International Review of Research in Open and Distributed Learning

An Item Response Theory Analysis of the Community of Inquiry Scale

Mehmet Barış Horzum et Gülden Kaya Uyanik

Citer cet article

Résumé de l'article

The aim of this study is to examine validity and reliability of Community of Inquiry Scale commonly used in online learning by the means of Item Response Theory. For this purpose, Community of Inquiry Scale version 14 is applied on 1,499 students of a distance education center's online learning programs at a Turkish state university via internet. The collected data is analyzed by using a statistical software package. Research data is analyzed in three aspects, which are checking model assumptions, checking model-data fit and item analysis. Item and test features of the scale are examined by the means of Graded Response Theory. In order to use this model of IRT, after testing the assumptions out of the data gathered from 1,499 participants, data model compliance was examined. Following the affirmative results gathered from the examinations, all data is analyzed by using GRM. As a result of the study, the Community of Inquiry Scale adapted to Turkish by Horzum (in press) is found to be reliable and valid by the means of Classical Test Theory and Item Response Theory.

An Item Response Theory Analysis of the Community of Inquiry Scale

Mehmet Barış Horzum and Gülden Kaya Uyanık
Sakarya University, Turkey

Abstract

The aim of this study is to examine validity and reliability of Community of Inquiry Scale commonly used in online learning by the means of Item Response Theory. For this purpose, Community of Inquiry Scale version 14 is applied on 1,499 students of a distance education center's online learning programs at a Turkish state university via internet. The collected data is analyzed by using a statistical software package. Research data is analyzed in three aspects, which are checking model assumptions, checking model-data fit and item analysis. Item and test features of the scale are examined by the means of Graded Response Theory. In order to use this model of IRT, after testing the assumptions out of the data gathered from 1,499 participants, data model compliance was examined. Following the affirmative results gathered from the examinations, all data is analyzed by using GRM. As a result of the study, the Community of Inquiry Scale adapted to Turkish by Horzum (in press) is found to be reliable and valid by the means of Classical Test Theory and Item Response Theory.

Keywords: Community of inquiry; social presence; teaching presence; cognitive presence; IRT

Introduction

Nowadays, online learning has become one of the most common applications used in distance learning. 75.9% of the institutions which have 7.1 million students taking at least one online course reported that online learning is critical as a long-term strategy (Allen & Seaman, 2013). There is a need to effectively plan, implement, and manage online learning which is a highly common application (Moore & Kearsley, 2012). To provide these, Garrison, Anderson and Archer (2000) designed Community of Inquiry (CoI) model as a guide to the provision of effective teaching in online learning studies and applications and to the qualities of learning outcomes. While CoI model helps to organize a theoretical frame of learning process in online learning environments (Garrison, Anderson & Archer, 2001), it displays the quality of a basically ideal education experience. Multi-elements like cooperation between the participants, interaction, and observable instructional indicators supporting inquiry are described within this experience (Bangert, 2009).

CoI Model and Components

The framework of CoI model was designed on the idea that the emergence of collaborative information configuration in online learning would occur through a community of inquiry (Shea, 2006). In this regard, it is a process-oriented model (Arbaugh, 2008; Arbaugh, Bangert, & Cleveland-Innes, 2010). CoI model highlights the significance of the fact that in order to enable sustainable deep learning and learning outcomes like critical inquiry in online learning environments, social interaction is not sufficient on its own, unless it is supported and integrated with cognitive and instructional elements (Garrison, Anderson, & Archer, 2000).

Within the framework of CoI, an environment to help reach a common point, diagnose misunderstood points, and enable responsibility in learning is created (Garrison & Anderson, 2003). The focus in CoI model is on the presence and belonging to a group (Joo, Lim, & Kim, 2011). The model includes three components, cognitive (CP), social (SP), and teaching presence (TP), which are emphasized as being the significant factors for the formation of a community.

One of the components of the model is TP, which is a key factor in terms of online teaching skills of the model (Garrison & Arbaugh, 2007), methods of the instructors (Bangert, 2009), and the behaviors necessary for creating a productive community (Shea, Li, Swan, & Pickett, 2006). By creating an appropriate teaching environment, collaboration of active participants in the community of teaching is aimed in the TP. Hence, design, facilitation, and direct instruction categories (Anderson, Rourke, Garrison, & Archer, 2001) are prioritized for online learning. The focus points are direct teaching activities like teaching process including the use of online learning tools, subject matter and outcomes, design and organization of the learning activities and tasks in the design category; providing participation of the learners, focusing on new terms, and simplifying the discussions to have consensus in the facilitation category; and relating the information from different sources, resolving misconceptions, and providing feedback in the direct instruction category (Shea, Fredericksen, Pickett, & Pelz, 2003). Not only the TP, but online

agents and automatic messages are also prioritized in online learning environments (Joo, Lim, & Kim, 2011). TP will be able to construct a bridge for the transactional distance between learner and instructor (Arbaugh & Hwang, 2006) by ensuring a deep learning process for the learner.

SP, another component of the model occurs with the learner's feeling of belonging to a course when taking part in online learning activities (Picciano, 2002). For SP, interaction is not enough on its own; thus, it should be supported with various interactional activities like developing inter-personal relationships and communicating purposefully in a trusting environment (Garrison, 2007). Emotional expression, open communication, and group cohesion are required to possess SP in online learning environments. The focus points of these categories can be stated as follows: emotions supporting inter-personal communication and relationships, use of affective response for humor and self-disclosure in the emotional expression category; recognition, encouragement, and interaction in the open communication category; addressing by name, salutations and use of inclusive pronouns in the group cohesion category (Garrison & Anderson, 2003). SP can be directly affected by the environment it is in and the tools used. SP of learners can decrease in a text-based environment, an agent application in which real people do not take place, or not individualized automatic text messages. Learners realizing their presence in the environment will create a positive effect and contribute to constructing a cooperative environment.

CP, another component of the CoI model, reflects learning and inquiry processes (Garrison, Cleveland-Innes, & Fung, 2010). The term CP based on the notion of Dewey's Practical Inquiry indicates critical and creative thinking process (Shea, Hayes, Vickers, Gozza-Cohen, Uzuner, Mehta, Valtcheva, & Rangan, 2010); thus, it is the reflection of meta-cognitive skills (Garrison, & Anderson, 2003). CP consists of four categories: triggering event, exploration, integration, and resolution. In the category of triggering event, a chaotic situation is created with the help of a problem to start an inquiry process. Exploration category includes generation of knowledge by brainstorming, clarifying the chaos and defining the problem. Sharing the generated knowledge, exchanging ideas, and reaching consensus are taking place in the category of integration. For the final phase, the implementation of the generated knowledge occurs in order to solve the problem (Akyol, Garrison, & Ozden, 2009; Joo, Lim, & Kim, 2011; Shea & Bidjerano, 2010). Use of reflective questions is highly significant for CP (Bangert, 2008). CP enables a learner to possess the meta-cognitive skills, which makes the learner more active and successful in a cooperative CoI.

When all three components of CoI are present, constitution of individual comprehension in learning and possession of the knowledge in a social process take place (Cleveland-Innes, Garrison, & Kinsel, 2007). In this regard, all three components of the model are interrelated elements that enhance each other (Anderson, Rourke, Garrison, & Archer, 2001; Arbaugh, 2008; Archibald, 2010; Conrad, 2009; Garrison & Cleveland-Innes, 2005; Kozan & Richardson, 2014b; Traver, Volchok, Bidjerano, & Shea, 2014). These studies also support the theoretical framework of the model. In an online learning environment, the occurrence of the three components effects the learning outcomes positively (Akyol, & Garrison, 2008; Horzum, In Press; Ke, 2010; Swan &

Shih, 2005). Consequently, measuring the formation of the components of CoI model in online learning environments is important.

Measurement of CoI Model's Components

In the studies related to the measurement of CoI model's components, the use of different qualitative and quantitative tools as measurement tools is observed. While the studies based on quantitative method include scales, qualitative studies consist of interviews and transcripts. It is possible to state that the studies using scales as measurement tools are in great numbers. In some of the studies using scales, components are measured separately while all components are measured in the other studies. For instance, the related literature includes not only the studies on separate components like SP (Kim, 2011), CP (Shea, & Bidjerano, 2009), and TP (Ice, Curtis, Phillips, & Wells, 2007), but also the studies on all components (Arbaugh, 2008; Arbaugh, Cleveland-Innes, Diaz, Garrison, Ice, Richardson, & Swan, 2008; Bangert, 2009; Burgess, Slate, Rojas-LeBouef & LaPrarie, 2010; Carlon, Bennett-Woods, Berg, Claywell, LeDuc, Marcisz, Mulhall, Noteboom, Snedden, Whalen, & Zenoni, 2012). The studies using scales on one component are regarded as not preferable in terms of the integrated structure and inter-relationships of the components. The measurement tools for all components are preferred to measuring the formation of a community. Among these measurement tools, the most used and cited one is the scale developed by Arbaugh and colleagues (2008). This scale is used in various studies (Akyol, Ice, Garrison, & Mitchell, 2010; Arbaugh, Bangert, & Cleveland-Innes, 2010; Bangert, 2009; Ice, Gibson, Boston, & Becher, 2011; Kovalik & Hosler, 2010 etc.). Besides, there are some studies re-examining the structure of CoI model and construct validity of the scale (Kozan & Richardson, 2014a; Shea, Hayes, Uzuner-Smith, Gozza-Cohen, Vickers, & Bidjerano, 2014; Swan et al., 2008).

Use of exploratory and confirmatory factor analyses based on classical test theory is observed in the scale development studies and re-examinations of validity and reliability of the scale (Arbaugh, Bangert, & Cleveland-Innes, 2010; Bangert, 2009; Diaz, Swan, Ice, & Kupczynski, 2010; Swan, Shea, Richardson, Ice, Garrison, Cleveland-Innes, & Arbaugh, 2008). In another study, the structure is tested by asking the significance of items in the scale and components of the model to the learners (Diaz, Swan, Ice, & Kupczynski, 2010). However, absence of a study based on item response theory examining the scale in terms of substantive results from the group is noticed.

Item Response Theory

Item Response Theory (IRT) was developed to resolve the deficiency of Classical Test Theory (CTT). The basis of IRT, classical measurement models have some limitations. These limitations can be summarized as follows:

- Ability of examinee and characteristics of test cannot be separated in CTT. Ability of the examinees depends on the test items and whether an item is hard or easy depends on the

ability of the examinees. In other words in CTT test items are group dependent and ability of examinees are test-dependent (Hambleton, Swaminathan & Rogers, 1991).

- In CTT reliability is defined as "the correlation between test scores on parallel forms of a test". But, in practice it is nearly impossible to have parallel tests. So that, reliability coefficients that are procured by CTT are lower bound (Hambleton & van der Linden, 1982).

- Standard error of measurement is assumed to be the same for all examinees in CTT. But scores on any test cannot be equal for examinees of different ability. Therefore, the same standard error of measurement for all examinees is implausible (Lord, 1984).

- Classical test theory is test oriented rather than item oriented.

Because of these limitations, some alternative theories and models have been sought. According to Hambleton, Swaminathan and Rogers (1991), this alternative theory would include: (a) item characteristics but not group-dependent, (b) examinee ability scores but not item-dependent, (c) reliability but does not require test to be parallel, (d) a measure of precision for each ability score. It has been shown that all these features appear within IRT (Hambleton & Swaminathan, 1985; Embretson & Reise, 2000; Baker, 2001).

According to IRT, there is a correlation which can be expressed mathematically between unobservable abilities in a certain area or features of individuals and answers of the test items related to these areas. IRT which has superiority over CTT can be used for test development, test equating, identification of item bias, CAT and standard-setting in the studies.

IRT has different models for binary and polytomous data. Numerous measurement instruments especially in attitude assessment include items with multiple ordered response categories. For this kind of data, polytomous item response models are needed to represent trait level. Likert-type scales can be analyzed by Graded Response Model (GRM). GRM is a kind of polytomous IRT model that can be used when item responses are ordered categorically just like in Likert-Scales. In this research, we use GRM to have item and scale parameters.

In online learning, the most widely used scale for measuring the components of CoI was developed by Arbaugh, Cleveland-Innes, Diaz, Garrison, Ice, Richardson, and Swan (2008) (Horzum, In press). The scale was adapted to different languages, mainly Turkish. The problem of this research is to determine whether the form of the Turkish version of this CoI scale is a valid and reliable scale when used IRT.

Aim of the Research

Whether the students perceive as a part of their community, teaching, cognitive, and social presence level could be calculated by using the CoI scale to the students in online learning applications. Making this calculation will reveal areas that need improvement in applications. In

this way, the online learning programs will be able to create a sense of community. Furthermore the data obtained from the scale for which categories of presence are missing will offer tips on both the content design, the components used in the learning management system and precautions that tutorials and administrators should take. With the findings from the scale, designing, planning and implementation to give more effective online learning outputs will be provided. Therefore it is important to have evidence that is valid and reliable in terms of the qualifications that the scale measures.

Exploratory and confirmatory factor analyses are focused in the validity and reliability studies of CoI model whereas another study examines the opinions of learners on the significance of the items. The aim of this study is to examine item and test characteristics of the scale used for components of CoI model by the means of IRT.

Method

Participants

The study group of this research consists of 1,499 learners in online learning programs provided by Sakarya University distance education center. 587 (39.2%) of these learners are female whereas 912 (60.8%) of them are male. Age of the participants ranges from 18 to 57, the mean age (M) is 27.48, and the standard deviation (s) is 6.70.

Process

Item and test features of the scale were examined by the means of Graded Response Theory. In order to use this model of IRT, after testing the assumptions out of the data gathered from 1,499 participants, data model compliance was examined. Following the affirmative results gathered from the examinations, all data was analyzed by using GRM.

Instrument

CoI was used in the study as a measurement tool. CoI scale was developed by Arbaugh, Cleveland-Innes, Diaz, Garrison, Ice, Richardson, & Swan in 2008. In this study, 34-item scale and the structure of 3 sub-factor components of CoI was scrutinized and analyzed by using exploratory factor analysis. Turkish adaptation of the scale was developed by Horzum (in press). For the adaptation study, construct validity of the scale was examined by exploratory and confirmatory factor analyses. In the exploratory factor analysis, it was found out that the scale had three factors structure and total variance explained 67.63%. Subsequently, as a result of confirmatory factor analysis, the fit index of 34-items and three factors structure was found as $\chi^2/df=1.74$, RMSEA=0.071, CFI=0.98, NFI=0.96, and NNFI=0.98. The first factor of CoI scale SP consists of 9 items, the second factor CP contains 12 items, and the last one TP includes 13 items. There are in total 34 items and 3 sub-factors in the scale, implementation of which takes 10 to 30 minutes.

The 5-Likert scale includes questions requiring the answers as in the rank from 'I Completely Disagree' (1) to 'I Strongly Agree' (5). Cronbach's Alpha factors meaning the internal consistency of scale reliability was 0.97 for overall of the scale, 0.90 for SP, 0.94 for CP, and 0.94 for TP (Horzum, In press).

Data Collection and Analysis

With the aim of collecting data, firstly Sakarya University Distance Education Center was consulted for getting necessary permissions related to the implementation of the scale. Following the permission process, CoI scale was changed into an online form, and published on learner management system to which the learners were registered. The scale was filled voluntarily, and no credential information was required.

There are three subtitle in the analysis of this research. First assumptions of model then model-data fit checked, later items were analyzed. In the first step of analysis, unidimensionality and local independence were checked as model assumptions. Factor analysis was applied for the assumption of unidimensionality. To define first factor as dominant, eigenvalues and scree plot was considered (Önder, 2007). To convince unidimensionality assumption a dominant first factor is needed (Hambleton et al., 1991). Confirmatory Factor Analysis (CFA) was conducted in LISREL 8.80 (Jöreskog & Sörbom, 1999) program to see if the scale verifies unidimensional structure. Pursuant to the results of CFA, model-data fit was computed by using the CFA, RMSEA and NNFI values. RMSEA has a value between 0 and 1, and perfect fit occurs when it approaches to 0 (Bollen & Curran, 2006; Tabachnick & Fidell, 2007). Having higher values than 0.90 from CFI and NNFI which are among other fit indexes indicates perfect fit (Tabachnick & Fidell, 2007).

Another assumption of IRT is "local independence". Local independence points to the statistical independence of the responses given by sub-groups of a certain ability level to an item. This assumption is verified if an individual's performance on an item does not affect the performances for other items. Local independence means that ability is not sufficient by itself to explain the relationships between the items (Hambleton & Swaminathan, 1985). Violation of this assumption leads to the violation of the unidimensionality, as well. Therefore, it can be acknowledged that unidimensional 34-item scale verifies local independence.

For the second step of data analysis, model-data fit was examined for GRM which is used for attitude items. At this stage, compliance levels of attitude items were analyzed by the means of the differences between observed and expected frequencies. The differences between observed and expected frequencies are also referred as "residuals". Embretson and Reise (2000) state the residuals approaching zero (<0.1) proves to be a solid criterion for the goodness of model-data fit.

Item calibration was conducted by applying GRM in MULTILOG (Thirsten, 1991) program for the last step of the analysis. Item parameters of the model were identified through Marginal Maximum Likelihood (MML) method. In addition to total information and standard error functions for three sub-scales, information curves of some examined items are given.

Results

Checking Model Assumptions

Polytomous models as well as binary models are required to meet the local independence and unidimensionality assumptions of IRT (Tang, 1996). Principal components factor analysis was conducted on CoI scale to determine if it satisfies the unidimensionality assumption. Findings of eigenvalues and variance proportions related to the conducted factor analysis are displayed in Table 1.

Table 1

Factor Analyses Results

Factor	Eigenvalue	% of Variance	Cumulative %
1	18.301	53.826	53.826
2	1.966	5.783	59.609
3	1.244	3.660	63.269

Table 1 clearly indicates that the 53.826% of variance is explored by the first factor. Besides, the fact that there is a sharp decrease in the proportions between eigenvalues and accounting for the variance after the first factor and that the difference between them becomes nine-fold is remarkable. Based on these results, CoI Scale may be said to be unidimensional or to measure only one structure.

Additionally, CFA was implemented to determine if the scale was unidimensional. The findings of the model were computed as RMSEA=0.054, NFI=0.99, NNFI=0.99, SRMR= 0.032, GFI=0.90, and CFI=0.99, which verifies the unidimensionality of the scale. In order to be able to identify unidimensionality visually, the eigenvalue factor chart is shown in Figure 1.
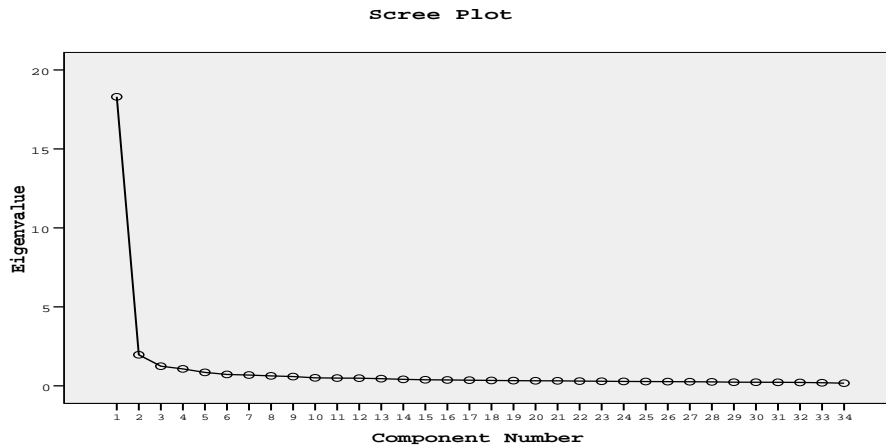
Scree Plot

Figure 1. Scree plot.

According to Figure 1, the slope forms a plateau after the second point, which means that the contributions made by factors after this point are small and almost the same. For this reason, it is thought that the number of factors available is one. In this case, it might be said that the scale is unidimensional.

Another assumption to be checked is local independence. If the assumption of local independence is interrupted then also assumption of unidimensionality is interrupted too. Thus, the assumption of unidimensionality for the 34 items in CoI scale accomplishing could be said as accomplishing the local independence assumption.

Model-Data Fit

Negative log likelihood (-2*LL) value is found as 89129.2 as a result of the calibration of the data gathered from Community of Inquiry Scale Instrument with Graded Response Model. Negative likelihood value in maximum likelihood estimation indicates the degree of data divergence from the model Maximum likelihood estimation (Embretson & Reise, 2000). Marginal reliability coefficient is found as 0.9768. Marginal reliability represents total reliability obtained from the average of the expected conditional standard errors of the students from all competency levels (DCAS 2010–2011, Technical Report).

Item-data fit level can be scrutinized by the means of the differences between observed and expected proportions. The differences between observed and expected frequencies are also referred as "residuals". Embretson and Reise (2000) state the residuals approaching zero (<0.10) proves to be a solid criterion for the goodness of model-data fit. The highest difference between the observed and expected frequencies gathered from the data is 0.0453. When explored the differences between the observed and expected frequencies gathered from each subcategory of 34

items in the scale, it can be observed that all residuals are lower than the value of 0.10. Based on these findings, it was concluded that the Graded Response Model (GRM) was congruous in terms of model-data fit.

Item Analysis

Both CTT and IRT parameter estimates for 34 items and three subscales are shown in Table 2. For each item, the mean, SD, and Corrected Item-Total Correlation (CITC) obtained from CTT, item difficulty and *a* parameter as the item discrimination gathered from IRT are clearly presented. Although there is no certain cutoff criterion for *a* parameter, admissible value can be said as 1 (Zickar, Russel, Smith, Bohle, & Tilley, 2002). But also if there are fewer items in a scale higher discrimination coefficient may be needed (Hafsteinsson, Donovan, & Breland, 2007). In our scale there is between 9 and 13 items per subscales, so we defined *a* values as moderate quality if the values are between 1.0 to 2.0 and as high quality if the values are more than 2.0. Over the 34 item set *a* parameter values are between 1.81 to 3.59 and only two items' *a* value is below 2. While in the teaching presence subscale discriminant value is between 2 and 3 for eight of the items (1, 2, 3, 4, 5, 10, 12, 13), it is over 3 for five of the items (6, 7, 8, 9, 11).

In the nine-item subscale of social presence, two items (15, 16) have *a* value slightly less than 2; the four items (14, 17, 21, 22) are in the range of 2 to 3. Besides, three items (18, 19, 20) in the social presence subscale have discriminating value over 3. While seven items (23, 24, 26, 27, 32, 33, 34) of cognitive subscale presence including 12 items in total have *a* value in the range of 2 to 3, other five items (25, 28, 29, 30, 31) have discriminating over 3.

Table 2

Item Statistics and Item Parameters

Scale	Item	Mean	SD	CITC	α	b_1	b_2	b_3	b_4
Teaching Presence	1	3.7578	1.07250	0.723	2.40	-1.94	-1.22	-0.17	0.84
	2	3.7932	1.08758	0.726	2.45	-1.89	-1.17	-0.24	0.78
	3	3.6785	1.10437	0.746	2.54	-1.76	-1.01	-0.13	0.94
	4	3.7865	1.11180	0.684	2.17	-1.95	-1.15	-0.25	0.79
	5	3.6024	1.12296	0.766	2.89	-1.58	-0.91	-0.05	0.98
	6	3.6117	1.16755	0.776	3.12	-1.50	-0.81	-0.04	0.86
	7	3.5530	1.17680	0.775	3.20	-1.41	-0.78	0.03	0.92
	8	3.5143	1.16251	0.789	**3.33**	-1.39	-0.79	0.07	0.99
	9	3.5697	1.15935	0.792	3.25	-1.45	-0.80	0.01	0.93
	10	3.4950	1.19375	0.742	2.61	-1.51	-0.76	0.08	1.01
	11	3.5384	1.16712	0.778	3.01	-1.48	-0.79	0.04	0.98
	12	3.4330	1.18570	0.733	2.44	-1.48	-0.76	0.12	1.16
	13	3.4683	1.19390	0.668	**2.00**	-1.66	-0.82	0.09	1.17
Social Presence	14	3.5163	1.17618	0.666	2.01	-1.79	-1.00	-0.02	1.02
	15	3.4636	1.13823	0.654	1.90	-1.80	-1.06	0.01	1.26
	16	3.6771	1.16058	0.621	**1.81**	-2.04	-1.18	-0.24	0.86
	17	3.7198	1.07250	0.697	2.56	-1.71	-1.07	-0.29	0.74
	18	3.6778	1.15413	0.740	3.12	-1.62	-0.96	-0.18	0.72
	19	3.7071	1.13048	0.767	**3.56**	-1.63	-1.01	-0.17	0.67
	20	3.6838	1.11433	0.765	3.44	-1.64	-1.02	-0.15	0.75
	21	3.5417	1.06408	0.730	2.67	-1.84	-1.06	0.01	1.08
	22	3.5497	1.11581	0.759	2.82	-1.74	-0.94	-0.04	0.98
Cognitive Presence	23	3.5784	1.15187	0.737	2.52	-1.71	-0.93	-0.02	0.97
	24	3.5223	1.16353	0.753	2.64	-1.61	-0.84	0.04	1.02
	25	3.5844	1.13979	0.794	3.08	-1.62	-0.86	-0.04	0.95
	26	3.6531	1.14608	0.701	**2.34**	-1.76	-0.99	-0.15	0.92
	27	3.5724	1.11277	0.778	2.95	-1.69	-0.89	-0.01	1.00
	28	3.6384	1.13036	0.774	3.00	-1.68	-0.92	-0.09	0.88
	29	3.7372	1.06543	0.807	**3.59**	-1.74	-1.05	-0.19	0.82
	30	3.5844	1.10711	0.788	3.22	-1.65	-0.92	-0.03	0.98
	31	3.6044	1.09508	0.786	3.14	-1.70	-0.94	-0.04	0.99
	32	3.6111	1.06475	0.776	2.99	-1.82	-0.96	-0.04	1.03
	33	3.5604	1.08985	0.764	2.78	-1.75	-0.93	-0.02	1.09
	34	3.7145	1.10762	0.719	2.40	-1.88	-1.11	-0.17	0.87

Note. For each subscale, values fall into the following ranges: Teaching Presence 2.00 to 3.33; Social Presence 1.81 to 3.56; Cognitive Presence 2.34 to 3.59. Of the 34 items, 32 can be characterized as "high quality" since they had *a* values higher than two. Only two items had *a* values as 1.81 and 1.90, which can be referred as "moderate quality".

There are values of IRT parameter and the corrected item total correlation for each item of subscales in Table 2. There is a linear positive relation between CITC and *a*. For example in Table 1 item 16 "Online or web-based communication is an excellent medium for social interaction" has

lowest *a* (1.81) and CITC (0.621) value. Similarly, item 29 "Combining new information helped me answer questions raised in course activities" has the highest values as an *a* of 3.59 and a CITC of 0.807.

Nevertheless, *a* values give much more detailed information than CITC values (Scherbaum et al., 2006). For example, Items 6 and 32 both have the same value of CITC (0.776), but the *a* for item 6 is 3.12 whereas for item 32 it is 2.99.

For each item, the between category threshold parameters (b) are ordered, which must occur in GRM. These parameters determine the location of the operating characteristic curves and where each of the category response curves for the middle response options peaks.
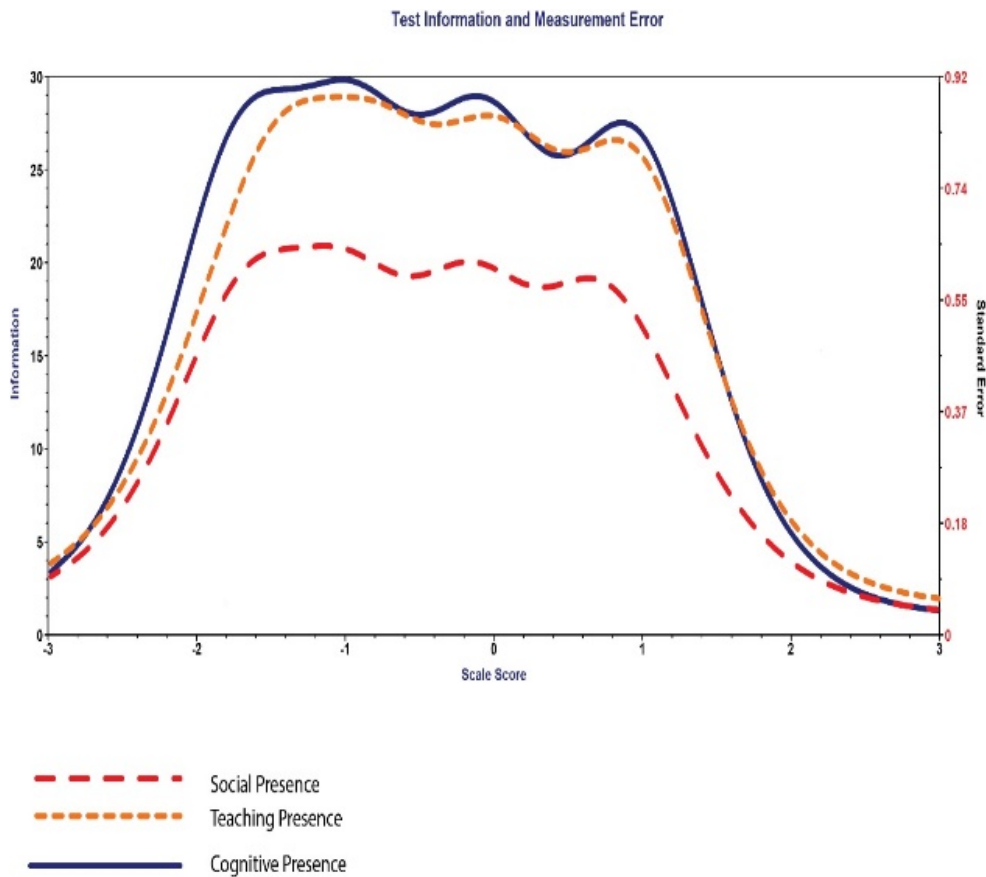

Figure 2. Test Information Function for each of the three subscales.

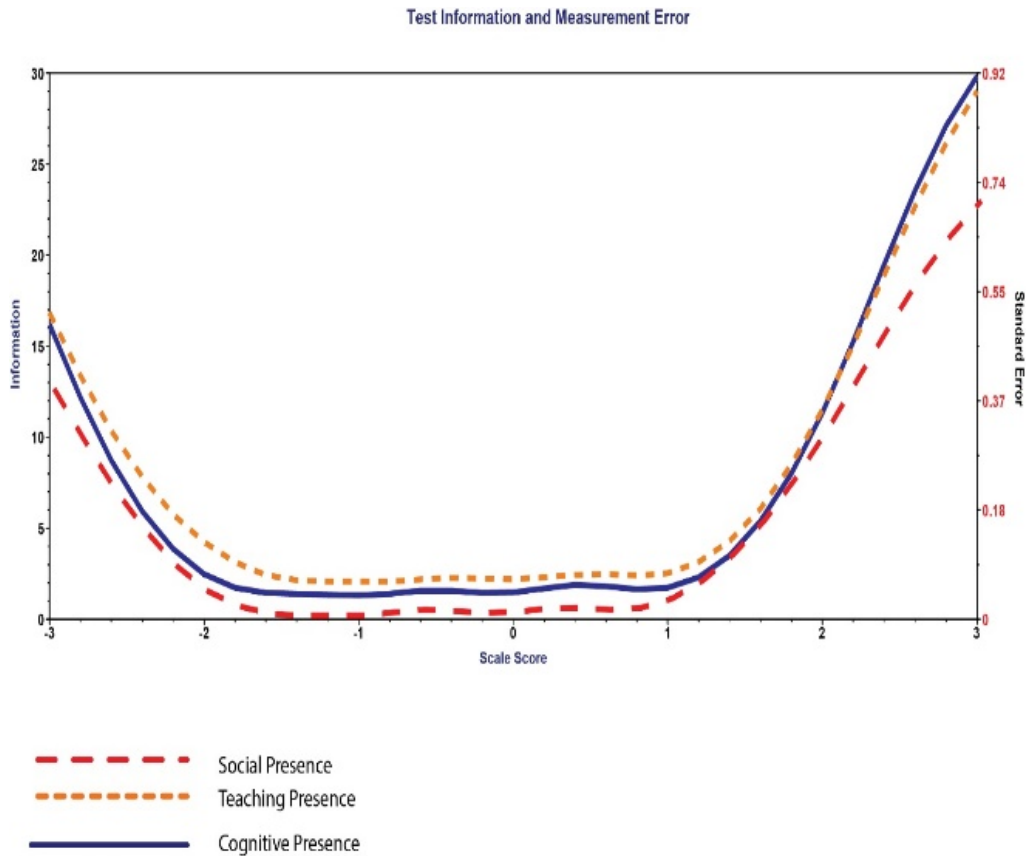**Test Information and Measurement Error**

Figure 3. Standard Error Function for each of the three subscales.

TIFs and structural equation modeling (SEM) curves for each subscale are shown in Figure 2 and 3 respectively. In the figures it is obvious that near the middle and upper ends of the trait distribution each subscale provides rich information.

TP has the highest information at the level of -1.0; SP has the highest information at the level of -1.2; and CP has the highest information at the level of -0.8 θ. Distributions of all subscales are unimodal. In plots of SEM it is obvious that the best measurement of all subscales is at the high end of the distribution. CP has the highest distribution of information. Second scale which gives more information and has less error is TP, and the third one is SP.

Coefficient alpha is the reliability index which is widely used in CTT and it is beneficial to compare information functions and standard errors with it. The alphas for the three factors range from 0.90 to 0.94. The highest alpha, 0.94 is for the CP. For TP, alpha is 0.94 while alpha coefficient is 0.90 for SP. It is obvious that there is a relation between information functions and

alpha values. The highest alpha has the most information function and lowest standard error. Also for CP and TP, alpha values obtained with the IRT analysis reveal the difference between them.

Discussion

The main purpose of this study is to compare the results of the CoI scale which has teaching, social and cognitive presence subscales by the means of CTT and to define the item and test parameters with IRT.

Firstly, *a* is the item discrimination parameter which gives information about item quality. Zickar et al. (2002) suggest that acceptable discriminability for *a* parameter should be higher than 1.0. However, according to Hafsteinsson et al. (2007) *a* values higher than 2.0 provide a higher item quality. In the dataset of this research there are 34 items totally and *a* values of the parameters range from 1.81 to 3.59. There are only two (6%) items which have *a* values below 2.0. All items of cognitive, teaching, and social presence subscales' *a* values are higher than 2.0, except for two items of social presence subscale's item 15 and 16 which are respectively 1.90 and 1.81. Although these items' *a* values are lower than 2.0 they are very close to 2.0.

There is a linear positive relation between CITC and *a*. For example, item 29, "Combining new information helped me answer questions raised in course activities," has an *a* of 3.59 and a CITC of 0.792 which has the highest values for both *a* and CITC. On the other hand, *a* values give more information about item quality than the CITCs (Scherbaum et al., 2006).

The alphas for the three subscales range from 0.90 to 0.94. These values show that scale has high quality. However, instead of the coefficient alpha, the scale information functions and SEMs give much more information. In the figures it is shown that near the middle and upper ends of the trait distribution each subscale gives rich information. -3 to +3 range is for 99% of the population and in this range values of SEM are entirely low for θ. At the low end of the trait continuum the most information and least error are taken by the teaching information function. On the other hand at the high end of the distribution social presence provides the best measurement. This situation is defined with alpha by only .04 (.94 vs. .90), but there is a considerable difference in characteristics of measurement.

Scale adaption studies give information about how the study is done properly in order to identify the better results of items in which the ability ranges. In this study, the parameters of scale and item are obtained and compared with both CTT and IRT. As a result of the comparison, it is seen IRT analysis has given similar results with CTT, but more detailed information has been reached on IRT. In this case, it is concluded that IRT analysis would be more appropriate in this scale study. Throughout examination of TP subscale it was revealed that, four items which are 6th, 7th,

8th, and 9th, have higher values and they form one of TP's indicators which is facilitation. In a similar observation one of CP's indicators integration with items 29, 30, and 31, and one of SP's indicators open communication with items 17, 18, and 19 were found to have higher values.

These findings suggest that facilitation to create TP, integration to create CP, open communication to create SP have great importance. In the light of these findings it is recommended to online learning designers and/or instructors

- to have higher TP, facilitation

- to have higher CP, integration

- to have higher SP, open communication

should have highest importance while designing training tools, and instructional processes.

As a conclusion, in the light of all these findings, both CTT and IRT have verified that the scale adapted to Turkish by Horzum (in press) is valid and reliable. This tool can be used to measure the formation of communities in different online learning applications and the relationship between the sense of community and different variables. Moreover, independent studies from culture and language for whether the scale forms in different cultures and languages are similar with the Turkish form of the scale could be carried out. With obtaining of the forms in different languages, comparisons with different countries about the usage of the scale in the online learning applications in different cultures and countries could be made.

References

Akyol, Z., & Garrison, D. R. (2008). The development of a community of inquiry over time in an online course: Understanding the progression and integration of social, cognitive and teaching presence. *Journal of Asynchronous Learning Networks, 12*(2-3), 3-23.

Akyol, Z., Garrison, D. R., & Ozden, Y. (2009). Online and blended communities of inquiry: Exploring the developmental and perceptional differences. *International Review of Research in Open and Distance Learning, 10*(6), 65-83.

Akyol, Z., Ice, P., Garrison, D. R., & Mitchell, R. (2010). The relationship between course socio-epistemological orientations and student perceptions of community of inquiry. *The Internet and Higher Education, 13*(1-2), 66-68.

Allen, I. E. & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States.* Sloan Consortium. Retrieved August 13, 2014 from http://www.onlinelearningsurvey.com/reports/changingcourse.pdf

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, 5(2). Retrieved August 13, 2014 from http://auspace.athabascau.ca/bitstream/2149/725/1/assessing_teaching_presence.pdf.

Arbaugh, J. B. (2008). Does the community of inquiry framework predict outcomes in online MBA courses? *International Review of Research in Open and Distance Learning, 9*(2), 1-21.

Arbaugh, J. B., & Hwang, A. (2006). Does "teaching presence" exist in online MBA courses? *The Internet and Higher Education, 9*(1), 9−21.

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K.P. (2008). Developing a community of inquiry instrument: Testing a measure of the Community of Inquiry Framework using a multi-institutional sample. *The Internet and Higher Education, 11*(3-4), 133-136.

Arbaugh, J., Bangert, A., & Cleveland-Innes, M. (2010). Subject matter effects and the community of inquiry (CoI) framework: An exploratory study. *The Internet and Higher Education, 13*(1−2), 37−44.

Archibald, D. (2010). Fostering the development of cognitive presence: Initial findings using the community of inquiry survey instrument. *The Internet and Higher Education, 13*(1-2), 73−74.

Baker, F. B. (2001). *The basics of item response theory* (2nd Edition). USA, Heinemann.

Bangert, A. W. (2008). The influence of social presence and teaching presence on the quality of online critical inquiry*. Journal of Computing in Higher Education, 20*(1), 34–61.

Bangert, A. W. (2009). Building a validity argument for the community of inquiry survey instrument. *The Internet and Higher Education, 12*(2), 104-111.

Bollen, K. A. & Curran, P. J. (2006). *Latent curve models: A structural equation perspective.* New Jersey: John Wiley & Sons, Inc.

Burgess, M. L., Slate, J. R., Rojas-LeBouef, A., & LaPrairie, K. (2010). Teaching and learning in Second Life: Using the community of inquiry (CoI) model to support online instruction with graduate students in instructional technology. *The Internet and Higher Education, 13*(1-2), 84–88.

Carlon, S., Bennett-Woods, D., Berg, B., Claywell, L., LeDuc, K., Marcisz, N., ... Zenoni, L. (2012). The community of inquiry instrument: Validation and results in online health care disciplines. *Computers & Education, 59*(2), 215–221.

Cleveland-Innes, M., Garrison, D. R., & Kinsel, E. (2007). Role adjustment for learners in an online community of inquiry: Identifying the needs of novice online learners. *International Journal of Web-based Learning and Teaching Technologies, 2*(1), 1-16.

Conrad, D. (2009). Cognitive, teaching, and social presence as factors in learners' negotiation of planned absences from online study. *International Review of Research in Open and Distance Learning, 10*(3), 1-18.

DCAS (2010–2011). *Delaware Comprehensive Assessment System. Evidence of Reliability and Validity.* Volume 4. American Institutes for Research.

Diaz, S. A., Swan, K., Ice, P., & Kupczynski, L. (2010). Student ratings of the importance of survey items, multiplicative factor analysis, and validity of the community of inquiry survey. *The Internet and Higher Education, 13*(1-2), 22-30.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Lawrence Erlbaum Associates Publishers, Mahwah, New Jersey.

Garrison, D. R. (2007). Online community of inquiry review: Social, cognitive, and teaching presence issues. *Journal of Asynchronous Learning Networks, 11*(1), 61-72.

Garrison, D. R., & Anderson, T. (2003) *E-learning in the 21st century: A framework for research and practice.* London: Routledge/Falmer.

Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education, 10*(3), 157-172.

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education, 19*(3), 133–148.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2-3), 87-105.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of distance education, 15*(1), 7-23.

Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships between teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education, 13*(1-2), 31-36.

Hafsteinsson, L. G., Donovan, J. J., & Breland, B. T. (2007). An item response theory examination of two popular goal orientation measures. *Educational and Psychological Measurement, 67*, 719–739.

Hambelton, R.K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of item response theory.* Newbury Park, CA: Sage Publications.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications.* Kluwer Nijhoff Publishing, Baston.

Hambleton, R. K., & van der Linden, W. J. (1982) Advances in item response theory and applications: An introduction. *Applied Psychological Measurement, 6, 373-378.*

Horzum, M. B. (In press). Online learning students` perceptions community of inquiry based on learning outcomes and demographic variables. *Croatian Journal of Education.*

Ice, P., Curtis, R., Phillips, P., & Wells, J. (2007). Using asynchronous audio feedback to enhance teaching presence and students' sense of community. *Journal of Asynchronous Learning Networks, 11*(2), 3-25.

Ice, P., Gibson, A. M., Boston, W., & Becher, D. (2011). An exploration of differences between community of inquiry indicators in low and high disenrollment online courses. *Journal of Asynchronous Learning Networks, 15*(2), 44-69.

Joo, Y. J., Lim, K. Y., & Kim, E. K. (2011). Online university students' satisfaction and persistence: Examining perceived level of presence, usefulness and ease of use as predictors in a structural model. *Computers & Education, 57*(2), 1654-1664.

Jöreskog, K. G. & Sörbom, D. (1999). *LISREL 8: Structural equation modeling with the Simplis Command Language.* USA: Scientific Software ınternational, Inc.

Ke, F. (2010). Examining online teaching, cognitive, and social presence for adult students. *Computers & Education, 55*(2), 808-820.

Kim, J. (2011). Developing an instrument to measure social presence in distance higher education. *British Journal of Educational Technology, 42*(5), 763-777.

Kovalik, C. L., & Hosler, K. A. (2010). Text messaging and the community of inquiry in online courses. *MERLOT Journal of Online Learning and Teaching, 6*(2). Retrieved August 13 2014 from http://jolt.merlot.org/vol6no2/kovalik_0610.htm

Kozan, K., & Richardson, J. C. (2014a). Interrelationships between and among social, teaching, and cognitive presence. *The Internet and Higher Education, 21*, 68-73.

Kozan, K., & Richardson, J. C. (2014b). New exploratory and confirmatory factor analysis insights into the community of inquiry survey. *The Internet and Higher Education, 23*, 39-47.

Lord, F. M. (1984). Standard errors of measurement at different ability levels. *Journal of Educational Measurement*, 21, 239-243.

Moore, M. G., & Kearsley, I. G. (2012). *Distance education: A systems view of online learning* (3rd Ed.). New York: Wadsworth Publishing Company.

Onder, I. (2007). An investigation of goodness of model data fit. *Hacettepe University Journal of Education, 32*, 210-220.

Picciano, A. G. (2002). Beyond student perceptions: Issues of interaction, presence, and performance in an online course. *Journal of Asynchronous Learning Networks, 6*(1), 21-40.

Shea, P. & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster "epistemic engagement" and "cognitive presence" in online education. *Computers & Education, 52*(3), 543-553.

Shea, P. (2006). A study of students' sense of learning community in online environments. *Journal of Asynchronous Learning Networks, 10*(1), 35-44.

Shea, P. J., Fredericksen, E. E., Pickett, A. M., & Pelz, W. E. (2003). A preliminary investigation of "teaching presence" in the SUNY learning network. *Elements of quality online education: Practice and direction, 4*, 279-312.

Shea, P., & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers & Education, 55*(4), 1721-1731.

Shea, P., Hayes, S., Uzuner-Smith, S., Gozza-Cohen, M., Vickers, J., & Bidjerano, T. (2014). Reconceptualizing the Community of Inquiry framework: Exploratory and confirmatory analysis. *The Internet and Higher Education, 23*, 9-17.

Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., ... Rangan, P. (2010). A re-examination of the community of inquiry framework: Social network and content analysis. *The Internet and Higher Education, 13*(1-2), 10-21.

Shea, P., Li, C. S., & Pickett, A. (2006). A study of teaching presence and student sense of learning community in fully online and web-enhanced college courses. *The Internet and Higher Education, 9*(3), 175-190.

Swan, K., & Shih, L. F. (2005). On the nature and development of social presence in online course discussions. *Journal of Asynchronous Learning Networks, 9*(3), 115-136.

Swan, K., Shea, P., Richardson, J., Ice, P., Garrison, D. R., Cleveland-Innes, M., & Arbaugh, J. B. (2008). Validating a measurement tool of presence in online communities of inquiry. *E-mentor, 2*(24), 1-12.

Tabachnick, B. G. & Fidell, L.S. (2007). *Using multivariate statistics* (5th Ed). USA: Pearson.

Tang, K. L. (1996). *Polytomous item response theory models and their apllications in large-scale testing programs: Review of literature.* Educational Testing Service, Princeton, New Jersey.

Thissen, D. (1991). *MULTILOG user's guide* (Version 6.0). Mooresville, IN: Scientific Software.

Traver, A. E., Volchok, E., Bidjerano, T., & Shea, P. (2014). Correlating community college students' perceptions of community of inquiry presences with their completion of blended courses. *The Internet and Higher Education, 20*, 1-9.

Zickar, M. J., Russel, S. S., Smith, C. S., Bohle, P., & Tilley, A. J. (2002). Evaluating two morningness scales with item response theory. *Personality and Individual Differences, 33*, 11–24.

Athabasca University