



## **Instruments for Measuring the Influence of Visual Persuasion: Validity and Reliability Tests**

**Nurulhuda Ibrahim<sup>1\*</sup>**

**Mohd Fairuz Shiratuddin<sup>2</sup>**

**Kok Wai Wong<sup>2</sup>**

<sup>1</sup>School of Multimedia Technology and Communication,  
Universiti Utara Malaysia

<sup>2</sup>School of Engineering and Information Technology,  
Murdoch University, Australia

\*Email: nurulhuda@uum.edu.my

DOI: 10.26417/ejser.v4i1.p25-37

### **Abstract**

In User Experience (UX) research, the instruments are often measured by means of rating scales such as Likert scale and semantic differential scale. The validity of the findings and conclusions rely heavily on the instruments used in the questionnaires. This paper provides the assessment of the validity and reliability of a new set of measures to evaluate the influence of visual persuasion on web users. The instruments will be used to assess web users' perceptions of credibility, engagement, informativeness, satisfaction, social influences, usability, and visual aesthetic. Firstly, 85 items are pilot tested by expert and novice users in an offline and online settings. Secondly, the exploratory factor analysis is carried out in which 44 items representing 12 latent variables are reduced to 39 items with some of the latent variables are combined into one. The results show: Kaiser-Meyer-Olkin (KMO) of 0.901, significant Bartlett's test, communalities range between 0.470 - 0.829, nine factors (also known as the latent variables) emerged with eigenvalues greater than 1, explaining more than 60% of the total variance, factor loadings of 0.466 and above, factors correlations of less than 0.7, and Cronbach's alphas are well above the limit of 0.70. Finally, a confirmatory factor analysis is carried out on the first-order and second-order latent variables using the PLS-

SEM. The instruments exceed the minimum requirement of the assessments for the convergent validity, discriminant validity, reliability and collinearity. The findings suggest that the proposed 39 items are valid and reliable for measuring the persuasiveness of visual persuasion.

**Keywords:** reliability test, validity test; visual persuasion measures, PLS-SEM

## Introduction

A number of researchers proposed that visual design on an interface have certain impact on users' motivation to stay longer on a website which consequently improves user experience (e.g. Hao, Tang, Yu, Li, & Law, 2015; Chu, Deng, & Chuang, 2014; Cyr, 2013; Horvath, 2011; Winn & Beck, 2002). User Experience (UX) is defined by three main characteristics; 1) involves a user, 2) the user is interacting with a product, system, or anything with a User Interface (UI), and 3) the user's experience is of interest, and observable or measurable (Albert & Tullis, 2013). A research on UX can be considered if behaviour or potential behaviour is expected to take place. Previous studies investigating UX have assessed variety of factors e.g. performance, usability, and satisfaction; which also includes specific factors e.g. credibility, engagement and visual aesthetic (Albert & Tullis, 2013).

This research examines the association between user's perception of web design characteristics and behavioural intention of persuasive visual design. The research extends the model of first impression formation toward tourism destination websites by Kim & Fesenmaier (2008). In the model, informativeness and usability are labelled as the hygiene (essential) factors, and persuasive features namely credibility, inspiration (visual appeal/aesthetic), involvement (or engagement), and reciprocity represent the motivating factors. The result of this study shows that only inspiration, usability, and credibility are significant in first impression formation. It is inspected that the cause of declining factors are resulted from their research design as they used an animated clip of webpage screenshots, and the participants of the study merely viewed the clips instead of browsing the identical webpage. Moreover, Kim & Fesenmaier (2008) work is specific for instant first impression formation with the display time span for the webpage was set to 7 seconds each. This research takes a different research design approach in which an actual web environment setting is used for the experiment. Furthermore, instead of using existing tourism websites, the web samples are personally developed to enable A/B testing method (see section 2.0 Research Design and Method). As such, investigation is more specific towards pointing out which visual persuasion has positive influence on users; thus reducing the potential of persuasion clutter. Besides, in this research, first impression is defined as the event when a user first encounters a new website, and forms a mental image of that website. Users are free to browse the website as long as they like to encourage actual surfing behaviour. In this research, users generally take about 2 minutes surfing the website (data is taken from Google Analytic).

The research conceptual model includes the web design characteristic's dimensions i.e. credibility, engagement, informativeness, satisfaction, usability, and visual aesthetic. The new model includes the social influence principles by Cialdini (2007) as an additional value to the motivating factors in the model. In the model, social influence is defined as a second-order variable; a latent variable whose indicators are themselves latent variables. The indicators of social influence are authority, commitment, liking, reciprocity, scarcity, and social proof (Cialdini, 2007). This research adapts existing measures from the web design evaluation metrics, and develops new measures for assessing the social influence specific. This is due to the fact that the existing measures of UX are suitable for a comprehensive evaluation of a UI. However, this research specifically measures users' perception of visual persuasion; focusing on the impact of certain visual (pictorial cues or textual messages) on users' motivation and behavioural intention. Moreover, users' perception is recorded based on the users' first impression of the visual design. As such, users may observe differently compared to the common web assessment.

The paper aims to evaluate the reliability and validity of the instruments used in the research by conducting the exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA explores and examines the inter-correlations that exist between a pool of items and in doing so, reduces the items into smaller groups, also known as the dimensions or factors (Hooper, 2012). On the other hand, CFA verifies the factor structure of a set of variables. The measures are tested during an online experimental design where participants are required to evaluate a tourism website.

## **Research Design and Method**

### **Participants**

Facebook users of the age 18 or older participate in the web survey. Participants are recruited through an advertisement on Facebook. Participants are also encouraged to invite their Facebook friends to participate in the research. Thus, the survey is non-representative, and relies heavily on volunteers who hear about it through Facebook's News Feed.

### **Procedures**

Once the users click on the URL to participate in the survey, they are taken to the survey website. Participants read the information and participating conditions, and give their consent. They need to fill up the demographic section first before being taken to the next page to evaluate the website. The survey employs an A/B test method, a common type of live-site study in which the researchers manipulate elements of the page that are presented to the users (Albert & Tullis, 2013). This method involves an online experiment in which the participants are required to evaluate a website, which is randomly assigned. In this experiment, the non-persuasive website represents the control design and the persuasive website represents the treatment design. The A/B test helps to identify which web sample increases favourable users' perception of the website and gives significant insight into

which visual elements 'work' and which 'do not work' on the website (Albert & Tullis, 2013). After browsing the website, the participants return to the survey to answer the questionnaire.

### **Instruments**

Initially 61 items of 12 latent/independent variables and 4 items of 1 observed/dependent variable go through a pilot test by 2 field experts, 5 novice users, and 3 postgraduate students. The objectives of the test are to check the content and to identify suitable items that are specific for evaluating visual interface. As a result, instruments related to measuring visual design are short listed and 20 new items are included. Then the instruments are checked and approved by Murdoch University Human Research Ethics Committee (Approval 2013/155). Another pilot test on the actual online environment is carried out to identify any problems with the questionnaire in terms of their clarity and to investigate the ability of potential respondent to understand the items. At the same time, any functionality issues and possible errors or bugs with the web samples can be identified. 30 individuals are invited via Facebook, but only 10 participants responded. Repeated measures approach is used in the pilot test and users evaluated both websites, in which the order is randomly assigned. Thus, the time period needed to complete the questionnaire is estimated. Each website is evaluated with the same set of questionnaire. Although each website design is different in terms of the visual design used, it is estimated that participants will be able to understand the items if they have more than one year of Internet and web experience. This is due to the reason that the social influence principles were long applied in the online marketing's websites. Comments or suggestions from the respondents are also recorded.

Modifications are made based on the feedback received from the pilot test. Several questions are deleted, or included some examples. To shorten the time needed for the survey completion, the random assignment approach is used instead of repeated measures. Thus, each participant evaluates only one website; adopting the method used by Tang (2009). In the end, 48 items remain in the list. UX is assessed by 24 items adapted from various authors. Another 6 latent variables representing the principles of social influence are measured using 2 adapted items and 18 newly developed items. The 44 psychometric items representing 12 latent variables are administered for the reliability and validity test with EFA. Another four indicators of 1 observed variable are validated with CFA as the items are well established through previous studies and no longer need to be explored.

### **Validity and Reliability Tests**

While the data collection stage is still ongoing, some completed responses are retrieved from the online database server to test the validity of the instruments using EFA. This due to the fact that even though most instruments used in the study are adapted from previous work and have been tested, however as the items are used in a new meaning in this research, and that some of the items are revised, the items are

re-examined to ensure the quality of the findings and conclusion of this research. Cronbach's alpha is used to measure if the items reliably measure the same latent variable.

Data is cleaned before conducting further analysis. Standard deviation is calculated using the function tools available in excel. Cases with standard deviation of below 0.7 which represent unengaged responses on a seven-point scale are eliminated (Gaskin, 2012a). In this research, unengaged responses refers to suspicious response pattern such as when respondent marks the same response for several group of items e.g. 55555 44444 6666. Then the data is imported into the IBM SPSS 19 software. Following the suggestion in Sekaran & Bougie (2010), incomplete responses with missing data of more than 25% are deleted. Negative items are reverse-coded and missing data is replaced with the median of nearby points. The median of nearby points method is used for data replacement because the data of the surveyis in the form of discrete number. In this case, the mean is irrelevant to be used as replacement value as the mean may be a decimal number. Finally, cases with high risk outliers are identified and removed. In the end, a total of 212 usable cases are identified which consisted of 84 responses in the non-persuasive sample, and 128 responses in the persuasive sample. Sample size exceeds the minimum of 100 participants and minimum ratio of 5 participants per item (Hair, Black, Babin, Anderson, & Tatham, 2009). Normality check using Shapiro-Wilk statistics shows that the data distribution for each item is not normally distributed.

### **Exploratory Factor Analysis (EFA)**

Initially, the factorability of the 44 items is examined. The criteria for the factorability of a correlation recommended in Hooper (2012) is used (please refer Table 1). Firstly, all of the 44 items correlate at more than 0.3 with at least one other item, suggesting reasonable factorability. Secondly, the Kaiser-Meyer-Olkin (KMO) measure of overall sampling adequacy is equal to 0.905, well above the recommended value of 0.5, and Bartlett's test of sphericity is significant with the p value of  $< 0.05$  (Hair et al., 2009). The communality for each item is set to be above 0.4 (Leimeister, 2010) to confirm that each item shares some common variance with other items. First round of the analysis meets the minimum requirement of items correlation, KMO, and Bartlett's test of sphericity. However, four items show communalities below 0.4. Thus, the four items are deleted one by one, and the factor analysis is repeated each time. With 40 items remaining in the list, the new factor analysis shows stronger results with KMO of 0.906 and communalities range between 0.470 - 0.830.

The next decision relates to the number of factors to be retained. As the data distribution is significantly not normal, Principle Axis Factoring (PAF) is preferred as the extraction method. PAF is recommended by Costello & Osborne (2005) to bring the best results for non-normal data. To determine the number of factors, two methods are considered; 1) eigen one rule or Kaiser-Guttman, and 2) scree plot graph of the eigenvalues. A predetermine level of cumulative variance is set to a minimum

of 60% representing the satisfactory percentage of variance criterion in social sciences (Hair et al., 2009). It is known that orthogonal rotations produced factors that are uncorrelated and resulted in a loss of valuable information if the factors are actually correlated (Costello & Osborne, 2005). Thus, an oblique rotation is preferred for the rotation method as in the social sciences some correlation among factors is expected (Kock, 2015b). Contrarily, In IBM SPSS 19, direct oblmin and promax are the available oblique rotation. In this research, promax rotation with the default Kappa (4) is used.

Nine factors emerge with eigenvalues greater than 1, explaining 66.943% of the total variance. The result shows that 66.943% of the common variance shared by the 40 items can be accounted for by nine factors. The scree plot graph also shows that the last significant break point in the graph shape above the 10th point. This also means that only 9 factors should be extracted as only factors above and excluding the break point should be retained (Hooper, 2012). When using oblique rotation, the pattern matrix is examined for factor/item loadings and the factor correlation matrix shows the correlation among factors. Small coefficients below 0.4 were suppressed so that results will be easier to interpret. An item with loading less than 0.4 (in the case of sample size,  $N = 212$ ) on all factors indicates that the item is insignificant and should be deleted (Hair et al., 2009; Hooper, 2012). An observation is also made to inspect any sign of cross-loading in which an item with coefficients greater than 0.4 on more than one factor.

Convergent validity is evident with the factor loadings of 0.57 and above on each factor. Convergent validity means that the items within a single factor are highly correlated (Gaskin, 2012b). Two primary methods are used to determine discriminant validity, which is the extent to which factors are distinct and uncorrelated. The first method is to examine the pattern matrix. The rule was that the items should be related more strongly to their own factor than to another factor. If "cross-loadings" exist (items load on multiple factors), the cross-loadings should differ by more than 0.2 (Gaskin, 2012b). The pattern matrix shows that there is no significant discriminant validity issue as the result is free from significant cross-loadings. Secondly, the factor correlation matrix is examined. The factor correlation matrix shows no present of correlation greater than 0.7, which implied that each factor assesses a unique variable. It is noted that one credibility item does not load properly with the rest of the credibility instruments. Instead, the item loads on the visual aesthetic and engagement dimensions. This indicates a face validity issue in which the items that are not theoretically similar in nature, load together on the same factor (Gaskin, 2012b). Thus, the item is deleted from the list.

The factor labels are revised as several items of six proposed latent variables load on the same factors. The engagement and visual aesthetic items load significantly together. This may be due to the fact that the users' perception are made instantly based on their short impression of the visual design of the website, thus, neglecting

the website interactivity. This new factor is labelled as 'visual engagement'. The items representing reciprocity and commitment also load together, and labelled as 'gratitude' (being thankful). In theory, reciprocity and commitment are about personal feelings of (1) being thankful and (2) being committed. This justified the reason of the items sharing the similar variance. Three items of authority, one item of social proof, and two items of liking significantly load together on another factor. The items of the new factor are all related to human figures, thus the factor is labelled as 'human persona'. Another two items of social proof are grouped separately and labelled as 'wisdom of crowds' as the items are related to testimonials or ratings from a large group of people. The 'visual engagement' factor is re-analysed with a separate EFA and the result shows that only one factor can be extracted. The factorability of the social influence related factors i.e. gratitude, human persona, wisdom of crowds, and scarcity is also repeated and the result produces the same set of factors as mentioned above. This indicates that the new factors are stable.

The wisdom of crowds and the satisfaction factors are represented by two items each, which is lower than the general requirement of 3 items per factor (Costello & Osborne, 2005). However, Hayduk & Littvay (2012) recommend the use of the few best items. They argue that one or two items are sufficient for latent variable to be included in a structural equation model. Furthermore, this research is the extension of an existing structural theory and the proposed model will be further analysed using the Partial Least Squares - Structural Equation Modelling (PLS-SEM). PLS-SEM allows for fewer items (1 or 2) per factor. Thus, the two factors are retained in the research model.

In particular, nine dimensional factor structures for assessing visual persuasion are discovered. Factor analysis is repeated with 39 items. The result shows slightly lower KMO (0.901), yet communalities index is quite similar with the new range between 0.470 - 0.829. As any KMO of above 0.9 is marvellous, 39 items are finalised as the instruments to be further examined with CFA. Internal consistency for each factor is examined using Cronbach's alpha. The alphas are moderate and well above the lower limit of 0.70 and none of the item shows increases in alpha if the item is deleted. Corrected item-total correlations are well above 0.5, and all inter item correlations exceed 0.3; meeting the minimum requirement in Joseph F Hair et al. (2009). Table 1 shows the EFA's assessment criteria and Table 2 shows the result of EFA and Cronbach's alpha. The finalised instruments of this research are shown in Table 3.

Table 1. Instruments assessment's guide for EFA

Criterion	Note	Reference
Inter items correlation	> 0.3 with at least one other item	Hooper (2012), Joseph F Hair et al. (2009)
Kaiser-Meyer-Olkin (KMO)	> 0.5	
Bartlett's test of sphericity	p < 0.05	

Communalities	> 0.4	Leimeister (2010)
Cumulative variance	> 60%	Joseph F Hair et al. (2009)
Factor loading	> 0.4 (sample size > 200)	
Cross-loading	< 0.4	Gaskin (2012b)
	significant cross-loadings should differ by more than 0.2	
Factor correlation matrix	< 0.7	
Cronbach's alpha	> 0.7	Joseph F Hair et al. (2009)
Corrected item-total correlations	> 0.5	

Table 2. Factor Loadings Based On A Principle Axis Factoring With Promax Rotation, and Cronbach's Alpha for 39 Items (N = 212)

Pattern Matrix				Cronbach's Alpha
Proposed Factors	New Factors	Item code	loadings	
Informativeness		Info1	0.672	0.869
		Info2	0.76	
		Info3	0.803	
		Info4	0.748	
Usability		Use1	0.961	0.835
		Use2	0.886	
		Use3	0.736	
Visual Aesthetic	Visual engagement	VisEng1	0.7	0.938
		VisEng2	0.787	
		VisEng3	0.611	
Engagement		VisEng4	0.466	
		VisEng5	0.745	
		VisEng6	0.764	
		VisEng7	0.844	
		VisEng8	0.898	
Credibility		Credib1	0.609	0.715
		Credib2	0.786	
		Credib3	0.695	
Satisfaction		Satisfy1	0.893	0.862
		Satisfy2	0.873	
Reciprocity	Gratitude	Gratit1	0.792	0.897
		Gratit2	0.855	



Commitment		Gratit3	0.726	
		Gratit4	0.712	
		Gratit5	0.608	
		Gratit6	0.649	
		Gratit7	0.638	
Liking	Persona	Person1	0.93	0.889
Authority		Person2	0.722	
		Person3	0.84	
		Person4	0.727	
		Person5	0.75	
		Person6	0.775	
Social Proof	Crowds	Crowd1	0.817	0.928
		Crowd2	0.796	
Scarcity		Scarce1	0.596	0.833
		Scarce2	0.516	
		Scarce3	0.94	
		Scarce4	0.89	

Table 3. Finalised Survey's Instruments

Factors	Label	Web Design Instruments	Adopted/adapted/newly constructed	Note
Informativeness	Info1	The information on this website is sufficient.	WebMAC Business (Small & Arnone, 1998)	
	Info2	The information on this website is useful.	Tang (2009)	
	Info3	The information on this website appears to be relevant and up-to-date.	WebMAC Business (Small & Arnone, 1998)	
	Info4	The travel information on this website appears to be accurate.		
Usability	Use1	This website is easy to use.	USE as in Albert & Tullis (2013)	
	Use2	I quickly familiarise		

		myself with this website.		
	Use3	This website makes it easy to go back and forth between pages.	Tang (2009)	
Visual engagement	VisEng1	This website has an attractive appearance.	WebMAC Business (Small & Arnone, 1998)	
	VisEng2	This website has good use of colour and layout.	WebMAC Professional (Small & Arnone, 2000)	
	VisEng3	The content in this website is well designed.	WebMAC Business (Small & Arnone, 1998)	
	VisEng4	The varieties of visuals (e.g. text, images, animation etc.) help to maintain attention.		
	VisEng5	The visuals included in this website enhance the presentation of the travel information.		
	VisEng6	This website provides opportunities for interactivity.		
	VisEng7	This website stimulates curiosity and exploration	WebMAC Professional (Small & Arnone, 2000)	
	VisEng8	The main page (i.e. the first webpage) of this website is interesting enough to	WebMAC Business (Small & Arnone, 1998)	

		continue browsing further.		
Credibility	Credib1	I think that this website has sufficient expertise in providing travel information and services.	Cugelman, Thelwall, & Dawes (2009)	Reverse-coded
	Credib2	I think some of the information in this website seems suspicious (e.g. misleading, fictitious, or made-up information).		
	Credib3	I need to verify some of the information (e.g. with friends or travel agents) before I can put my trust to this website.	WebMAC Business (Small & Arnone, 1998)	Reverse-coded
Satisfaction	Satisfy1	This website quickly loads all the text and graphics.		
	Satisfy2	In overall, I am satisfied with this website.	USE as in Albert & Tullis (2013)	
Gratitude	Gratit1	I want to give reviews about my past travel experiences on this website.	Newly constructed	
	Gratit2	I want to register as a member of this website.	Kim (2008)	
	Gratit3	I want to receive special	Kim (2008)	

		offers from this website. So, I will provide my address or contact number if asked by this website.		
	Gratit4	I will provide my friends' addresses if asked by this website.	Newly constructed	
	Gratit5	I want to search for travel destinations, flights, and hotels information on this website.		
	Gratit6	I am tempted to click on the result links from my searching activities on this website.		
	Gratit7	I will visit other related websites that are recommended by this website (e.g. via image or web link).		
	Person1	I will like the website more if I see some pictures of other people that share something similar with me on the website (e.g. picture of hikers - if you		
Persona				

		like adventurous activity).	
	Person2	I will like the website more if I see some pictures of friendly persons on the website.	
	Person3	I will trust the website more if there are some pictures of celebrities on the website.	
	Person4	I will trust the reviews that come from celebrities.	
	Person5	I will trust the information that come from an authoritative person (e.g. flight's staff, chef, representative of local Tourism Ministry etc.).	
	Person6	I will like the website more if I see familiar faces on the website (e.g. friends, or friends of friends).	
Crowds	Crowd1	I will trust the reviews (positive or negative) from other travellers on the website.	
	Crowd2	I will trust the	

		information more if I see other people paid attention to it as well (e.g. number of 'likes' at a Like button).	
Scarcity	Scarce1	I think that price is one of the most important information in a travel website.	
	Scarce2	I think that discount highlight is also important for a travel website.	
	Scarce3	I think that I will act fast to purchase a travel package if I see the 'Limited Offer' or 'Ending Soon' sign on the advertisement.	
	Scarce4	I believe that I will be missing out on some good deals if I fail to act quickly with my purchasing.	

### Confirmatory Factor Analysis (CFA)

When a satisfactory number of responses required by the research are achieved, the raw data are downloaded from the Survey Monkey server in Microsoft Excel format. Complying with the ethic's consideration, data from the web server is deleted once it is secured in a password-protected computer. Data cleaning is carried out, resulted to 290 usable cases in which 181 responses are from the persuasive sample group. The 39 items retrieved from the EFA section, and four items of an observed variable that

represent the behavioural intention dimension, are re-examined with CFA. The behavioural intention dimension includes three items that measure the intention to use, to purchase, and to recommend, and one item to measure the attitude towards the destination. In this section, the factors obtained from EFA are referred as the latent variables, whereas the term indicator is used to refer to the survey item.

CFA is conducted in conjunction with PLS-SEM analysis using the software called WarpPLS version 5.0. WarpPLS software provides the users with features of which are not available from other SEM software (Kock, 2015b). The software is the first to explicitly identify nonlinear functions connecting pairs of latent variables in SEM models and calculate multivariate coefficients of association accordingly. It is also the first software to provide classic PLS algorithms together with factor-based PLS algorithms for SEM (Kock, 2014). Factor-based PLS algorithms generated estimates of both true composites and factors, fully accounting for measurement error. Original PLS design based its model estimation only on the composites; the linear combinations of indicators (Kock, 2014). Composite based do not take measurement error into consideration. With composite based the path coefficient tends to be weaker, thus leads to biased model parameter estimates particularly on the path coefficients and loadings (Kock, 2015a). On the other hand, factor-based incorporates measurement errors. Factor scores also accounted nonlinearity and estimate best-fitting nonlinear functions which lead to stable and reliable path coefficients. Moreover, factor-based PLS algorithms combine the precision of covariance-based SEM algorithms under common factor model assumptions with the nonparametric characteristics of classic PLS algorithms (Kock, 2014). These advantages enable the data from this research to be analysed as the data distribution is not normal, whereas normality is a major requirement for CB-SEM software.

There are five main steps to be taken to analyse data with the software (Kock, 2010). Firstly, a project file is created. Then, the raw data is imported into the software. Data imported into WarpPLS automatically go through data pre-processing. The software checks and corrects missing values, zero variance problem, identical columns (also known as the indicators) names, and rank problems. The data are also standardised in step three. Standardised data columns have means that equal zero and standard deviations that equal one (Kock, 2010). As the pre-process are automatically carried out, it is crucial to correct missing values prior to importing the data into WarpPLS so that the percentage of corrected missing values do not exceed 25% as recommended by Sekaran & Bougie (2010). The data used in this research do not appear to have the respective above mentioned problem.

In step four, five reflective latent variables, one formative latent variable, and one observed variables are defined in the SEM model. All the variables namely informativeness, usability, credibility, satisfaction, visual engagement, and behavioural intention are assigned with four, three, three, two, eight, and four indicators, respectively. The indicators of reflective latent variables are expected to be highly correlated with the latent variable score. The social influence latent variable

is defined as the formative latent variable. In the model, social influence is a second-order latent variable in which the indicators are made of other latent variables; i.e. gratitude, human persona, wisdom of crowds, and scarcity. Thus the indicators of social influence variable are expected to measure certain attribute of social influence, but they are not expected to be correlated among themselves. A direct link connects each latent variable to the observed variable. In this software, the variables are called as the outer model whereas the model links are referred as the inner model. WarpPLS estimates collinearity, measurement error and composite weights before the SEM analysis is executed in step five. If any of the above mentioned assessments appears to be too high, users are warned about possible unreliability of the results. In the current research, none of the issues arise. Based on the analysis with WarpPLS, the validity and reliability of the measurement scales are assessed in two stages: 1) evaluation of first-order latent constructs, and 2) evaluation of second-order latent constructs.

**First-order latent variables' evaluation**

Several criteria are being considered during the assessments as shown in Table 4. For reflective constructs, the combined loadings and cross-loadings provided by WarpPLS software are used to describe the convergent validity of the measurement scales. Convergent validity can be obtained if the instruments are understood by the respondents in the same way as they are intended by the researchers (Kock, 2015b). In this respect, two criteria are assessed: 1) the p values associated with the loadings must be equal to or lower than 0.05 (Kock, 2015b), 2) the loadings must be equal to or greater than 0.5 (Hair et al., 2009). As shown in Table 5, the p values for all indicators are lower than 0.05, and the loadings for the indicators of each latent variable are well above 0.5. No significant cross-loading is found, indicating acceptable convergent validity.

Table 4. First-order assessment's guide for WarpPLS

Assessment	Criterion	Note	Reference
Convergent validity	Individual item standardised loading on parent factor	Min. of 0.50	Hair et al. (2009)
	Loadings with significant p value	< 0.05	Kock (2015)
Reliability	Cronbach's alpha	> 0.70	Hair et al. (2009)
	Composite reliability	> 0.70	
Discriminant validity	Square-root of AVE	More than the correlations of the latent variables	Hair et al. (2009)
	Average variance extracted (AVE)	> 0.50	
Collinearity	Variance Inflation Factor (VIF)	< 3.3	Kock (2015)



Following the recommendation in Kock (2015b), measurement reliability is assessed with Cronbach's alpha coefficient, and Dillon–Goldstein rho coefficient (DG's rho), also known as the composite reliability coefficient (Tenenhaus, Vinzi, Chatelin, & Lauro, 2005). The average variances extracted (AVEs), and the square-root of AVEs are used to assess discriminant validity, and full collinearity variance inflation factors (VIFs) is used to assess the overall collinearity. Based on Table 6, all the first-order latent variables exceed all the minimum requirement of reliability, collinearity, and part of discriminant validity with the Cronbach's alpha and DG's rho coefficients of above 0.7, VIFs of less than 3.3, and AVEs of above 0.5, respectively. However, the square root of AVEs for gratitude is slightly lower than the correlation of the variable with the behavioural intent variable. For each latent variable, the square root of the average variance extracted should be higher than any of the correlations involving that latent variable (Fornell & Larcker, 1981). In the actual research model, gratitude serves as an indicator to the social influence variable; a formative second-order latent variable. It is noted that formative latent variable is the most likely to lead to discriminant validity problems (Kock, 2015b). As a second-order variable's indicator, gratitude is not expected to directly correlate with the intent variable. Therefore, the variable is retained as the second-order variable will be evaluated separately on the next section.

	Intent	Usability	Informative	Engage	Satisfy	Credibility	Gratitude	Crowd	Scarce	Persona	P value
Intent 1	<b>0.893</b>	0.053	-0.018	0.003	0.057	-0.062	0.132	-0.056	0.061	-0.053	<0.001
Intent 2	<b>0.880</b>	-0.004	0.004	-0.084	-0.097	0.047	0.146	0.073	0.019	-0.005	<0.001
Intent 3	<b>0.917</b>	-0.012	-0.046	0.085	0.049	-0.032	-0.025	-0.062	0.053	0.031	<0.001
Intent 4	<b>0.845</b>	-0.038	0.065	-0.007	0.094	0.051	-0.264	0.049	0.027	0.028	<0.001
Use1	0.181	<b>0.809</b>	-0.025	0.006	0.116	-0.213	-0.189	-0.003	0.02	-0.051	<0.001
Use2	-0.064	<b>0.879</b>	0.091	0.051	0.015	0.081	0.101	-0.011	0.092	-0.026	<0.001
Use3	-0.11	<b>0.825</b>	-0.072	-0.06	0.098	0.123	0.078	0.015	0.079	0.078	<0.001
Info1	-0.065	0.071	<b>0.793</b>	-0.009	0.126	0.046	-0.063	0.028	0.152	0.002	<0.001
Info2	-0.213	0.111	<b>0.828</b>	0.117	0.053	-0.039	-0.108	0.158	0.093	0.054	<0.001
Info3	0.068	-0.061	<b>0.838</b>	-0.034	0.132	-0.1	0.067	-0.057	0.116	0.006	<0.001
Info4	0.21	-0.119	<b>0.813</b>	-0.076	0.041	0.097	0.103	-0.129	0.124	-0.064	<0.001
VisEng1	0.224	-0.061	0.144	<b>0.826</b>	-0.023	0.009	-0.023	0.003	0.121	-0.148	<0.001

VisEn g2	0.01 3	-0.024	-0.202	<b>0.79</b> <b>8</b>	0.09 1	-0.042	0.023	0.07 6	- 0.10 3	-0.102	<0.00 1
VisEn g3	0.03 3	0.077	0.084	<b>0.73</b> <b>1</b>	0.02 8	-0.008	-0.144	0.04 5	- 0.16 8	0.141	<0.00 1
VisEn g1	0.27	0.006	0.224	<b>0.71</b> <b>4</b>	0.08 6	-0.069	-0.11	- 0.21 5	0.04 2	0.1	<0.00 1
VisEn g2	- 0.26	0.148	0.069	<b>0.80</b> <b>5</b>	- 0.03 1	0.118	0.008	0.08 4	0.18 1	0.001	<0.00 1
VisEn g3	- 0.26 1	0.062	-0.151	<b>0.76</b> <b>7</b>	- 0.02 4	0.122	0.078	0.09 1	- 0.02 9	0.092	<0.00 1
VisEn g4	- 0.26 8	-0.115	-0.13	<b>0.79</b> <b>7</b>	- 0.06 6	-0.114	0.239	0.07 8	0.13 1	0.014	<0.00 1
VisEn g5	0.24 9	-0.077	-0.018	<b>0.84</b> <b>9</b>	- 0.04 5	-0.02	-0.085	- 0.16 8	0.05 5	-0.063	<0.00 1
Satisf y1	0.02 5	-0.027	0.031	- 0.11 9	<b>0.89</b> <b>8</b>	-0.029	-0.038	0.03 2	- 0.00 9	0.069	<0.00 1
Satisf y2	- 0.02 5	0.027	-0.031	0.11 9	<b>0.89</b> <b>8</b>	0.029	0.038	- 0.03 2	0.00 9	-0.069	<0.00 1
Credib b1	0.01 6	0.046	0.153	0.21 7	- 0.04 5	<b>0.743</b>	0.088	0.00 5	0.11	-0.058	<0.00 1
Credib b2	- 0.01 2	0.016	-0.061	- 0.08 2	- 0.08 4	<b>0.861</b>	-0.119	- 0.04 9	0.06 5	0.148	<0.00 1
Credib b3	- 0.00 2	-0.059	-0.074	- 0.11 1	0.13	<b>0.818</b>	0.046	0.04 7	- 0.16 8	-0.103	<0.00 1
Gratit 1	- 0.25 6	-0.123	-0.042	0.07 2	- 0.00 3	-0.006	<b>0.712</b>	0.22 2	- 0.12 2	-0.002	<0.00 1

Gratit 2	0.16 2	-0.168	-0.007	0.14 4	- 0.10 1	-0.033	<b>0.817</b>	- 0.00 5	- 0.17 8	0.02	<0.00 1
Gratit 3	0.19 2	-0.055	-0.001	- 0.10 7	- 0.14 2	0.166	<b>0.705</b>	- 0.06 1	- 0.16 2	-0.085	<0.00 1
Gratit 4	0.17 5	-0.256	-0.2	0.11 7	0.00 6	0.185	<b>0.655</b>	- 0.07 5	- 0.21 7	-0.038	<0.00 1
Gratit 5	- 0.20 5	0.243	0.087	- 0.06 4	0.12 5	-0.03	<b>0.747</b>	- 0.07 5	0.22 9	0.093	<0.00 1
Gratit 6	- 0.06 4	0.243	-0.017	0.09	0.06 3	-0.067	<b>0.778</b>	0.00 9	0.16 5	-0.034	<0.00 1
Gratit 7	0.00 5	0.084	0.16	- 0.26 6	0.05 3	-0.182	<b>0.720</b>	- 0.01 8	0.26 2	0.037	<0.00 1
Crow d1	- 0.01 3	-0.003	0.023	- 0.03 1	0.04 9	-0.016	0.055	<b>0.95 1</b>	- 0.03 9	0.05	<0.00 1
Crow d2	0.01 3	0.003	-0.023	0.03 1	- 0.04 9	0.016	-0.055	<b>0.95 1</b>	0.03 9	-0.05	<0.00 1
Scarc e1	- 0.38	0.094	0.182	0.03 9	0.05 4	-0.01	0.019	- 0.06 1	<b>0.78 2</b>	0.141	<0.00 1
Scarc e2	- 0.13 8	0.159	0.081	0.03	0.03 4	0.054	-0.135	- 0.10 2	<b>0.82 1</b>	0.058	<0.00 1
Scarc e3	0.19 7	-0.165	-0.13	-0.07	- 0.00 2	-0.053	0.099	0.07	<b>0.75 5</b>	-0.126	<0.00 1
Scarc e4	0.36 5	-0.11	-0.154	- 0.00 4	- 0.09 5	0.005	0.03	0.11	<b>0.71 8</b>	-0.087	<0.00 1
Perso n1	- 0.07 5	-0.061	-0.016	0.05	0.12 9	0.076	-0.066	0.12 2	0.12	<b>0.806</b>	<0.00 1

Perso n2	0.09 4	0.05	-0.003	- 0.14 2	0.02 7	0.221	-0.082	0.19	0.00 8	<b>0.763</b>	<0.00 1
Perso n3	- 0.12 3	0.066	0.017	0.09 7	- 0.02 5	-0.129	-0.008	- 0.32 3	- 0.01 2	<b>0.727</b>	<0.00 1
Perso n4	- 0.12 2	0.071	-0.076	0.01 6	- 0.08 2	-0.159	0.26	- 0.25 6	- 0.05 5	<b>0.710</b>	<0.00 1
Perso n5	0.00 3	0.055	-0.074	- 0.11 1	0.05 7	-0.012	-0.055	0.33 5	- 0.04 5	<b>0.711</b>	<0.00 1
Perso n6	0.24 5	-0.19	0.162	0.09 7	- 0.13 2	-0.023	-0.035	- 0.09 7	- 0.03 3	<b>0.668</b>	<0.00 1

Table 5. First-Order Latent Variables Loadings, Cross-Loadings, and Cronbach's alpha  
Table 6. Assessment of First-Order Latent Variables

	Inte nt	Usabili ty	Informat ive	Enga ge	Satis fy	Credibil ity	Gratitu de	Crowd	Scar ce	Perso na
Cronbach' s α	0.90 7	0.788	0.835	0.912	0.76 1	0.733	0.857	0.89 5	0.77 0	0.826
DG's rho	0.93 5	0.876	0.890	0.928	0.89 3	0.850	0.891	0.95 0	0.85 3	0.874
Intent	0.88 4									
Usability	0.32 3	0.838								
Informati ve	0.46 8	0.423	0.818							
Engage	0.71 6	0.459	0.635	0.787						
Satisfy	0.36 0	0.265	0.402	0.513	0.89 8					
Credibilit y	0.39 4	0.265	0.279	0.339	0.15 9	0.809				
Gratitude	0.75 0	0.434	0.512	0.694	0.32 8	0.380	0.735			
Crowd	0.39 2	0.100	0.254	0.325	0.17 9	0.179	0.309	0.95 1		
Scarce	0.43 9	0.293	0.471	0.472	0.22 7	0.244	0.478	0.16 0	0.77 0	
Persona	0.41 7	0.222	0.277	0.382	0.17 4	0.265	0.404	0.34 6	0.39 7	0.732

AVEs	0.78 2	0.703	0.669	0.619	0.80 7	0.654	0.540	0.90 5	0.59 3	0.536
VIFs	3.00 7	1.394	1.901	3.269	1.38 5	1.241	2.832	1.26 8	1.53 9	1.400

**Second-order latent variable's evaluation**

The social influence variable is a formative second-order latent variable. The evaluation of second-order latent variable is conducted based on the approach used by Schmiedel, vom Brocke, & Recker (2014). Three criteria are assessed: 1) p values and variance inflation factors (VIF) associated with indicator weights, 2) adequacy coefficient ( $R^2_a$ ), and 3) conceptual redundancy. The p values of indicator weights are the indication that the formative latent variable measurement items are properly constructed. In this research, all p values meet the desirable requirement for formative indicators of lower than 0.05 (Kock, 2011). The strength of the relationship between the first-order and the second-order variables are assessed with the adequacy coefficient ( $R^2_a$ ).  $R^2_a$  is calculated by summing the squared correlations ( $R^2$ ) between the variable and its indicators and dividing by the number of indicators (Mackenzie, Podsakoff, & Podsakoff, 2011). In this case, the value of  $R^2_a$  for social influence is 0.553 which is greater than the cut off value of 0.50. This indicated that on average, a majority of the variance in the indicators is shared with the construct. In formative latent variables, indicators are expected to measure different facets of the same construct, which means that they should not be redundant. VIF are used for redundancy assessment. VIFs below 2.5 are desirable for formative indicators (Kock, 2011); which is met by all indicators of the research. The assessment for the second-order latent variables is shown in Table 7.

Table 7. Assessment of Second-Order Variables

	Social Influence	Type	SE	P value	VIF	WLS	ES	R	$R^2$	$R^2_a$
*lv_Gratitudo	0.480	Formative	0.067	<0.001	1.897	1	0.418	0.871	0.759	0.553
*lv_Persona	0.260	Formative	0.071	<0.001	1.973	1	0.206	0.793	0.629	
*lv_Crowds	0.242	Formative	0.071	<0.001	1.381	1	0.134	0.553	0.306	
*lv_Scarcity	0.268	Formative	0.07	<0.001	1.817	1	0.193	0.728	0.518	

\*latent variable indicators

SE: standard error, VIF: variance inflation factors ( ), WLS: weight-loading signs, ES: effect size, R: correlation coefficient,  $R^2$ : squared correlation coefficient,  $R^2_a$ : adequacy coefficient

General assessment

A negative weight-loading sign (WLS) of an indicator suggests the existence of a Simpson's paradox instance (Pearl, 2009; Wagner, 1982) in the outer model. A

Simpson's paradox instance indicates a causality problem. It is recommended that all indicators' WLS values be positive, for both formative and reflective latent variables. In this case, all WLS values are equal to one, which is a positive value. The indicator's effect sizes (ES) shows the extent of contribution it had on the R<sup>2</sup> coefficients of the latent variable to which each indicator is associated with. The indicator effects are small, medium, or large with the usual recommended value of 0.02, 0.15, and 0.35; respectively (Cohen, 1988). Values below 0.02 suggesting effects that is too weak to be considered relevant, even when the corresponding p values are statistically significant. It is recommended that all indicator effect sizes be equal to or greater than 0.02, for both formative and reflective latent variables (Kock, 2015b). In this case, the indicators had the minimum effect size of 0.094, well above recommended value. Based on the result of the CFA, all measurement scales used in this research are appropriate to be used for further PLS-SEM analyses.

## Conclusion

This paper demonstrates the EFA and CFA to assess the validity and reliability of the research instruments. In particular, nine dimensional factor structures for assessing visual persuasion are discovered. The result shows that participants perceived differently between a general web evaluation and the evaluation specifically based on visual design. This implies that users' perception varies according to evaluation goals. This also means that the social influence principles by Cialdini (2007) can be applied to online persuasion, specifically to online visual persuasion. However, the variation of data shows that the pattern of the data relies heavily on the visual elements used in the study and may be different from the existing theory. For example, even though the instruments representing the authority, liking, and social proof principles are theoretically separated, yet the items that relate to human figures tend to load on one factor, while other textual or symbol cues tend to load on another factor. The findings suggest that the proposed 39 items are valid and reliable for measuring the persuasiveness of visual persuasion.

## References

- [1] Albert, W., & Tullis, T. (2013). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics* (2nd Editio.). Elsevier Science.
- [2] Chu, H.-L., Deng, Y.-S., & Chuang, M.-C. (2014). Persuasive Web Design in e-Commerce. In F.-H. Nah (Ed.), *HCI in Business SE - 47* (Vol. 8527, pp. 482–491). Springer International Publishing. doi:10.1007/978-3-319-07293-7\_47
- [3] Cialdini, R. B. (2007). *Influence: The Psychology of Persuasion*. Collins (Revised ed., Vol. 55). HarperBusiness;
- [4] Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1–9. doi:10.1.1.110.9154

- [5] Cugelman, B., Thelwall, M., & Dawes, P. L. (2009). The Dimensions of Web Site Credibility and Their Relation to Active Trust and Behavioural Impact. *Communications of the Association for Information Systems*, 24, 455–472.
- [6] Cyr, D. (2013). Website design, trust and culture: An eight country investigation. *Electronic Commerce Research and Applications*, 12, 373–385. doi:10.1016/j.elerap.2013.03.007
- [7] Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research (JMR)*. Feb1981, 18(1), 39–50. 12p. 1 Diagram. doi:10.2307/3151312
- [8] Gaskin, J. (2012a). Data screening. Gaskination's StatWiki. Retrieved March 30, 2015, from <http://statwiki.kolobkreations.com>
- [9] Gaskin, J. (2012b). Exploratory Factor Analysis. Gaskination's StatWiki. Retrieved April 5, 2015, from <http://statwiki.kolobkreations.com>
- [10] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Multivariate Data Analysis*. Prentice Hall.
- [11] Hao, J.-X., Tang, R., Yu, Y., Li, N., & Law, R. (2015). Visual Appeal of Hotel Websites: An Exploratory Eye Tracking Study on Chinese Generation Y. In I. Tussyadiah & A. Inversini (Eds.), *Information and Communication Technologies in Tourism 2015 SE - 44* (pp. 607–620). Springer International Publishing. doi:10.1007/978-3-319-14343-9\_44
- [12] Hayduk, L. A., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology*. doi:10.1186/1471-2288-12-159
- [13] Hooper, D. (2012). Exploratory factor analysis. In H. (Ed. . Chen (Ed.), *Approaches to Quantitative Research - Theory and its Practical Application: A Guide to Dissertation Students*. Cork, Ireland: Oak Tree Press.
- [14] Horvath, J. (2011). Persuasive Design: It's Not Just about Selling Stuff . In A. Marcus (Ed.), *Design, User Experience, and Usability. Theory, Methods, Tools and Practice (Vol. 6770, pp. 567–574)*. Springer Berlin / Heidelberg. doi:10.1007/978-3-642-21708-1\_63
- [15] Kim, H. (2008). Persuasive Architecture of Tourism Destination Websites: An Analysis of First Impression.
- [16] Kim, H., & Fesenmaier, D. R. (2008). Persuasive Design of Destination Web Sites: An Analysis of First Impression. *Journal of Travel Research*, 47(1), 3–13. doi:10.1177/0047287507312405
- [17] Kock, N. (2010). Using WarpPLS in E-collaboration Studies: An Overview of Five Main Analysis Steps. *Int. J. E-Collab.*, 6(4), 1–11. doi:10.4018/jec.2010100101



- [18] Kock, N. (2011). Using WarpPLS in e-collaboration studies: Mediating effects, control and second order variables, and algorithm choices. *International Journal of E-Collaboration*, 7(3), 1–13. doi:10.4018/jec.2011070101
- [19] Kock, N. (2014). A note on how to conduct a factor-based PLS-SEM analysis.
- [20] Kock, N. (2015a). One-Tailed or Two-Tailed P Values in PLS-SEM? *Int. J. E-Collab.*, 11(2), 1–7. doi:10.4018/ijec.2015040101
- [21] Kock, N. (2015b). *WarpPLS 5.0 User Manual*. Laredo, TX: ScriptWarp Systems.
- [22] Leimeister, S. (2010). IT outsourcing governance: Client types and their management strategies. *IT Outsourcing Governance: Client Types and Their Management Strategies*. doi:10.1007/978-3-8349-6303-1
- [23] Mackenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293–334. Retrieved from <http://dl.acm.org/citation.cfm?id=2017510\papers3://publication/uuid/E4D3717C-7F3F-4791-8835-141D4309976B>
- [24] Schmiedel, T., vom Brocke, J., & Recker, J. (2014). Development and validation of an instrument to measure organizational cultures' support of business process management. *Information & Management*, 51(1), 43–56. doi:10.1016/j.im.2013.08.005
- [25] Sekaran, U., & Bougie, R. (2010). *Research methods for business: a skill-building approach*. Chichester: Wiley.
- [26] Small, R. V., & Arnone, M. P. (1998). *Website Motivational Analysis Checklist for Business (WebMAC Business)*. New York: The Motivation Mining Company.
- [27] Small, R. V., & Arnone, M. P. (2000). *Website Motivational Analysis Checklist (WebMAC): Professional*. New York: Motivation Mining.
- [28] Tang, L. (2009). *Destination websites as advertising: An application of Elaboration Likelihood Model*. ProQuest, UMI Dissertations Publishing.
- [29] Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis*, 48(1), 159–205. doi:10.1016/j.csda.2004.03.005
- [30] Winn, W., & Beck, K. (2002). The Persuasive Power of Design Elements on an E-Commerce Web Site. *Technical Communication*, 49(1), 17–35.