

Semantic Analysis Process: Quantification and Validity of Translation Process

Shah Rukh Tariq and Iffat Batool

Govt. College University

In present study semantic analysis approach having a distinct calculation method as compared to Landauer's (1999) approach was applied for translating Curiosity and Exploration Inventory-II (Kashdan et al., 2009). The new Semantic analysis comprised of five steps: (a) identification of the significant words in the item (b) conceptual or contextual meaning of the words rather than literal meaning (c) categorization of responses (d) comparison of three types of response (e) percentages of favorable (item value) and other responses. The study was conducted in three phases. In phase-I, two samples were used: sample 1 consisted of 15 university students for the process of semantic analysis whereas sample 2 consisted of 50 students to check the psychometric properties of the original CEI-II. During the second phase, the process of forward-backward translation was done and 50 university students were recruited to check the psychometric properties of the Urdu version of CEI-II. For the third phase, 250 university students were recruited to check the item analysis and another sample consisted of 150 university students was recruited to check the confirmatory factor analysis of CEI-II Urdu version. The samples was comprised of university students with age ranging from 18 to 24 years ($M = 21.24$, $SD = .91$). In present study, nonsignificant value of SAV ($n = 15$) = $-.79$, $p = .45$ suggested to translate the CEI-II and reevaluation of its psychometric properties. However, significant SAV with respect to required cut-off value indicated no need of further psychometric optimization for instrument.

Keywords. Semantic analysis, translation, validity, item analysis

It has been important to consider that the process of translating an instrument commonly required eliminating all forms of cultural biases and language related barriers and hindrances. The same process also helps in the optimization of the validity as well as reliability of

Shah Rukh Tariq and Iffat Batool, Department of Psychology, Government College University, Lahore, Pakistan.

Correspondence concerning this article should be addressed to Shah Rukh Tariq, Department of Psychology, Government College University, Lahore, Pakistan. Email: tariq.s.edu@gmail.com

instruments (Rode, 2005). Commonly, the translation process becomes a necessity when the source and target languages happen to be different (Harkness, Vijver, & Mohler, 2003). In terms of psychometric domains, the initial steps should be executed in a strategic manner, because any problems during the initial phases can result in negative effects on the entire assessment process. For this reason, it is more often said that assessment is a process in which each step has a role to play. One such technique to minimize the response biases in translation process is known as semantic analysis. Semantic analysis is a technique used to determine conceptual themes and meanings from the content (Landauer & Dumais, 1997). This method facilitates an objective understanding and evaluation of the entire process of translation. For this reason, semantic analysis can be employed as a specialized screening technique. In initial phase of the translation process, use of semantic analysis (or item analysis) can be helpful to eliminate the translation errors. Secondly, semantic equivalence of an instrument can be improved by adjusting significant indicators obtained from the process of semantic analysis or item analysis (Tariq, 2013; Tariq, Batool, & Khan, 2013). Translation of an instrument is not only the process to improve psychometric properties instead it is also related to other psychometric evaluation techniques such as reliability (r corrected and alpha coefficient), item discrimination, and difficulty index (Rode, 2005).

On the other hand, the use of item analysis is common for diagnosing items as well as for witnessing an improvement in the psychometric and other related properties in an instrument (Heo, Moser, Riegel, Hall, & Christman, 2005; Kashdan et al., 2009). In this regard, Muraki (1990) suggested the use of two parametric logistic (2PL) models for Likert type items and responses, and this process provides two primary parameters including: difficulty level of items and specific discrimination thresholds. Each of these are referred to as "a" and "b" parameters in a respective manner (DeVellis, 2012). It should be pondered that this model happens to be an extended version of the renowned classical testing theory. Commonly, item analysis has been applied to item with multiple choice responses instead of rating scales (Hambleton & Jones, 1993; Samejima, 1969). Therefore, it is difficult to apply this technique on psychological instruments (Chernyshenko, Stark, Chan, Drasgow, & Williams, 2001). However, in present era advance softwares can solve this problem (e.g., IRT Pro JMetrik, and SPSS). Item analysis can be computed for Likert type scales with the help of 2PL models. It is an appropriate technique to use after translation of the instrument (Rogelberg, 2002).

Another technique to improve psychometric properties of an instrument is semantic analysis (Tariq, 2013; Tariq et al., 2013). Several studies have supported use of Latent Semantic Analysis (LSA) approach in different contexts. For instance, it had been used in neuropsychological deficits for categorization of words (Laham, 1997), predicting learning from text (McNamara, Kintsch, Songer, & Kintsch, 1996), text comprehension (Foltz, 1996), simulating semantic priming (Landauer & Dumais, 1997), simulating word sorting, and relatedness judgment and synonym test (Landauer et al., 1998). Semantic equivalence process was also supported by some psychological researches, such as Mallinckrodt and Wang (2004) suggested dual-language, split half quantitative approach (split half items of source and target language filled by sample of bilinguals) and Tariq et al. (2013) approach to translate Curiosity And Exploration Inventory-II with the help of semantic analysis. In another way forward-backward translation (correlation between source and target language) was commonly used to demonstrate the semantic equivalence of translated instruments (Maneesriwongul & Dixon, 2004; Willgerodt, Kataoka-Yahiro, Kim, & Ceria, 2005). Present study developed a new way to utilize semantic analysis process as compared to Landauer's and other approaches.

Theoretically, latent semantic analysis was founded on the basis of learning and language models (Landauer & Dumais, 1997). According to learning and language perspective, the interesting and curious question was to know how people extract maximum possible knowledge from available information. In this regard, cognition was the main attention for psychologists, linguists, and social scientists (Chomsky, 1991; Jackendoff, 1992; Shapard, 1987; Vygostsky, 1962). The experts in cognitive perspective had proposed similar and unique theories of language acquisition, and among them, the general theory of similarity and generalization was noticeable (Chomsky, 1991; Jackendoff, 1992; Shapard, 1987; Vygostsky, 1962). However, this theory had problem which was called as induction problem. Moreover, social scientists came to unitary thought to cope this problem by understanding of the acquisition of language. For instance, Chomsky (1991) suggested that children disregard the language rules and provide the best model of underlying structure. In contrast, Saeed (2009) suggested that changing semantics or syntax affects the way one extract the meaning from information.

The core principal of LSA is that the language structures can be learned from language itself. In other words, Landauer, McNamara, Dennis, and Kintsch (2007) suggested that meanings can be extracted from the sentences without any prior rules and learning. In LSA, it

was not important that the meaning had been extracted based on prior learning, instinct or both, instead, main assumption was that how one solves the inductive problem of meaning making (Landauer et al., 2007). Therefore, it was regarded as a theory of memory, learning, and knowledge (Landauer & Dumais, 1997). Psychologically, the relationship between stimuli (words) and local context (meanings with respect to context) has importance for LSA model. Moreover, it is also evident that language acquisition rules can be applied to mathematical models, for example, singular value decomposition (Landauer & Dumais, 1997; Landauer et al., 2007). Therefore, based on experts' suggestions it can be said that semantic analysis had been an appropriate method to extract meanings from the information (Chomsky, 1991; Jackendoff, 1992; Landauer & Dumais, 1997). However, mathematical model of LSA has been still a debatable topic, and commonly criticized due to its extraction method of meanings than other semantic techniques (Flanagan, 2010). Further, it was important to consider how LSA extract meaning from a text, which has been described below.

Latent semantic analysis uses unique process of deriving contextual words from the raw text, and during the process, the data based knowledge (e.g., vocabulary in dictionaries, semantic networks, & grammar) was prohibited to use in semantic process (Landauer & Dumais, 1997). Landauer, Foltz, and Laham (1998) emphasized that in initial phase of latent semantic analysis, contingency table was computed to calculate the frequency.

Secondly, in latent semantic analysis singular value decomposition can be applied to the contingency form. For instance, the value obtained from decomposition of rectangular matrix (denoted by X) is equal to the product of three matrices, which are denoted by W , S and P . The LSA process in this phase is similar to the factor analysis that is, reduction of dimensions (Madsen, Hansen, & Winther, 2004). Two out of three component matrices reflect the row and column entities as vector of orthogonal factor values (dependent domains). The third component matrix is diagonal which describes scaling values (Landauer, Foltz, & Laham, 1998). In the final phase of LSA, interpretation of results can be done by evaluating correlation coefficients (Spearman correlation method) between the word (row) and context (column), and singular value decomposition. The obtained cosine value closer to one indicates similarity between the two parts, and in opposite case a value of cosine closer to zero indicates nonsimilarity (Jorge-Botana, Leon, Olmos, & Escudero, 2010). According to Landauer et al. (1998), LSA is a practical technique in

judgment and extracting human cognition to obtain significant components of the lexical and passage meanings.

Difference between LSA and Newly Developed Semantic Analysis Process

The present study has proposed a new method to conduct semantic analysis which can examine the efficiency of language translation (psychological instruments), and presents this efficiency in a statistical form. Following are some major differences and similarity between present semantic analysis process and Landauer's approach (see Table 1).

Table 1

Similarity and Differences between Landauer's LSA and New Semantic Analysis Approach

Factors	Characteristics
Words Similarity	Word to word relations (Landauer's approach); categorization of similar contextual meaning of the word (applied in the present study)
Formula	Statistical formula in Landauer's approach was correlation, while the new semantic analysis approach uses the test of difference, called as semantic analysis validity (see Equation 1).
Common use	Landauer's model is commonly used in discourse passages, while the new semantic analysis has been used in translation process of psychological instrument (items)
Domain Extraction	There are multiple dimensions in Landauer's approach. In contrast, newly proposed semantic analysis represents only two dimensions (named as favorable domain-item value, based on similar contextual meanings of the word in an item; average based or most frequently used words; and secondly, unfavorable domain-nonsimilar contextual meanings of the word in an item)
Utility	(a) The new semantic analysis could determine translation equivalence, while in Landauer's approach no bilingual context has been used. (b) It could assess the individual and overall translation efficiency, while Landauer's approach assesses the text in discourse passages. (c) The new approach is easy to calculate and interpretable (manual calculations are possible) as compared to Landauer's approach-machine based analysis.

Note. See method section for the new SA process.

The present study was an effort to establish the reliable and valid method which can measure the translation validity by using the inferential statistics. It is also evident that some of the commonly available techniques could assess semantic equivalence of the translated instruments (e.g. correlational approach). In contrast, new semantic approach can be done with small sample size, and without correlation between instruments of source and target language. Another reason was to develop the new semantic analysis technique because previous techniques only applied in language or discourse perspectives (Landauer, & Dumais, 1997; Landauer et al., 1998). And these techniques cannot provide the depth of analysis as much the new semantic analysis process. For instance, item analysis and correlation between the forms could only tell that particular item is not up to the standards. However, the new semantic analysis process provides word by word, item by item, and as well as overall index to screen psychological instruments (Tariq 2013, Tariq et al., 2013).

At a more general level, the process of test development, assessment, and application demand the use of objective, strategic, and reliable measures (DeVellis, 2012). There was also a need for a specific method to ensure that the process of initial screening always brings in the required results. Keeping this into perspective, semantic analysis validity was established. Furthermore, to achieve present study's objectives, a short instrument was required. The authors of the present study had experience regarding brief instruments (e.g., instrument with 10 or less items with sub-scales), which usually produce complexity in psychometric properties (Gosling, Rentfrow, & Swann, 2003; Tariq, 2013; Tariq et al., 2013). Therefore, Curiosity and Exploration Inventory-II (CEI-II) was selected conveniently as a study instrument. Moreover, cultural variations were also expected in CEI-II (Kashdan et al., 2009) because this instrument was not developed indigenously, and it was the second reason that CEI-II was selected as a present study instrument.

Objectives of the Study

For this research, the following were the objectives:

1. To explore, understand and evaluate the semantic analysis process from psychological perspectives that is, for the quantification of the entire translation process
2. To examine and determine the validity of the entire process (through the use of descriptive as well as inferential statistics)

3. To evaluate the psychometric and other related properties of the version translated in Urdu language
4. To explore and evaluate the item response perspective for the version in Urdu language

Method

Research Design

The present research was completed in three specific phases. For the first phase, the process of semantic analysis had been applied. During the second phase, the process of forward-backward translation had been completed. Moreover, the psychometric and other related properties and domains of the instrument were examined in close detail. For the third phase, item analysis and confirmatory factor analysis were subject to implementation.

Phase I

In phase one, two sample were used. Sample-I was used for the process of semantic analysis whereas sample-II was recruited to check the psychometric properties of the original CEI-II.

Sample. Convenience sampling was used to recruit participants from GCU, Lahore. Sample 1 consisted of 15 university students (young women = 10, young men = 5) with an age range from 19 to 21 years ($M = 20.7$, $SD = .07$). Sample 2 was consisted of 50 students (young women = 40, young men = 10) with an age range from 19 to 24 years ($M = 21.78$, $SD = 1.15$). In phase I, both samples were consisted of BS (Hons.) students. Minimum age was 18 years and at least education of 15 years, as an inclusion criterion for sample.

Instrument. In phase I, original English version of Curiosity and Exploration Inventory-II (CEI-II) was used (Kashdan et al., 2009). CEI-II encompasses a total of 10 items, with 5-point rating scale (*completely disagree* = 1 to *completely agree* = 5). It measures the trait curiosity with two subscales, named as Embracing and Stretching. The first subscale 'Stretching' includes 5 items (odd numbered items), and it is concerned with the search for knowledge as well as new ideas and experiences. The second subscale 'Embracing' includes 5 items (even numbered items), and it includes the levels to which uncertainties and ambiguities can be accepted. The scoring can be done through adding

up the scores for each of the subscales. The total score of instrument indicates the overall levels of curiosity in an individual. There was no specific cut-off score for this instrument. However, mean and standard deviation based cut-off score can be used to interpret the low and high curiosity levels. High score on CEI-II indicates superior curiosity trait.

Procedure. Firstly, permission was taken from the heads of respective departments and then questionnaire of semantic analysis was distributed among the students. For semantic analysis specific instructions were given to the participants i.e., read the statement carefully and then write the underlined word's meaning in Urdu. Therefore the meaning obtained from the word (indicator) was expected to be contextual rather than literal meaning. Secondly, in sample II curiosity and exploration inventory- II was administered. The participants were assured that their data would be used for research purposes only.

The Process of Semantic Analysis

The following phases and steps have been recommended for the process of conducting semantic analysis and these steps are (a) identification of the significant words in the item (b) conceptual or contextual meaning of the words rather than literal meaning (c) categorization of responses (d) compare three types of response (e) count percentages of favorable and other responses (Tariq, 2013). These steps are further explained.

Step 1. The first step involves the identification of specific meanings and words in items. The main point to consider in this phase is to identify and outline the problematic words. The selection process of the words can be based on the use of face validity. Furthermore, expert judgments and opinions of the targeted population segment can also be used. For this specific research, all three approaches were used. For instance, in item number seven "I am always looking for experiences that challenge how I think about myself and the world" the word "challenge" was expected to be problematic.

Step 2. In this step, the most important thing was to focus on the contextual as well as conceptual meanings of words. In other words, the literal meaning of the words or sentences holds no significance in this phase. This step is initiated by presenting all items to the participants and then asking them to write meanings of each of the words in their own language. For this study, the participants were explicitly asked to write the meanings of words in Urdu, as it was their

primary language. For example, CEI-II item no. seven showed that the word “challenge” obtained the four similar meanings (“*Mukabla*”) out of total fifteen responses. Other responses were quite different in meanings such as “*Muskil Kam*”, “*Pecheda*”, “*Dawa*” and one missing response was observed. It should be noted that there are times in which literal meanings are mixed with conceptual and contextual meanings for entire sentences. It happens because the local meanings are not made accessible from the side of the target population.

Step 3. This step is concerned with response categorization. It should be noted that this phase can be termed the core component of the entire semantic analysis process. It involves making categories of favorable as well as non-favorable responses, that is, the meanings of the words in Urdu. For this step, screening based on the use of statistics happens to be helpful. The statistical methods that might be used here include: determining the frequency of words, word to word discrimination that is, and homogenous categorization or through the use of categorical analysis. For present study, the use of words frequency method turned out to be the most optimal. It also ensures an easy interpretation. The next level of categorization of words is based on expert opinion and judgments keeping in view the face validity. Another point to ponder in terms of scoring is the specific categorization of a number or multiple indicators. All of the items which are marked by multiple indicators need to be scored in a careful manner. For example, if an item happens to be have indicators and if anyone has received a non-favorable, that is, zero score, then the score of the entire item will be considered “zero”. Alternatively, item value obtained by the average of multiple indicators can be a better solution (e.g., item 10). It has been suggested that the responses are to be scored “1” for favorable and “0” for non-favorable. However, this phase has some technicalities and complexities which should be considered. Moreover, this step should always be executed in a careful manner in order to avoid any problems during the later phases of the entire process.

Step 4. This step involves the comparison between three types of responses. For this stage, possible categorizations were: focusing on the significant meanings of all of the words, focusing on the nonsignificant meanings and no-response categories. Each of these categories were made through a detailed analysis of the responses. For instance, in item no. seven, four responses were favorable, and eleven responses were non-favorable.

Step 5. This step involves the process of determining the % favor and % other. In order to determine the % favor, it is imperative to identify and outline homogenous meanings. The obtained value is

then to be divided by the total no. of responses. For the other domain, the main focus should be on the nonsignificant meanings together with all other missing responses. The obtained item values are actually the favorable responses reported as percentage. For instance, item number seven obtained 27% favorable response, and 73% as non-favorable response in the SA process.

Individualistic Semantic Analysis Values

The process of semantic analysis yields a great deal of important information such as the item values for each participant. This value can be determined by adding up the favorable responses and then dividing them with the total no. of responses. It could help to identify those individuals who have poor ability to respond on the test.

Validity of the Semantic Analysis (SAV)

For the validation of the entire process of semantic analysis, the following formula dimensions should be considered. The favorable responses mean (Fm) can be determined by computing mean of the item values. On the other hand, the nonfavorable responses mean can be determined by subtracting favorable responses mean from one ($1 - Fm$). For the inferential testing of semantic analysis, the following formula was developed:

$$SAV = \frac{Fm - \Delta}{s/\sqrt{k}} \quad \text{Eq. 1)}$$

The new formula can be used for determining the semantic analysis validity. It happens to be similar to the one sample t test. However, it should be noted that there are some distinct notions and values which make it different (the symbol k represents number of items, and s is the standard deviation of the item values). The cut off range for SAV could be .90 to .70 for already validated instruments, and for new instruments .50 to .55 is optimal (Tariq, 2013). In terms of this research .70 (value of delta) was used as the criterion value.

Results of Phase-I

The main objective of the present study was to demonstrate newly developed semantic analysis process, and for this purpose CEI-II (Kashdan et al., 2009) was used. Descriptive statistics were used for this study in order to report important statistics for the CEI-II. For

computing the analyses, SPSS, MS Excel, AMOS and IRT Pro (student version) were used.

Table 2

Semantic Analysis for Curiosity and Exploration Inventory-II

Item no. CEI-II	Original indicator	Sample one ($n = 15$)		Item value	r	Alpha if item deleted
		% favor	% other			
1	Actively	73%	27%	.73	.16	.65
2	a. Type of person b. Uncertainty	73%	27%	.73	.44	.58
3	a. Complex b. Challenging	73%	27%	.73	.16	.65
4	Out looking	87%	17%	.87	.46	.59
5	Opportunity to grow	73%	27%	.73	.28	.62
6	Frightening	80%	20%	.80	.36	.61
7	Challenge	27%	73%	.27	.13	.65
8	Excitingly unpredictable	53%	47%	.53	.43	.59
9	Grow as a person	80%	20%	.80	.54	.57
10	a. Embraces b. Unfamiliar c. Event	27%	73%	.27	.16	.66

Note. r = item-total correlation (corrected), item value = ratio of favorable responses; % favor reflects the ratio of contextual meanings in an item. For multiple indicators single item value was presented.

The new semantic analysis has been used to statistically examine the contextual meanings of items. The cut off value was .50 for the item value. The results indicate that the items 7 and 10 are problematic in terms of the conceptual understanding and comprehension of meanings of each of the items. In all, a total of 2 items are found to be below the specific cut off value (see Table 2).

Item Value for Individuals

It is found that students' item values range from .30 to 1.0 ($n = 15$). Moreover, the cut-off value used for the identification of good and poor responses happened to be the same as item values.

Thus, on the basis of the cut off value, five participants did not exhibit an accurate understanding of the meanings of specific words and items.

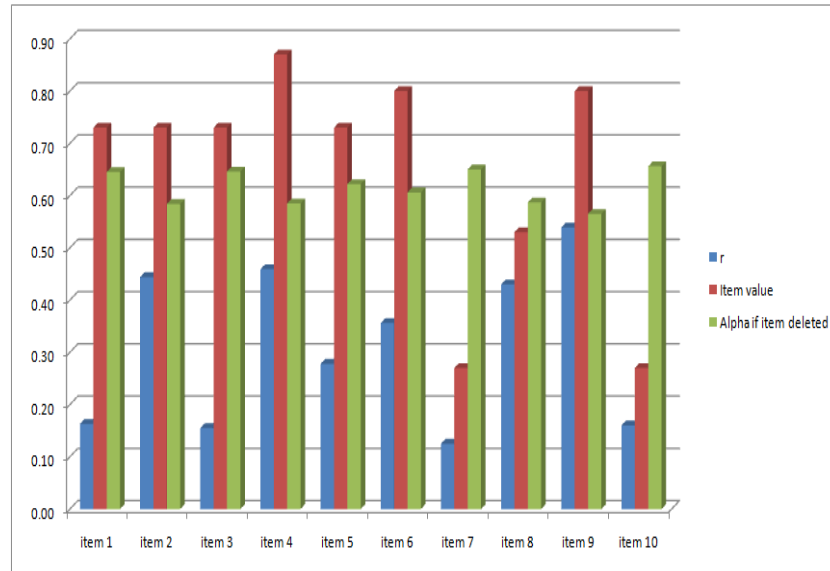


Figure 1. Comparison among alpha value, item-total correlation (r) and item values.

The Figure 1 show that items no. 1, 3, 7, and 10 demand improvement. The value for the item-total correlation was determined to be $r = .2$. However, the value of $r = .3$ can be used or might be optimal for larger samples (Field, 2009). A higher alpha value points out need improvement while lower values point to satisfactory item performance. It should be noted that overall alpha values are not expected to fall below .50 (George & Mallery, 2006). It was also found that the sample-II's ($n = 50, k = 10$) computed $\alpha = .64$. After the exclusion of item number 1, 3, 7, and 10 (i.e., indicated in SA process), the value was computed to be .72 ($k = 6$).

Validity of Semantic Analysis

For Equation 1, the determined item values were added to SPSS. One sample t -test was the option selected for the analysis. The criterion value selected for the command was $\Delta = .70$. It has been determined that semantic analysis validity, $SAV = -.79$, 95% $CI = -.20$ to $.10$, $p = .45$, offers an indication of the need for translation (if $Fm <$

Δ). The resultant favorable mean needs to be greater than the overall test value. The mean calculated for the favorable responses is helpful in suggesting the overall levels of meaningfulness of the items as perceived from the side of the target population. The non-significant value of *SAV* suggested that the instrument need improvements in the identified items by SA process.

Phase II

The objectives of phase II was to perform process of forward-backward translation and then checking the psychometric properties of the Urdu version of CEI-II.

Forward and backward translations. For the second phase, an important facet to consider was the translation of the instrument. Semantic analysis was used for the purpose of determining the understanding and conceptualization of the target population for the level of items. It has been thus found that there is a need of engaging in translation of the instrument (see Table 2). It needs to be considered that the main language of the population was Urdu due to which the conceptualization of items and meanings might be influenced. For the purpose of forward translation into Urdu, two bilingual experts were contacted. One of them had done his M.Phil in Urdu while the second one was a Masters in English literature. They were specifically asked to engage in translation on the basis of conceptual equivalence rather than just focusing on the available literal meanings. For the purpose of keeping the Urdu as simple and understandable as possible, precise as well as simple issues related to equivalence were discussed thoroughly with each expert. For the process of backward translation, the same procedure was employed. Then, two other bilingual experts were asked to translate the obtained version from Urdu into English language. It was made sure that these experts had no prior knowledge about the instrument which is important from translation and research concerns. For the final phase of approval of the instruments, the versions were presented to the Department Committee. The translated versions were subject to comparisons and evaluations keeping in view the cultural contexts and conceptual equivalence. Moreover, surplus and other misfit words had to be removed from the items. These words were replaced as the semantic analysis had advised improvements. After the translation process had been completed, the reliability of the versions was examined (see Table 4 and 5).

Sample. In phase II, sample-III consisting of 50 university students (young women = 33, young men = 17) was recruited by using

convenience sampling technique, and participants were recruited from various universities of Lahore. The age range of participant was 19 to 24 years ($M = 21.70$, $SD = 1.14$), and they were students of BS (Hons.).

Instrument. In phase II, Urdu translated Curiosity and Exploration Inventory-II (CEI-II) was used (see details in phase I).

Procedure. The permission of higher authorities and consent from was taken in advance for the data collection. In phase II, Urdu translation of CEI-II was carried out, and psychometric properties were compared between the Urdu and English version of the curiosity and exploration inventory-II

Results of Phase-II

Table 3

Curiosity and Exploration Inventory-II Psychometric Properties (Urdu version)

Scale	k	α	Actual	Range		Skew
				Potential		
CEI-II	10	.80	1-5	1-5		-.83

Note. Sample III ($n = 50$)

The Cronbach's alpha was found satisfactory after translation of the instrument in Urdu language (see Table 3). It was also noticeable at this stage that original Curiosity and Exploration Inventory-II had lower alpha coefficient ($\alpha = .64$) as compared to the translated version ($\alpha = .80$).

Phase III

For the third phase, two samples were used: Sample-IV ($n = 250$) was recruited to check the item analysis whereas sample-V ($n = 150$) was recruited to run confirmatory factor analysis of CEI-II Urdu version.

Sample. In phase III, convenience sampling was used, and participants were recruited from various universities of Lahore. Sample-IV consisted of 150 students (young women = 110, young men = 40) with an age range from 19 to 24 years ($M = 21.01$, $SD = 1.03$). In contrast, sample-V ($n = 250$, women = 137, men = 113)

was consisted of age range 18 to 24 years ($M = 21$, $SD = 1.14$), and they were students of BS (Hons.).

Instrument. In phase III, Urdu translated version of CEI-II was used (see details in phase I).

Procedure. Permission from higher authorities and consent was taken before data collection from the participants. For phase III, confirmatory factor analysis on Urdu version of CEI-II was computed, and further item analysis was administered (see Result section). It should be considered that between samples II to V, less than 2% missing responses had been observed. In order to eliminate the possibility of experiencing any problems and errors in the phases, all of the missing responses were adjusted in relevance to the mean series (Rubin, Witkiewitz, Andre, & Reilly, 2007).

Results of Phase-III

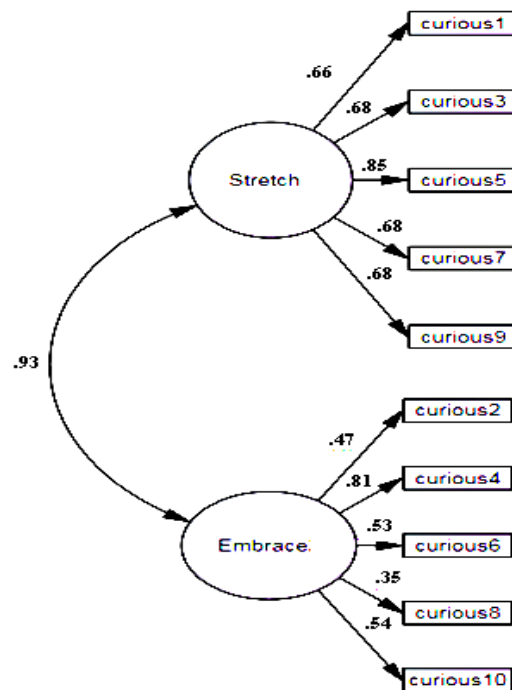


Figure 2. Complete standardized solution of CEI-II in fourth sample ($n = 150$).

Confirmatory factor analysis of CEI-II Urdu version showed that the RMSEA (root mean square error of approximation) value was .086

while the goodness of fit index value was .91. The CFI (comparative fit index) value was calculated to be .93 together with χ^2 (34, $n = 150$) = 71.39, $p < .01$. The normed fit index was .90. In terms of the overall indices, it is evident that this model was acceptable (Byrne, 2010).

Table 4

Psychometric Properties of Curiosity and Exploration Inventory-II (CEI-II)

Scale	k	M (SD)	α	Skew
Overall curiosity	10	36.90 (7.20)	.81	-.61
Stretch	5	19.85 (4.13)	.80	-1.00
Embrace	5	16.96 (4.14)	.68	-.32

Note. $N = 250$ (fifth sample). k = number of items.

The Table 4 shows the values of standard deviation, Cronbach's alpha, mean and skewness. The values for the Cronbach's alpha were found satisfactory (α range was .68 to .81), and the calculated values of skewness were acceptable (Tariq, 2011; Tariq, 2013).

Table 5

Item Analysis for Curiosity and Exploration Inventory-II

Items	M (SD)	a	b1	b2	b3	b4	Top two ratings
1	4.03(1.10)	1.64	-2.40	-1.74	-1.12	.37	4 (38.4%) and 5 (40.4%)
2	3.35(1.29)	.71	-2.89	-1.63	-.34	2.28	4 (36.8%) and 5 (18.8%)
3	4.10(1.01)	1.46	-2.89	-2.2	-1.31	.39	4 (40.4%) and 5 (40.4%)
4	3.87(1.15)	1.90	-2.05	-1.54	-.76	.55	4 (38.4%) and 5 (34.0%)
5	4.00(1.15)	2.72	-1.85	-1.42	-.83	.24	4 (38.4%) and 5 (34.0%)
6	3.04(1.30)	.94	-1.94	-.83	.34	.23	3 (22.8%) and 4 (29.6%)
7	3.80(1.21)	1.92	-2.03	-1.35	-.58	.51	4 (32.0%) and 5 (35.2%)
8	3.09(1.18)	.60	-3.75	-1.53	.92	3.35	3 (32.8%) and 4 (23.6%)
9	4.02(1.09)	1.47	-2.69	-1.86	-1.16	.41	4 (38.4%) and 5 (30.4%)
10	3.60(1.34)	.90	-2.61	-1.55	-.83	1.12	4 (34.8%) and 5 (30.4%)

Note. $N = 250$; a = parameter reflects discrimination; b = parameter reflects item difficulty. IRT with two parametric graded-response (2PL) model.

Table 5 shows the results of the items analysis for the instrument. It was found that item number 1, 3, 4, 5, 7, and 9 are marked for high discrimination values. At the same time, items 2, 6, and 10 are marked for having moderate and item 8 for its low discrimination value. In order to indicate the thresholds for the resultant item difficulty, item no. 5 had been indicated in the category response curve. According to Kashdan et al. (2009, p. 11) “these difficulty thresholds represents the trait levels at which someone has a 50% chance of scoring at or above a scale response”. In present study item number seven has a value of -2.03 (b1), which indicates that an individual who happens to be a low trait of curiosity is more likely to select from the options 5, 4, 3, or 2 (see Table 5). The value of b1, at the same time, suggests that an individual with curiosity trait standing at -2.03 might select the response 1. The value of b2 suggests that an individual who happens to have a low curiosity trait has almost 50 % chance of selecting responses from 5, 4 or 3. The b3 value is indicative of a chance of selection of 5 or 4. Furthermore, the b4 value is indicative of a chance of selecting response 5. Therefore, it could be suggested that item seven is an easy item, and similarly other items of the instrument can be interpreted in this manner.

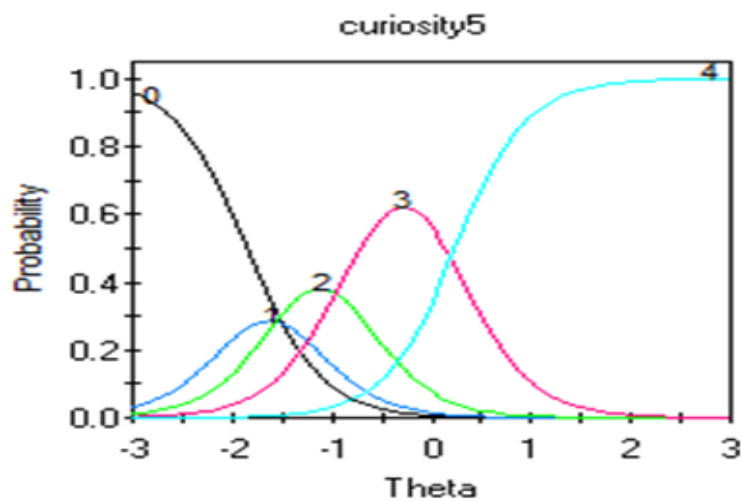


Figure 3. Category response curve (CRC) for item 5. The normalization of latent trait (presented on x-axis, theta symbol) and probabilities level presented on y-axis.

The complete descriptions of CRC for total items was not possible, and thus for demonstration purpose only item 5 was explained. Baker (1985) suggested that probabilities level change with

the function of latent trait in category response curve. In Figure 3, most of the responses were appeared slightly left side of the distribution. It was also found that participants mostly answered with a response of 4, and it could be assessed by the peakness of curves (see Table 5). Moreover, most of the responses were found between -1.5 and -.5 area of the distribution.

Table 6

Percentile Scores for Curiosity and Exploration Inventory-II Urdu Version (N = 250)

Percentile	Subscales		
	Overall	Stretching	Embracing
5	23.00	12.00	10.00
25	33.00	17.00	14.00
50	38.00	21.00	17.00
75	42.25	23.00	20.00
95	47.00	25.00	23.00

Table 6 shows the percentile and raw scores for the CEI-II. The percentiles scores present the relative position of each of the individuals (Cohen, 2001). In terms of psychological assessment, common percentiles are reported to be around 25th to 95th. Similarly, Anastasi and Urbina (1997) suggested that individuals could be divided according to their level of traits or abilities into three categories i.e. high, average and low. For this instrument, someone who achieves a score of 38 will stand at the 50th percentile (for overall curiosity level).

Discussion

The prime intent behind the use of semantic analysis in the present study was to facilitate the translation process. Secondly, the present study evaluated the semantic validity, psychometric properties, and item analysis statistics produced under the new semantic process for Urdu translated version of CEI-II (Kashdan et al., 2009). The major question was that how one can capture the cognitive domain of human made translation process. In this regard, semantic analysis can be compromised as a suitable technique (Tariq et al., 2013). However, the previous technique of semantic analysis was solely developed for the discourse purposes (Landauer, & Dumais,

1997; Landauer et al., 1998), and therefore, the new technique was required for translation purpose. It was also required to understand the semantic analysis in a psychological perspective.

The first objective of the study was achieved by translation of CEI-II and quantification of the semantic process (see Table 2). This process has been called as psychological because cognitive domain of the participants was evaluated in this phase, and particularly at this stage one can examine how much deviations in the translation process was present (see Table 3 and 4). In contrast, social researchers used correlation between the instruments of target and source language (Maneesriwongul & Dixon, 2004; Nosheen & Jami, 2013; Willgerodt et al., 2005) for the cross validation of instruments, but this method had its own limitations. For instance, correlation coefficient obtained by cross-validation did not provide details about words or part of an item which lacks the understanding in term of semantics. In such case psychological instruments need some alternative strategy to cover this gap. The present study was the first attempt to evaluate this process in a new way. Therefore, first necessary part to eliminate cultural and language biases was to translation or semantic analysis process (Rode, 2005, Tariq et al., 2013).

After semantic analysis process, next step was to understand validity of the whole semantic process. Therefore, second objective of the study was to validate the whole semantic process. This objective was successfully achieved due to the Equation 1. Secondly, in common practice newly developed or translated instruments required factor analysis for verification of the latent structure (DeVellis, 2012), and thus confirmatory factor analysis was conducted for CEI-II Urdu version. In present study, phase II supported second objective regarding Urdu translated CEI-II factor structure (see Figure 2). Findings indicated that the present study had acceptable model fit values of confirmatory factor analysis for Urdu version of CEI-II (Byrne, 1994). There was slight disagreement between the GFI, CFI, NFI, RMSEA in the present study, but they were overall acceptable (Kenny, 2014; Lei and Wu, 2007). These findings were consistent with other CFA models of CEI-II (Kashdan et al., 2009; Tariq, 2013, Tariq et al., 2013).

The third objective of the study was achieved by examining the mean, standard deviation and alpha confident (i.e., psychometric properties) of the CEI-II Urdu version. In present study, it was found that initially some items has problematic reliability but after using of semantic analysis and translation process they become optimized (see Table 4). In social sciences commonly alpha coefficient of .70 is used as a set standard for a reliable domain, but alpha coefficient of .50

could be used as minimum criterion of acceptable alpha coefficient (George & Mallery, 2006). The fourth objective of the study was achieved by examining the item analysis of the Urdu version of the CEI-II. It was necessary because it gives the items by item evaluation of the instrument. And these values could be compared in term of semantic analysis. For instance, it was expected that low difficulty level of items could have high item-value (see Figure 1).

Alternatively, a close review of item discrimination shows that all of the item scores change in response of variation in the latent traits. At the same time, the values describe the accuracy to which these items reflect the trait that needs to be measured (Kashdan, et al., 2009). Moreover, phase III findings (item analysis) supported that phase I and II psychometric properties was fine (Baker, 1985), and all items were in the range of acceptable discrimination threshold (Nevo, 1980). Another interesting point of the item difficulty was that it reflects the individual differences (low, moderate and high responses), and in present study these differences has been observed (see Table 5). It was also evident from the present study that item analysis combined with the new semantic analysis can yield more accurate efficiency of psychological instrument (Tariq, 2013; Tariq et al., 2013). Kashdan et al. (2009) findings also support the fourth objective.

Overall, present study emphasized that at the initial stage of instrument translation semantic analysis can be helpful. Secondly, the entire process of semantic analysis can be evaluated by conducting semantic analysis validity coefficient (SAV) and factor analysis. For already validated instrument confirmatory factor analysis (CFA) was suitable instead of exploratory factor analysis (DeVellis, 2012). Thirdly, after examining and comparing original (source) and translated version of an instrument one can use more advanced technique to evaluate the efficiency of psychometric properties such as item difficulty and discrimination index along with percentiles. The percentiles can determine the cut-off score for the newly developed or translated instrument. Percentiles also helped in interpretation of the scores. All the process was interlink to each other. For instance (a) semantic analysis values were comparable with reliability coefficients (see Figure 1) (b) reliability provided item-total correlation and alpha coefficients which were interlink with item analysis i.e. indicated acceptable and non-acceptable item (see Table 2) (c) psychometric properties (high or low alpha coefficient) might or might not be related to better or poor factor loadings (no item was discarded as item analysis suggested; see Table 5). In other words, a reliable measure could be valid but a valid measure not necessary could be reliable (Wooldridge, 2003).

Limitations and Suggestions

Semantic analysis validity cannot be efficiently tested if all of the item values turn out to be constant. Therefore, for testing semantic analysis validity, at least one of the favorable responses needs to be different from the other ones. Alternatively, computed mean of all the item values are expected to be greater than the specific cut off value. Secondly, Semantic analysis validity was based on favorable responses and categorization of the varying homogenous responses. Therefore, validity of categorization should be considered effectively in order to witness an optimal performance in terms of the semantic analysis. Large samples can be used in different scenarios keeping in view the guidelines of the experts (e.g., 30 or more). Another limitation of the study was the selection of small sample size in terms of the item response theory. However, in some cases small samples can be used for item analysis (Maydeu-Olivares, Drasgow, & Mead, 1994; Nevo, 1980). In future studies, researchers should examine the pre and post analysis of the semantic analysis validity (i.e., compute *SAV* before and after instrument translation). The pre and post analysis could provide better estimate of the semantic equivalence, and it was not included in the present study.

Conclusion

There was a need for a specific method to ensure that the process of initial screening always brings in the required results. Results of the present study indicated that semantic analysis can be used effectively for translation process. It was indeed one of the useful techniques that aid in a statistical measuring of the understanding level for all of the items under focus. At the same time, this method also comes in handy for the purpose of identifying varying levels of individual differences. It has now been empirically tested and validated that semantic analysis was useful and can improve psychometric properties of an instrument. In present study, nonsignificant value of *SAV* suggested to reevaluate psychometric properties of original CEI-II. However, significant *SAV* with respect to required cut-off value indicate no need of further psychometric optimization and translation.

Implications

The main implication of semantic analysis was for translation related purposes. In other words, it can be used for examining and

increasing translation validity. It was evident in the present findings that it can be used in educational and testing psychology domains for evaluation of the items. However, it should be noted that the semantic analysis for this study is a relatively new technique which raises the need for further and higher levels of empirical testing. The findings of semantic analysis validity can be generalized for the instrument, because it was tested with the help of inferential statistic i.e. the significant mean difference between the mean of items value (F_m) and test value. The inferential test suggests that obtained significant difference was not by chance. However, it can be used before and after the translation process to see the performance of translation process.

References

- Anastasi, A., & Urbina, S. (1997). *Psychological testing*. New Jersey: Prentice Hall.
- Baker, F. B. (1985). *The basics of item response theory*. Portsmouth: Heinemann.
- Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows*. Thousand Oaks: Sage Publications.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New Jersey: Lawrence Erlbaum Associates.
- Chernyshenko, O. S., Stark, S., Chan, K. Y., Drasgow, F., & Williams, B. (2001). Fitting item response theory models to two personality inventories: Issues and insights. *Multivariate Behavioral Research*, 36(4), 523-562.
- Chomsky, N. (1991). Linguistics and cognitive science: Problem and mysteries. In A. Kasher (Ed.), *The Chomskyan Turn* (pp. 26-53). London: Blackwell.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309-319.
- Cohen, B. H. (2001). *Explaining psychological statistics*. New York: John Wiley & Sons.
- DeVellis, R. F. (2012). *Scale development: Theory and application* (3rd ed.). London: Sage Publications.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.
- Flanagan, O. (2010). *The really hard problem: Meaning in a material world*. Massachusetts: MIT Press.

- Foltz, P. W. (1996). Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments, and Computers*, 28, 197-202.
- George, D., & Mallery, P. (2006). *SPSS for Windows step by step: A simple guide and reference, 13.0 update*. Boston: Pearson.
- Gosling, S. S., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-528.
- Hambleton, R. K., & Jones, R. W. (1993). *Comparison of classical test theory and item response theory and their applications to test development*. Retrieved from <http://ncme.org/linkserverid/66968080-1320-5CAE-6E4E546A2E4FA9E1/showMeta/0/>
- Harkness, J. A., Vijver, F. J. R., & Mohler, P. P. (2003). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Heo, S., Moser, D. K., Riegel, B., Hall, L. A., & Christman, N. (2005). Testing the psychometric properties of the Minnesota Living with Heart Failure Questionnaire. *Nursing Research*, 54(4), 265-272.
- Jackendoff, R. (1992). *Languages of the mind: Essays on mental representation*. Massachusetts: MIT Press.
- Jorge-Botana, G., León, J. A., Olmos, R., & Escudero, I. (2010). Latent semantic analysis parameters for essay evaluation using small-scale corpora. *Journal of Quantitative Linguistics*, 17(1), 1-29.
- Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The Curiosity and Exploration Inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, 43(6), 987-998.
- Kenny, D. (2014). *Measuring model fit*. Retrieved from <http://davidakennedy.net/cm/fit.htm>
- Laham, D. (1997). *Latent semantic analysis approaches to categorization*. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the 19th Annual Conference of the Cognitive Science Society* (p. 979). New Jersey: Lawrence Erlbaum Associates, Inc.
- Landauer, T. K. (1999). Latent semantic analysis: A theory of the psychology of language and mind. *Discourse Processes*, 27(3), 303-310.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211-240.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, 259-284.
- Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis*. New Jersey: Lawrence Erlbaum.
- Lei, P. W., & Wu, Q. (2007). *Introduction to structural equation modeling: Issues and practical considerations*. Retrieved from <http://ncme.org/linkserverid/47EFEB5A-1320-5CAE-6EC90BFDF09AA39E/showMeta/0/>

- Madsen, R. E., Hansen, L. K., & Winther, O. (2004). *Singular value decomposition and principal component analysis*. Retrieved from http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/4000/pdf
- Mallinckrodt, B., & Wang, C. C. (2004). Quantitative methods for verifying semantic equivalence of translated research instruments: A Chinese version of the experiences in close relationships scale. *Journal of Counseling Psychology, 51*(3), 368-379.
- Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: A methods review. *Journal of Advance Nursing, 48*(2), 175-186.
- Maydeu-Olivares, A., Drasgow, F., & Mead A. D. (1994). Distinguishing among parametric item response models for polychotomous ordered data. *Applied Psychological Measurement, 18*(3), 245-256.
- McNamara, D. S., Kintsch, E., Songer B. N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement, 14*(1), 59-71.
- Nevo, B. (1980). Item analysis with small sample sizes. *Applied Psychological Researches, 4*(3), 323-329. Retrieved from <http://conser.vancy.umn.edu/bitstream/100192/1/v04n3p323.pdf>
- Nosheen, I., & Jami, H. (2013). Translation, adaptation, and validation of contraceptive attitude scale. *Pakistan Journal of Psychological Research, 28*(2), 335-359.
- Rode, N. (2005). Translation of measurement instruments and their reliability: An example of job-related affective well-being scale. *Methodological Notebooks, 2*(1), 15-26.
- Rogelberg, S. G. (2002). *Handbook of research methods in industrial and organizational psychology*. London: Blackwell Publishers.
- Rubin, L. H., Witkiewitz, K., Andre, J. S., & Reilly, S. (2007). Methods for handling missing data in the behavioral neurosciences: Don't throw the baby rat out with the bath water. *Journal of Undergraduate Neuroscience Education, 5*(2), 71-77.
- Saeed, J. I. (2009). *Semantics* (3rd ed.). London: Blackwell Publication.
- Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores* (Psychometric Monograph No. 17). Richmond, VA: Psychometric Society. Retrieved from <http://www.psychometrika.org/journal/online/MN17.pdf>
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*(4820), 1317-1323.
- Tariq, S. (2011). *Identity styles, causality orientation and psychological well-being among adolescents: A cross-cultural comparison* (Unpublished

- doctoral dissertation). National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.
- Tariq, S. (2013). *Students' curiosity, self-regulation, and cognitive styles with academic achievement* (Unpublished MPhil dissertation). Department of Psychology, Government College University, Lahore, Pakistan.
- Tariq, S., Batool, I., & Khan, T. S. (2013). Curiosity, self-regulation and academic achievement among graduate students. *Pakistan Journal of Social and Clinical Psychology*, *11*(2), 28-35.
- Vygotsky, L. S. (1962). *Thought and language*. New York: Institute of Technology Press & John Wiley.
- Willgerodt, M. A., Kataoka-Yahiro, M., Kim, E., & Ceria, C. (2005). Issues of instrument translation in research on Asian immigrant populations. *Journal of Professional Nursing*, *21*(4), 231-239.
- Wooldridge, J. M. (2003). *Introductory econometrics: A modern approach*. Mason, Ohio: South-Western Publishers.

Received October 9th, 2014

Revision received July 11th, 2016