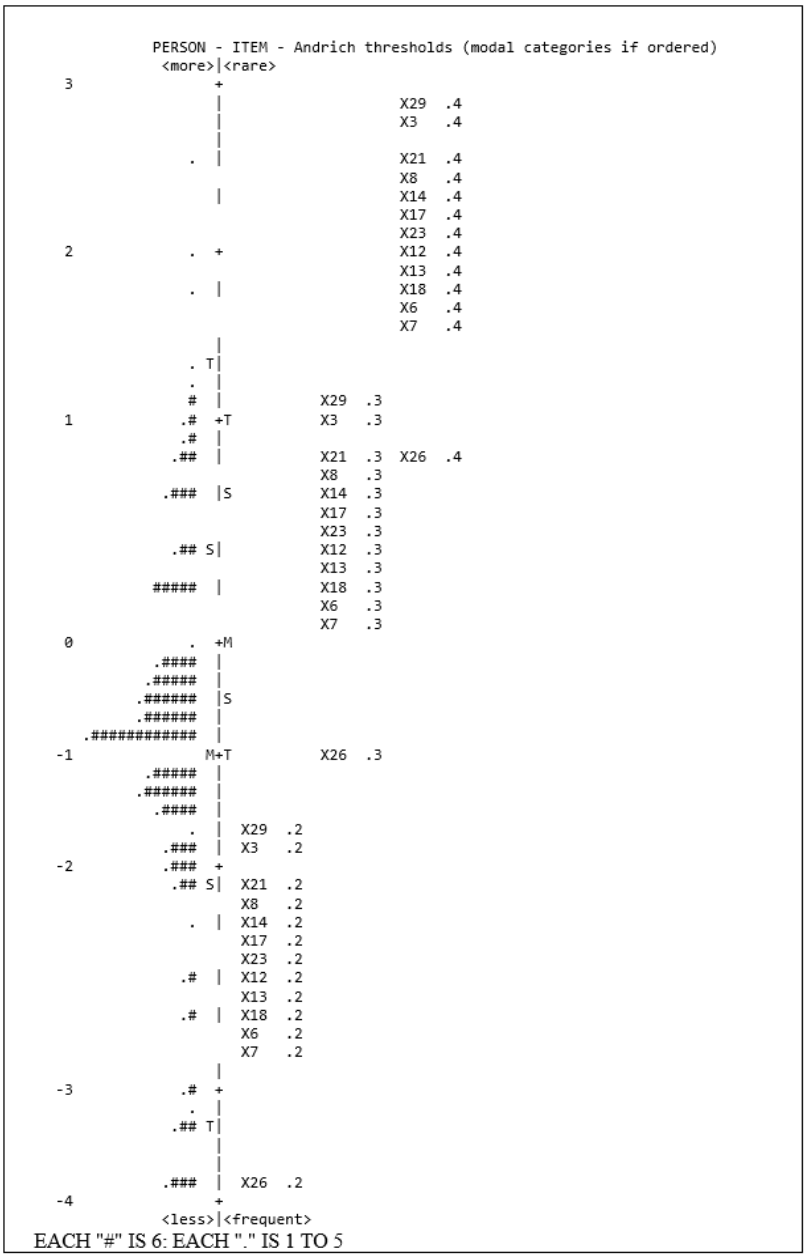

Figure 2. The affective dimension's item person map with Andrich thresholds. The map is truncated at the bottom.

Table 5. Raw Scores, Fit Statistics, and Item Endorsability Measures of Affective Dimension Items

Item	Rasch model measure	Raw score	Infit MNSQ	Outfit MNSQ
3	0.59	1043	1.012	1.026
6	-0.19	1204	1.073	1.080
7	-0.25	1206	0.897	0.905
8	0.31	1099	0.787	0.766
12	-0.01	1158	0.971	0.951
13	-0.07	1174	0.765	0.735
14	0.07	1137	0.978	0.861
17	0.1	1138	0.895	0.858
18	-0.27	1209	1.003	1.019
21	0.33	1094	0.834	0.786
23	0.11	1129	0.749	0.719
26	-1.46	1492	1.041	1.078
29	0.74	1021	1.032	0.965

Next, PCAR was performed to further explore the dimensionality of the test. The analysis showed that the raw variance explained by items and persons was 38.2% (7.4 eigenvalue units), whereas the first dimension in residuals explained 8.6% (1.8 eigenvalue units), indicating that the structure in residuals has no substance, as it accounts for a tiny proportion of the observed variance.

Finally, the response category features, thresholds, and category measures of the affective dimension were examined. Thresholds in this dimension increase monotonically (threshold 1 = -2.50; threshold 2 = 0.30; threshold 3 = 2.20). In addition, each response category has at least 10 observed counts and acceptable fit MNSQ indices (infit MNSQ: category 1 = 0.89; category 2 = 0.90; category 3 = 0.85; category 4 = 1.45), indicating that participants selected categories as expected by the model.

Behavioral Dimension

The item and person reliability indices in this dimension were. 99 (separation = 1.47) and .69 (separation = 10.26), respectively. The item person map with Andrich thresholds is also displayed in Figure 3. This figure shows a noticeable gap around the mean score. Otherwise, the spread of item thresholds corresponds to respondents' trait levels, though adding one or two items to fill in the gap would enhance the reliability of the dimension.

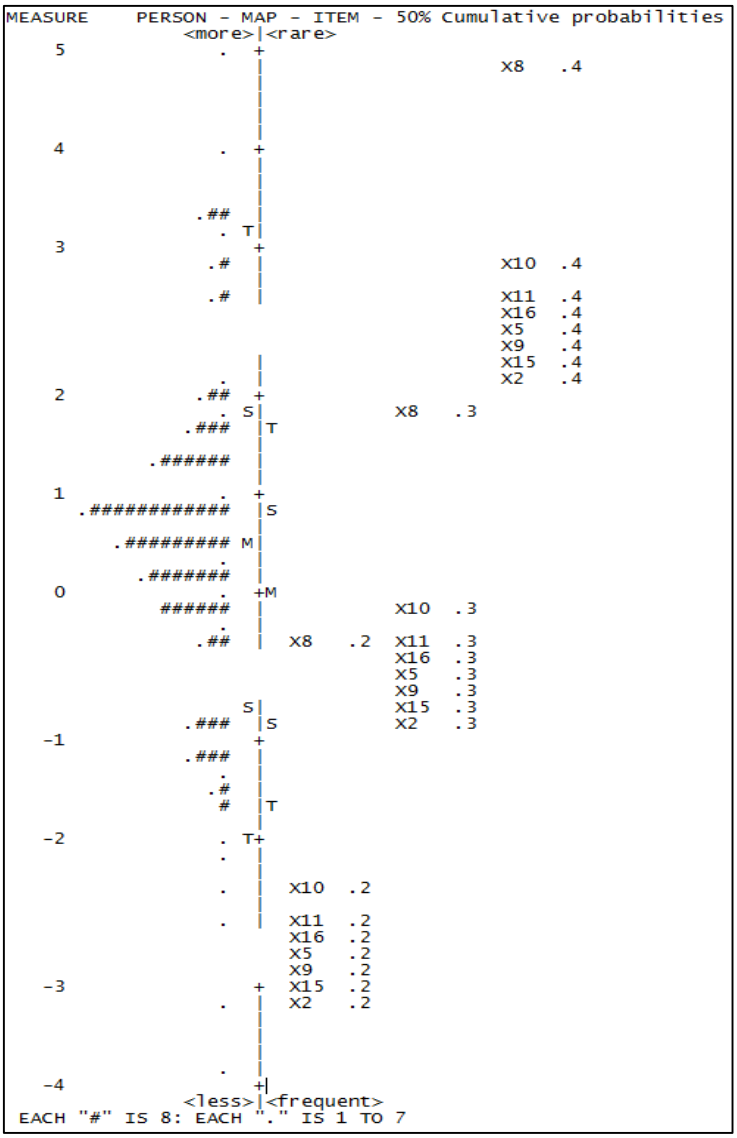

Figure 3. The behavioral dimension's item person map with Andrich thresholds. The map is truncated at the bottom.

Table 6 gives fit statistics, item endorsability measures, and raw scores for behavioral dimension items. Item measures (M = 0.00; SD = 0.80) range between -0.67 (Item 2) and 2.06 (Item 8) and person measures (M = 0.09; SD = 0.63) range between -5.04 and 3.27. Infit and outfit MNSQ statistics range between 0.70 and 1.35, suggesting that the items all fit the model, and that the data has not been contaminated by perturbations.

Table 6. Raw Scores, Fit Statistics, and Item Endorsability Measures of Behavioral Dimension Items

Item	Rasch model measure	Raw score	Infit MNSQ	Outfit MNSQ
2	-0.67	1663	0.864	0.827
5	-0.22	1561	0.904	0.895
9	-0.22	1566	0.954	0.930
8	2.06	1099	1.350	1.242
10	0.03	1511	0.816	0.999
11	-0.26	1580	0.982	0.964
15	-0.42	1609	0.884	0.856
16	-0.28	1585	0.717	0.687

Subsequently, PCAR was performed on the Rasch model residuals of this dimension to examine the presence of any substantive structure. The Rasch model item and person measures explained 44.5% of the observed variance, extracting 6.2 eigenvalues. The first contrast extracted only 1.5 eigenvalues, explaining 18.9% of the variance in data which is not a substantive amount (Linacre, 2013b). These results provide further evidence supporting the unidimensionality of this dimension.

Finally, the response category functions were examined. It was found a monotonic increment of the thresholds at -2.54, -0.23, and +2.77, with proper category infit MNSQ indices (1.25, 0.68, 0.91, 0.99). The fit statistics and ascending category measures provide evidence supporting the measurement features of response categories.

Perseverance and Confidence Dimension

Respective item and person reliability indices are 0.98 (separation = 6.25) and 0.41 (separation = 1.31). The rather low person reliability is likely due to the small number of items.

Figure 4 presents perseverance and confidence dimension's item person map with Andrich thresholds. There are two rather small gaps in the map, although there are not many respondents with trait levels corresponding to the gaps and there are sufficient thresholds distinguishing respondents.

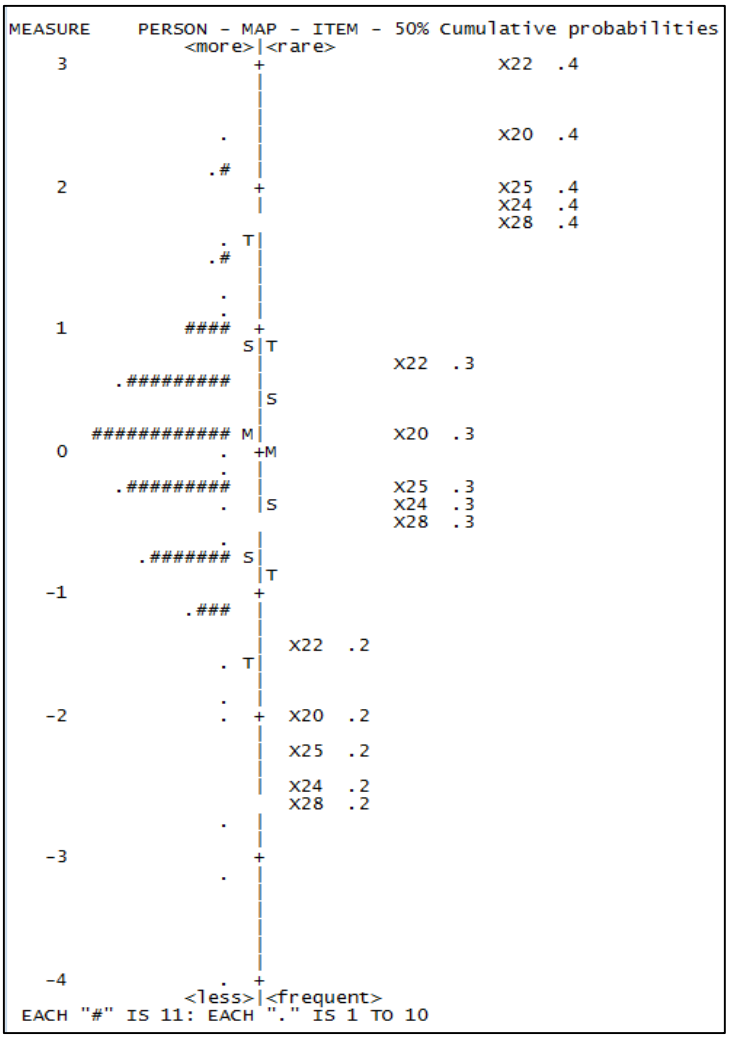

Figure 4. The perseverance and confidence dimension's item person map with Andrich thresholds.

Table 7 fit statistics, item endorsability measures, and raw scores for perseverance and confidence dimension items. The most difficult item to endorse is Item 22 (measure = 0.75) and the easiest is Item 28 (measure = -0.39). Infit and outfit MNSQ statistics fall between 0.7 to 1.20, suggesting that the items were not contaminated by construct-irrelevant factors.

Table 7. Raw Scores, Fit Statistics, and Item Endorsability Measures of Perseverance and Confidence Dimension Items

Item	Rasch model measure	Raw score	Infit MNSQ	Outfit MNSQ
20	0.20	1361	0.768	0.773
22	0.75	1228	1.106	1.151
24	-0.37	1500	0.819	0.817
25	-0.18	1453	0.890	0.904
28	-0.39	1506	0.711	0.713

Next, the dimension's residuals were examined through PCAR. Item and person measures explain 27% of the variance or 4.4 eigenvalue units, whereas 14.7% of the variance in the data or 1.00 eigenvalue units is accounted for by the first contrast. This further supports the unidimensionality of the perseverance and confidence dimension.

Finally, the functionality of response categories was examined. Like the other two dimensions, the thresholds had a monotonic increase: -2.03, -0.08, and 2.10. Infit and outfit MNSQ of the response categories fell between 0.80 and 1.20. This analysis provides support for the functionality of response categories.

Discussion

The present study was designed to develop and build a validity argument for P-BAQ. In this section, the findings and their contribution to the theory of attitude are discussed and are subsequently developed into a validity argument for the instrument.

P-BAQ's Dimensions

The data analysis procedures provided evidence supporting the presence of three dimensions underlying P-BAQ: affective, behavioral, and perseverance and confidence dimensions. Consistent with previous research, the present study yields support for the multidimensionality of attitude (e.g., Eagly & Chaiken, 1993; Hogg & Vaughan, 2011; McLeod, 2009; Mantle-Bromley, 1995). The results also lend partial support to the attitude's basic model, which comprises affective, behavioral, and cognitive dimensions (McLeod, 2009).

However, unlike the basic model, the study did not yield a cognitive dimension. This is in line with Kay (1993), Mantle-Bromley (1995), Min (1998),

Selwyn (1997), and Tsai et al. (2001) who conceptualize attitude as a set of "evaluative" responses to individuals, objects, or opinions but distinguish it from a cognitive dimension measuring individuals' beliefs and perceptions—although cognitive capacity is a precondition to forming and holding attitudes. Relatedly, Ajzen (2005, 2011) argues that there is considerable commonality between cognitive and affective dimensions, and that is why distinguishing the two components might be impossible. The partial disagreement over the constituent structure of attitude may also be attributed to the lack of clarity surrounding attitude's definition, as there is still no universally endorsed definition for attitude (Franzoi, 2003; Smith et al., 2000).

Findings further show that attitude toward blogs is highly similar to attitude toward computers, internet, and web-based technologies such as social networking sites (e.g., Fančovičová & Prokop, 2008; Liaw, 2002; Sam et al., 2008). However, the emergence of the perseverance and confidence dimension is new in this study. If attitude is conceptualized as a continuum with affective and behavioral ends, the perseverance and confidence dimension would fall between the two ends. It might be said that the individuals who score low on this dimension would likely score low on both affective and behavioral dimensions and vice versa; that is, performance on this dimension might predict students' performance on the other two dimensions. Future research can examine this possibility by using causal statistical models.

Our findings undermine Daud's (1995) and Smith's (1996) claims of attitude's unidimensionality. The results of PCA showed three separate dimensions, and CFA model testing confirmed that a unidimensional model would not fit the data. Although the correlation coefficients between the three dimensions were high in the CFA model, high correlations may not be confused with identicality of dimensions (Borsboom, 2008); two independent concepts such as age and height might be highly correlated yet distinct in definition and nature. It should be noted that dimensions are *not* psychological traits but a rank-ordering of people on the traits. Aryadoust (2013, p. 195) stated:

> Using the example of two "climbing" [dimensions], "holding onto ledges" and "pulling oneself up from ledges," the correlation between two factors represents how well individual differences on one sub-skill (holding onto ledges) *linearly* predicts individual differences on another (pulling oneself up from ledges). Whether that correlation is 0, 0.5, or 1 is immaterial to the question of whether or not the [dimensions] exist. It only addresses the

question whether individual differences on one predict individual differences on another.

Despite the high psychometric quality of the three dimensions, it would be useful to add a few items to each dimension to improve person reliability statistics. Overall, the multidimensional structure of attitude toward blogs is supported in the present study. A validity argument for P-BAQ will be developed in the following section using the evidence yielded at different stages of the study.

Validity Argument of P-BAQ

A validity argument provides a coherent treatment of the evidence gathered in the validation study. P-BAQ's validity argument is supported by the findings of the previously stated psychometric studies and is presented in Table 8. For example, the domain inference is based on the claim concerning the multidimensionality of P-BAQ. The claim is based on the assumptions that dimensions of P-BAQ can be determined and measured. It is supported by the postulations that "Attitude toward blogs is composed of multiple components which were identified through the survey of literature in this study" and "The pool of items adapted from previous attitude research contained a wide range of items tapping attitude's dimensions." The presented literature survey shows that the commonly adhered-to attitude construct is multidimensional (e.g., Selwyn, 1997; Tsai et al., 2001) and supports the representativeness of the item pool. In addition, items chosen for translation by experts would engage at least two dimensions of attitudes: affective and behavioral. However, Table 8 presents a potential counterargument: some researchers provided empirical evidence supporting attitude's unidimensionality. The counterargument is rebutted by the evidence from PCA and CFA, thereby supporting the domain definition inference and the theoretical representation of P-BAQ.

The remainder of the inferences are also supported by the available evidence. However, the generalization inference is undermined by low person reliability and separation statistics of the three dimensions which result from the gaps in the item person map. This would indicate a requirement to add a few items to P-BAQ to enhance the generalizability of individuals' attitude scores.

Table 8. Validity argument of the P-BAQ

Inference	Claim	Assumption	Postulations (warrants)	Backing
Domain definition	Iranian tertiary students' attitude toward blogs is a multidimensional construct including at least two dimensions: affective and behavioral.	(i) The components (dimensions) of attitude toward blogs can be determined; and (ii) the components are measurable.	(a) Attitude toward blogs is composed of multiple components which were identified through the survey of literature in this study; and (b) the pool of items adapted from previous attitude research contained a wide range of items tapping attitude's dimensions.	The presented literature survey shed light on the theory of attitude toward technology and supports the thoroughness of the item pool created. The literature also supported multidimensionality of attitude although a handful of studies have yielded unidimensional attitude scales. Finally, items which are chosen for translation by experts would engage at least two dimensions of attitudes: affective and behavioral.
Translation	The chosen items are clearly translated into Persian.	(i) The items are translatable; and (ii) the translated items are easy to understand and contain no ambiguous terms or phrases for Persian speakers.	Translation would have no adverse effect on the construct equivalence of the Persian instrument if it is approved by experts.	The approval of translation specialists of the clarity of the translated items and their lack of linguistic ambiguities was obtained.
Evaluation	Participants consistently choose their responses, producing reliable data.	Scoring categories (Likert scale) are appropriate and precise and thus can be established through psychometric and statistical analysis.	(a) Four-category Likert scales would prevent respondents from "flat-lining" the scale which happens when they endorse the mid-points on the scale; and (b) Psychometric attributes of the Likert scale are backed by psychometric analysis.	(1) Monotonic increment of Likert scales' response categories and thresholds was established; and (2) appropriate fit statistics for each response category were supported by RSM.

Table 8 continued on the next page

Generalization	The observed scores can be generalized to the universe of scores.	The performance on the survey is considered representative of performance on the universe of items.	Generalization is psychometrically measured.	High Rasch measurement scores on the item reliability and separation indices support that the observed measures can be generalized to the universe of measures. Undermining evidence: In contrast, low person reliability and separation indices on the three dimensions suggest that additional items should be added to the P-BAQ.
Explanation	A link is established between the observed scores and the underlying construct.	(i) Construct representativeness of P-BAQ can be examined; (ii) the correlation of the dimensions can be examined; and (iii) lack of construct-irrelevant factors can be investigated.	(a) The constituent structure of P-BAQ is supported by theory; (b) the sub-components (dimensions) of P-BAQ are distinguishable and related; and (c) P-BAQ has reliable psychometric features and dimensionality analysis can show that its components are not contaminated.	The results of psychometric analysis: PCA supported the multidimensional structure of the instrument; CFA verified the structure across a different sample, supporting the psychometric distinction of the sub-components; and RSM results lent themselves to the unidimensionality of the dimensions of P-BAQ, disproving the presence of construct-irrelevant factors. However, the item person map in the RSM analysis contained a few gaps, which should be filled in by further items in the future.

Applications of P-BAQ

Turning to the pedagogical overtones of the study, research shows that learners' progress largely depends on their motivation, which is a "composite of attitudinal variables" (Mantle-Bromley & Miller, 1991, p. 418). Positive attitudes toward learning tools such as blogs and classes can improve students' motivation. Measuring students' attitude toward blogs will give an indication of the potential causes of students' lack of motivation and help teachers take proper measures to motivate the students.

To our knowledge, P-BAQ is the first Persian blog attitude instrument that has a (strong) validity argument. Because it is a valid psychometric tool, the P-BAQ is suitable for use in educational programs that make use of blogs. It is also of potential use for research conducted in Afghanistan, Iran, Tajikistan, and other countries where Persian or Farsi are spoken.

Conducing continuous research into the attitude of learners who use blogs in low-resource countries is of paramount importance. It might be said that a lack of facilities and financial aid coupled with cultural factors and unfamiliarity with blogs (and other technology advances) can affect students' attitudes and motivations, and consequently their learning achievements.

Conclusion

A validity argument for the P-BAQ was constructed by examining claims and assumptions, proposing postulations, and offering backing (evidence) supporting these postulations. Overall, the validity argument is strong, though the structure of the instrument could be improved through future research. Validation of psychometric instruments is never completed; it is extended and "refined as new research and data analysis technologies are developed" (Aryadoust, 2013, p. 223). Future research can address other issues such as bias or differential item functioning which are related to the explanation inference.

Articulating the claims concerning the psychometric quality of instruments as well as uses and interpretations of scores would benefit validation research. However, research into measuring attitude has hardly sought to construct and assess validity arguments. To our knowledge, the only study attempting to construct a validity argument for a survey instrument was done by Bangert (2009). It is hoped that validity arguments will be applied more often in future attitude research and other educational studies.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*(4), 561-573. doi:10.1007/BF02293814

Al-Tamimi, A., & Shuib, M. (2009). Motivation and attitudes towards learning English: A study of petroleum engineering undergraduates at Hadhramout University of Sciences and Technology. *GEMA Online Journal of Language Studies*, *9*(2), 29-53.

Ajzen, I. (2005). *Attitudes, personality, and behaviour* (2nd ed.). New York: Open University Press.

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, *26*(9), 1113-1127. doi:10.1080/08870446.2011.613995

Aryadoust, V. (2013). *Building a validity argument for a listening test of academic proficiency*. Newcastle: Cambridge Scholars Publishing.

Aryadoust, V., Mehran, P., & Alizadeh, M. (2014). Validating a computer-assisted language learning attitude instrument used in Iranian EFL context: An evidence-based approach. *Computer Assisted Language Learning Journal*-First Online. doi:10.1080/09588221.2014.1000931

Ayres, R. (2002). Learner attitudes towards the use of CALL. *Computer Assisted Language Learning*, *15*(3), 241-249. doi:10.1076/call.15.3.241.8189

Bangert, A. W. (2009). Building a validity argument for the community of inquiry survey instrument. *Internet and Higher Education*, *12*(2), 104-111. doi:10.1016/j.iheduc.2009.06.001

Blackstone, B., Spiri, J., & Naganuma, N. (2007). Blogs in English language teaching and learning: Pedagogical uses and student responses. *Reflections on English Language Teaching*, *6*(2), 1-20.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences.* London: Lawrence Erlbaum Associates.

Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research and Perspectives*, *6*(1/2), 25-53. doi:10.1080/15366360802035497

Brodersen, J., Thorsen, H., & Kreiner, S. (2007). Validation of a condition-specific measure for women having an abnormal screening mammography. *Value in Health*, *10*(4), 294-304. doi:10.1111/j.1524-4733.2007.00184.x

Chen, W. C., Shih, Y., & Liu, G. Z. (2013). Learning task design and its induced learning effects in a cross-institutional blog-mediated telecollaboration.

Computer Assisted Language Learning, *28*(4), 285-305. doi:10.1080/09588221.2013.818557

Chu, S.-C., & Kamal, S. (2008). The effect of perceived blogger credibility and argument quality on message elaboration and brand attitudes: An exploratory study. *Journal Interactive Advertising, 8*(2), 32-33. doi:10.1080/15252019.2008.10722140

Daud, N. M. (1995). A computer attitude scale for language teachers. *Computer Assisted Language Learning, 8*(4), 355-363. doi:10.1080/0958822950080404

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, London, New Delhi: Sage Publications.

Dimroth, C. (2010). The acquisition of negation. In L. R. Horn (Ed.), *The expression of negation* (pp. 39-73). Berlin/New York: Mouton de Gruyter.

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York: Routledge.

Duda, G., & Garrett, K. (2008). Blogging in the physics classroom: A research-based approach to shaping students' attitudes toward physics. *American Journal of Physics, 76*(11), 1054-1065. doi:10.1119/1.2967707

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Fageeh, A. I. (2011). EFL learners' use of blogging for developing writing skills and enhancing attitudes towards English Learning: An exploratory study, *Journal of Language and Literature*, *2*(1), 2078-0303.

Fančovičová, J., & Prokop, P. (2008). Students' attitudes toward computer use in Slovakia. *Eurasia Journal of Mathematics, Science & Technology Education*, *4*(3), 255-262.

Farahzad, F. (1992). Testing achievement in translation classes. In C. Dollerup & A. Loddergard (Eds.), *Teaching translation and interpreting: Training, talent and experience* (pp. 271-278). Amsterdam/Philadelphia: John Benjamins.

Franzoi, S. (2003). *Social Psychology*. Boston: McGraw-Hill.

Gajewski, J. R. (2007). Neg-raising and polarity. *Linguistics and Philosophy, 30*(3), 289-328. doi:10.1007/s10988-007-9020-z

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). New Jersey: Pearson Educational Product.

Harkness, J. A. (2008). Comparative survey research: Goals and challenges. In E. D. De Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 56-77). New York, NY: Lawrence Erlbaum.

Harwood, C. (2010). Using blogs to practice grammar editing skills. *English Language Teaching World Online*, 2. Retrieved from http://blog.nus.edu.sg/eltwo/2010/08/11/using-blogs-to-practice-grammar-editing-skills/#more-1107

Hogg, M. A., & Vaughan, G. M. (2011). *Social psychology* (6th ed.). Harlow: Pearson.

Huang, H. C. (2013). From web-based readers to voice bloggers: EFL learners' perspectives. *Computer Assisted Language Learning*, *28*(2), 145-170. doi:10.1080/09588221.2013.803983

IBM Corp. (2012). IBM SPSS Statistics for Windows, Version 21.0 [Computer Software]. Armonk, NY: IBM Corp.

Jones, T., & Clarke, V. A. (1994). A computer attitude scale for secondary students. *Computers & Education*, *22*(4), 315-318. doi:10.1016/0360-1315(94)90053-1

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (pp. 17-64). Westport: American Council on Education/Praeger.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1-73. doi:10.1111/jedm.12000

Kay, R. H. (1993). An exploration of theoretical and practical foundations for assessing attitudes towards computers: The Computer Attitude Measure (CAM). *Computers in Human Behavior*, *9*(4), 371-386. doi:10.1016/0747-5632(93)90029-R

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kline, R. (2006). Reverse arrow dynamics: Formative measurement and feedback loops. In G. R. Hancok & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 43-68). Greenwich, CT: Information Age Publishing.

Liaw, S.-S. (2002). An Internet survey for perceptions of computer and the World Wide Web: Relationship, prediction, and difference. *Computers in Human Behaviour*, *18*(1), 17-35. doi:10.1016/S0747-5632(01)00032-2

Liaw S.-S., Huang, H. M., Chen, G. D. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education, 49*(4), 1066-1080.

Linacre, J. M. (2013a). WINSTEPS Rasch measurement [computer program]. Chicago: Winsteps.com.

Linacre, J. M. (2013b). *A user's guide to WINSTEPS*. Chicago: Winsteps.com.

Loyd, B. H., & Gressard, C. P. (1984). Reliability and factoral validity of computer attitude scale. *Educational and Psychological Measurement, 44*(2), 501-505. doi:10.1177/0013164484442033

Mantle-Bromley, C. (1995). Positive attitudes and realistic beliefs: Links to proficiency. *The Modern Language Journal*, *79*(3), 371-386. doi:10.1111/j.1540-4781.1995.tb01114.x

Mantle-Bromley, C., & Miller, R. B. (1991). Effect of multicultural lessons on attitudes of students of Spanish. *The Modern Language Journal*, *75*(4), 418-425. doi:10.1111/j.1540-4781.1991.tb05376.x

McIntosh, E. (2009). Blogging improves young people's confidence in their writing and reading. Retrieved from http://edu.blogs.com/edublogs/2009/12/blogging-improves-young-peoples confidence-writing-reading.html

McLeod, S. A. (2009). *Attitudes and behaviour*. Retrieved from http://www.simplypsychology.org/attitudes.html

Min, B. C. (1998). *A study of the attitudes of Korean adults toward technology assisted instruction in English language programs* (Unpublished doctoral dissertation). University of Northern Illinois, DeKalb.

Morse, B. J., Gullekson, N. L., Morris, S. A., & Popovich, P. M. (2011). The development of a general Internet attitudes scale. *Computers in Human Behavior*, *27*(1), 480-489. doi:10.1016/j.chb.2010.09.016

Pektas, S. T., & Erkip, F. (2006). Attitudes of Design Students toward Computer Usage in Design. *International Journal of Technology and Design Education, 16*(1), 79-95. doi:10.1007/s10798-005-3175-0

Pickens, J. (2005). *Attitudes and perceptions.* Retrieved from http://healthadmin.jbpub.com/borkowski/chapter3.pdf

Sam, H. K., Othman, A. E. A., & Nordin, Z. S. (2005). Computer self-efficacy, computer anxiety, and attitudes toward the Internet: A study among undergraduates in Unimas. *Journal of Educational Technology & Society*, *8*(4), 205-219.

Selwyn, N. (1997). Students' attitudes toward computers: Validation of a computer attitude scale for 16–19 education. *Computers & Education, 28*(1), 35-41. doi:10.1016/S0360-1315(96)00035-8

Shaft, T. M., Sharfman, M. P., & Wu, W. W. (2004). Reliability assessment of the attitude towards computers instrument (ATCI). *Computers in Human Behaviour, 20*(5), 661-689. doi:10.1016/j.chb.2003.10.021

Shahsavar, Z., & Tan, B. H. (2011). Does cognitive style affect bloggers' attitude in an online learning environment? *GEMA Online TM Journal of Language Studies, 11*(1), 159-171.

Shahsavar, Z., & Tan, B. H. (2012). Developing a questionnaire to measure students' attitudes toward the course blog. *Turkish Online Journal of Distance Education, 13*(1), 200-210.

Smith, R. M. (1996). A comparison of methods for determining dimensionality in Rasch measurement. *Structural Equation Modelling, 3*(1), 25-40. doi:10.1080/10705519609540027

Smith, B., Caputi, P., & Rawstorne, P. (2000). Differentiating computer experience and attitudes toward computers: An empirical investigation. *Computers in Human Behavior, 16*(1), 59-81. doi:10.1016/S0747-5632(99)00052-7

Teo, T. (2012). Measurement invariance of self-reported data across gender: A study of technology acceptance. *International Journal of Educational and Psychological Assessment, 11*(1), 43-54.

Tsai, C.-C., Lin, S. S. J., & Tsai, M.-J. (2001). Developing an Internet attitude scale for high school students. *Computers & Education, 37*(1), 41-51. doi:10.1016/S0360-1315(01)00033-1

Tu, C.-C., Chen, P.-C., & Lee, M.-Y. (2007). Fostering EFL learners' writing competence through web-Based guided writing. *WHAMPOA—An Interdisciplinary Journal, 53*, 225-244.

Vandewaetere, M., & Desmet, P. (2009). Introducing psychometrical validation of questionnaires in CALL research: The case of measuring attitude towards CALL. *Computer Assisted Language Learning, 22*(4), 349-380. doi:10.1080/09588220903186547

Wang, Q., & Woo, H. L. (2010). Investigating students' critical thinking in weblogs: An exploratory study in a Singapore secondary school. *Asia Pacific Education Review, 11*(4), 541-551. doi:10.1007/s12564-010-9101-5

Woo, H. L., & Wang, Q. (2009). Using weblog to promote critical thinking: An exploratory study. *World Academy of Science, Engineering and Technology, International Science Index 25, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 3*(1), 48-57.

Wright, B. D. (1994a). Comparing factor analysis and Rasch measurement. *Rasch Measurement Transactions, 8*(1), 350.

Wright, B. D. (1994b). Local dependency, correlations, and principal components. *Rasch Measurement Transactions, 10*(3), 509-511.

Zawilinski, L. (2009). HOT blogging: A framework for blogging to promote higher order thinking. *The Reading Teacher*, *62*(8), 650-661. doi:10.1598/RT.62.8.3

Zhang, Y. (2007). Development and validation of an Internet use attitude scale. *Computers & Education*, *49*(2), 243-253. doi:10.1016/j.compedu.2005.05.005

Zhao, Y., & Frank, K. A. (2003). Factors affecting technology uses in school: An ecological perspective. *American Educational Research Journal, 40*(4), 807-840. doi:10.3102/00028312040004807

Zimbardo, P. G., & Leippe, M. R. (1991). *The psychology of attitude change and social influence*. New York: McGraw-Hill.

Appendix

P-BAQ's items

		کاملاً مخالف	مخالف	موافق	کاملاً موافق
۱	کار کردن با وبلاگ مرا مضطرب نمی کند.				
۲	کار کردن با وبلاگ را دوست دارم.				
۳	کار کردن با وبلاگ مراعصبی می کند.				
۴	وقتی کسی در مورد وبلاگ صحبتی به میان می آورد احساس نگرانی نمی کنم.				
۵	انتخاب درسهایی که در رابطه با وبلاگ است، برای من خوشایند هست.				
۶	با وبلاگ نمی توانم خوب کار کنم.				
۷	تبادل اطلاعات از طریق وبلاگ برایم خوشایند نیست.				
۸	از کار کردن با وبلاگ احساس ناراحتی می کنم.				
۹	معمولاً وقتی برای تبادل اطلاعات از وبلاگ استفاده می کنم، احساس رضایت دارم.				
۱۰	در کلاسهایی که از وبلاگ استفاده می شود، احساس راحتی می کنم.				
۱۱	فکر می کنم کار کردن با وبلاگ خیلی لذت بخش و جالب باشد.				
۱۲	فکر نمی کنم از کار کردن با وبلاگ لذت ببرم.				
۱۳	تبادل نظر با دیگران از طریق وبلاگ برای من خوشایند نیست.				
۱۴	وقتی به کار کردن با وبلاگ فکر می کنم، احساس خوبی ندارم.				
۱۵	مطمئن هستم که وبلاگ در انجام دادن کارهایم مفید خواهد بود.				
۱۶	از کار کردن با وبلاگ احساس رضایت میکنم.				
17	من برای کارکردن با وبلاگ ساخته نشده ام.				
18	برای من قابل توجیه نیست که چگونه بعضی از افراد زمان زیادی را صرف کارکردن با وبلاگ میکنند و از کارکردن با آن لذت می برند.				
19	مطمئنم که می توانم کارکردن با وبلاگ را یاد بگیرم.				
20	وقتی که شروع به کارکردن با وبلاگ می کنم ، دوست ندارم از آن دست بکشم.				
21	تصور می کنم که کارکردن با وبلاگ برای من دشوار باشد.				
22	درمجموع تمایل چندانی به استفاده از وبلاگ ندارم.				
23	از کار کردن با وبلاگ احساس سردرگمی و نارضایتی می کنم.				
24	اگر مسئله ای در حین استفاده از وبلاگ در کلاس حل نشده باقی بماند ، بعد از کلاس به آن فکر میکنم.				
25	در درسهایی که مربوط به وبلاگ است، می توانم نمرات خوبی کسب کنم.				
26	از صحبت با دیگران در مورد وبلاگ لذت می برم.				
27	گذراندن درسهای مربوط به وبلاگ از عهده من خارج است.				
28	زمان کارکردن با وبلاگ از اعتماد به نفس بالایی برخوردارم.				
29	نسبت به وبلاگ احساس انزجار و تنفر می کنم.				

The English Version of the Blog Attitude Questionnaire

#	**Affective**	SD	D	A	SA
3	Keeping a blog would make me very nervous.				
6	I'm no good with blogs.				
7	The challenge of adding a new post on blogs and keeping them updated does not appeal to me.				
8	Blogs make me feel uncomfortable.				
12	I don't think I would enjoy doing advanced blog work (such as designing complex appearance).				
13	Figuring out any blog problem does not appeal to me.				
14	I get a sinking feeling when I think of trying to keep a blog.				
17	I'm not the type to do well with blogs.				
18	I don't understand how some people can spend so much time working with blogs and seem to enjoy it.				
21	I think using or keeping a blog would be very hard for me.				
23	Keeping blogs make me feel uneasy and confused.				
26	I don't enjoy talking with others about blogs.				
29	I feel aggressive and hostile toward blogs.				
	Behavioral				
2	I would like working with and keeping blogs.				
5	It wouldn't bother me at all to take courses on blogs.				
9	Generally, I would feel OK about adding a new post on the blog.				
10	I would feel at ease in a blog class.				
11	I think working with and keeping blogs would be enjoyable and stimulating.				
15	I am sure I could do work with blogs.				
16	I would feel comfortable working with a blog (e.g., keeping and updating them, referencing to other sources, etc.).				
	Perseverance and Confidence				
20	Once I start to work on a blog, I would find it hard to stop.				
22	I will do as little work through blogs as possible.				
24	If a problem with my blog is left unsolved, I would continue to think about it afterward.				
25	I could get good grades in blog courses if there are any.				
28	I have a lot of self-confidence when it comes to working with blogs.				