

## A REPLICATION STUDY ON THE MULTI-DIMENSIONALITY OF ONLINE SOCIAL PRESENCE

**David B. Mykota**

*College of Education, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 0X1 Canada*

*Tel.: 1-306-966-5258; fax: 1-306-966-7719.*

*david.mykota@usask.ca*

### ABSTRACT

The purpose of the present study is to conduct an external replication into the multi-dimensionality of social presence as measured by the Computer-Mediated Communication Questionnaire (Tu, 2005). Online social presence is one of the more important constructs for determining the level of interaction and effectiveness of learning in an online environment. This study is unique as it provides the opportunity to replicate previous research on the dimensionality of the instrument and offers insight as to the nature of social presence among students enrolled in higher education postgraduate online courses.

**Keywords:** replication analysis, social presence, online learning

### INTRODUCTION

To frame good pedagogical practices for online learning, Garrison, Anderson, and Archer (2000) developed the community of inquiry model recognizing the transactional relationship between instructors and learners through the interaction of cognitive presence (of the learner), teaching presence (i.e. the structure and process), and social presence (i.e. affective interpersonal communication). According to Garrison et al. (2000), these elements, which define the community of inquiry model, are fundamental to a successful higher education learning experience. A key component in the model is the concept of social presence that refers to the affective domain as it relates to interpersonal communications. If social presence is absent, learner frustration develops because of the poor quality of interpersonal interactions (Rifkind, 1992). Conversely, a high degree of social presence in online learning is viewed as synonymous with an interactive environment (Lobry de Bruyn, 2004; McIssac & Gunawardena, 1996).

In assessing social presence within a text-based, asynchronous environment, three categories of communicative responses have been identified (Rourke, Anderson, Garrison, & Archer, 2001): affective indicators (i.e., values, beliefs, feelings, and emotions); cohesive indicators (i.e. group presence and commitment); and interactive indicators (i.e., attending in a socially meaningful way). Although Rourke et al. (2001) recognize that the coding and analyzing of CMC text-based transcripts using the aforementioned indicators provides a measure of the density of social presence, they also contend future exploratory studies including factor analysis would aid in further defining the construct.

The construct of social presence is the critical affective ingredient for online learning. Although earlier research on social presence has found it to be related to learner satisfaction (Gunawardena & Zittle, 1997) others such as Wise, Chang, Duffy and del Valle (2004) argue that it is more of a correlational than causal variable. Nevertheless, several researchers have demonstrated that social presence is one of the more important constructs to determine the level of interaction and effectiveness of learning in an online environment (Garrison et al., 2000; Gunawardena & Zittle, 1997; McIssac & Gunawardena, 1996; Lobry de Bruyn, 2004; Richardson & Swan, 2003; Rourke et al., 2001; Tu & McIssac, 2002). For example, Kim, Kwon, and Cho (2011) found that students who valued media integration, quality instruction and interactivity had increased social presence perception and heightened learning satisfaction While So and Brush (2008) demonstrated that in a distance learning environment, student perception and overall satisfaction of collaborative learning is associated with social presence (So, 2008).

Part of the difficulty in aggregating findings is the varying way that researchers have measured and reported social presence. For example, earlier efforts were found lacking because they were created for the face-to-face environment and then adapted to online learning without proper validation procedures. Recently, Sung and Mayer (2012) found a five-factor solution for the Online Social Presence Questionnaire that included social respect, social sharing, open mind, social identity, and intimacy. However, according to Sung and Mayer (2012) a limitation of their study is that it focused only on university students in Korea and consequently future research needs to examine if similar findings would be replicated in samples with different ages, genders, grades, intellectual levels, and cultural backgrounds.

Another instrument that has been developed to measure the construct, social presence, in a computer-mediated communication (CMC) environment is the Computer Mediated Communication Questionnaire (CMCQ; Tu 2005). The exploratory factor analysis of the CMCQ resulted in a four factor multi-dimensional model of online social presence (Tu & Yen, 2006). With the increased frequency of research reports using the CMCQ (Mykota & Duncan, 2007; So, 2008; So & Brush, 2008; Stein & Wanstreet; 2003; Tu & McIssac; 2002; Tu & Yen, 2007) the need to explore the multidimensional characteristics of the instrument with different samples in varying contexts (i.e. replication analysis) is warranted so as to advance understanding of the construct’s operationalization and interpretation.

Problematic to replication analysis are the lack of standardized metrics available to make informed comparisons of exploratory factor analysis (EFA) procedures conducted with separate data sets. Good EFA studies detail what might be expected for replication results. This can include the number of factors present, whether the factors are correlated or not, the factor loadings for the communalities that comprise the factor and the names of the factors as informed by substantive interpretation which includes the theoretical underpinnings for the factors identified (Osborne & Fitzpatrick, 2012). With the lack of standardized metrics, a similar procedure is advocated for EFA replication purposes whereby items are examined to see if they load to the same factors and if the individual item factor loadings are equivalent (Osborne & Fitzpatrick, 2012).

**PURPOSE**

With procedures for analysis having been reported in the literature and a basis for the interpretation of EFA replication data established it was decided to undertake a replication analysis. The main objective of the study then is to conduct an external replication (i.e. an independent external sample) of the CMCQ (Tu, 2005) using the same EFA procedures as previously reported by Tu and Yen (2006). By doing so a better understanding of the instrument’s (CMCQ; Tu, 2005) generalizability will occur while adding to the extant research on the measurement of the construct social presence in an online learning environment.

**METHOD**

The participant sample is derived from students enrolled in the postgraduate special education program offered at the University of Saskatchewan. The postgraduate program is comprised of 9 courses that are delivered over two years. Using convenience sampling, 275 students (90% response rate) enrolled in the postgraduate program participated in the study. Participants in the study voluntarily completed the computer-mediated communication questionnaire (CMCQ; Tu, 2005). The sentence stems on the CMCQ were used to identify social presence in a text-based system with the CMC tools (email, discussion, and chat) The respondent was asked to complete each of the instrument’s 24 items on the basis of a five-point Likert scale converted to a numerical weighting ranging in options from 0 (uncertain); 1 (strongly disagree); 2 (disagree); 3 (agree); and 4 (strongly agree). The frequency counts for the demographic variables age, sex, and numbers of online classes taken are reported in Table 1.

Table 1. Frequency Demographics, N=275

	Frequency	Percent
Sex		
Male	20	8.3
Female	254	91.7
Age		
18-25	39	14.1
26-33	95	34.3
34-41	84	30.3
42 or older	59	21.3
Number of Online Courses		
1 course	113	40.8
2-3 courses	78	28.2
4 or more courses	86	31.0

To determine the adequacy of the sample for factorial analysis the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (i.e. KMO) and Bartlett’s Test of Sphericity were calculated. The replication analysis of the CMCQ’s multidimensionality also included a reliability analysis of the instrument’s items. A visual inspection of the corrected item total correlations for the CMCQ was conducted with items below .10 deemed as outlier test items with a Cronbach alpha reliability coefficient value equal to or above .70 considered to reflect a high internal consistency. Following the procedures previously reported by Tu and Yen (2006) a principal components analysis with varimax orthogonal rotation was conducted on those items deemed to reflect high internal

consistency. Subsequent to the orthogonal rotation a principal axis factor analysis using a promax oblique rotation was undertaken. In determining what factors were to be retained a two-line scree test of those the factors with eigen values >1 was applied. The methods described for the exploratory factor analysis replicates the methods of the initial validation study conducted (Tu & Yen, 2006). By following the same procedure undertaken by Tu and Yen (2006) a comparative analysis of whether the same factor structure with equivalent factor loadings and item communalities of the CMCQ can occur. The methods outlined are consistent with procedural aspects of replication analysis as reported by Osborne and Fitzpatrick (2012). As an external replication of the previously reported EFA undertaken (Tu & Yen, 2006) the present study adds value as it aids in determining the degree to which the CMCQ factor structure generalizes to a new data set and if there are any items considered challenging or problematic.

## RESULTS

A reliability analysis for the instrument (i.e. 24 items) was conducted with two items deemed as outliers and removed (i.e. items 4 & 19), see Table 2. The 22 remaining items reflected relatively high internal consistency with a Cronbach alpha reliability coefficient calculated at .732. In determining the appropriateness of conducting a factor analysis the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett’s Test of Sphericity were calculated (Tabachnick & Fidell, 2007). For the data the KMO statistic is between .7 and .8 (i.e. .734) which is deemed as being good, this indicates the sample should be adequate for the conducting of a factor analysis (Hutcheson & Sofroniou, 1999). Bartlett’s Test of Sphericity tests the null hypothesis that the correlation matrix is an identity matrix, which implies the variables are uncorrelated. In this case, Bartlett’s Test of Sphericity was significant ( $\chi^2(754) = 231 p < .05$ ) which leads to the rejection of the null hypothesis and to the conclusion that there are correlations in the data set and that the data is appropriate for the conducting of a factor analysis.

Table 2. Item-Total Statistics

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
q1	46.6606	87.463	.345	.719
q2	46.8358	89.588	.201	.728
q3	47.3321	86.977	.279	.723
q5	47.2226	87.221	.233	.727
q6	47.9343	86.897	.266	.724
q7	46.9088	88.464	.254	.724
q8	46.7482	89.237	.251	.725
q9	47.4197	85.285	.318	.719
q10	47.2080	85.374	.401	.714
q11	47.2810	82.086	.466	.706
q12	47.0620	85.912	.356	.717
q13	47.2847	88.842	.226	.726
q14	47.5328	85.700	.288	.722
q15	46.9197	85.788	.357	.717
q16	47.5328	88.067	.252	.725
q17	47.3905	87.748	.295	.722
q18	47.8978	88.671	.176	.731
q20	47.7263	85.328	.333	.718
q21	46.8686	89.052	.267	.724
q22	46.3102	91.021	.172	.729
q23	46.8029	85.023	.389	.714
q24	48.0073	88.601	.145	.736

Like the initial EFA validation study conducted by Tu and Yen (2006), a principal components analysis with varimax orthogonal rotation was used to explore the underlying factor structure. Using the Kaiser criterion the first stage of the analysis identified seven components with eigen values greater than 1. In total these components accounted for 51.4% of the variance in the test items. The five factor pattern that best met the criteria of simple structure (Thurstone, 1947) and was clearly interpretable, and in this instance replicable as reported in the previous EFA (Tu & Yen, 2006), was retained. In the rotated component matrix the five components were well defined by 3-4 test items that had high loadings ( $\geq .320$ ) on only one component with those items that did not have a high loading on any one component or loaded across components excluded, see Table 3.

In the previous EFA with the CMCQ (Tu & Yen, 2006) a promax oblique rotation was also conducted because aspects of the multidimensional construct social presence as measured by the CMCQ were correlated. For the second phase of the EFA replication analysis, a similar procedure was adopted to determine the degree of correlation between the factors. In the principal axis factoring analysis using a promax oblique rotation, seven factors with eigen values greater than 1 were extracted accounting for 51.4% of the variance. A process similar to that used for the identification of a simple structure that is clearly interpretable and replicable resulted in a five-component solution being retained. In the original EFA conducted by Tu and Yen (2006) the structure matrix was used for interpreting the oblique rotation. The structure matrix is a simple correlation of the variables with the factors. Following the procedures conducted in the original EFA, items with correlations  $\geq .320$  for a factor and not correlated with another factor were retained see Table 4.

Table 3. Principal Component Analysis Rotated Component Matrix

Item	Component				
	1	2	3	4	5
q7	.745			-.121	
q10	.654		.215		
q1	.559			.157	.101
q15	.539	-.144		.276	.160
q18		.770	-.106		.128
q14		.648	.211		.116
q24		.606		-.106	
q9	.293	.470	.131	.333	-.144
q3			.655		
q11	.266	.143	.636	.138	
q12	.236		.492		.381
q8	.152	.124	.489		
q16	.161	-.157		.754	
q20			.176	.609	.150
q6		.129		.513	
q22					.811
q23	.205	.253	.167		.636
q13			.189		.598
q2	.344		-.264		.229
q17	.134	.145	.217	.240	-.171
q5		.110		.207	
q21	.158			.247	.214

Table 4. Principal Axis Factor Analysis Structure Matrix

Item	Factor				
	1	2	3	4	5
q7	.698		.244	.109	.177
q1	.474		.322	.279	.235
q15	.412		.278	.270	.222
q18		.684			.194
q14	.119	.551	.311	.113	.236
q24		.396			
q11	.342	.167	.600	.217	.260
q3	.185		.464	.179	.150
q12	.297		.382	.157	.358
q17	.210	.139	.343	.288	
q8	.225	.147	.327	.164	.113
q21	.262	.131	.309	.268	.238
q10	.155		.292	.206	
q16	.216		.196	.624	.138
q20	.139	.116	.286	.453	.171
q6		.140	.287	.355	
q23	.375	.325	.390	.224	.658
q22		.121		.101	.571
q13	.502		.286	.230	.334
q9	.219	.349	.161	.252	

q5			.174	.202	
q2	.204		.138	.141	.135

Following the EFA procedures conducted by Tu and Yen (2006), both the item and factor loadings for the principal components analysis with varimax orthogonal rotation and the structure matrix for the principal axis factoring analysis using a promax oblique rotation were reported. In keeping with EFA replication practices (Osborne & Fitzpatrick, 2012) the factor structure and squared difference for item loadings on the factors for the orthogonal rotation and oblique rotation used in the original (Tu & Yen, 2006) and replication studies were compared. The factor structure and item loadings for the principal components analysis with varimax rotation are reported with the squared difference for item loadings determined, see Table 5. A similar procedure was used for the oblique rotation and is presented in Table 6.

Table 5. Principal Component Matrixes for CMCQ Factor Replication

Item	External Sample Factor Loadings					CMCQ Factor Loadings				Squared Difference
	F1	F2	F3	F4	F5	F1	F2	F3	F4	
q1	.559					.644				.007
q2										
q3			.655			.703				.002
q4								.516		
q5										
q6				.513						
q7	.745									
q8			.489				.369			.014
q9		.470								
q10	.654								.727	.005
q11			.636							
q12			.492							
q13					.598		.699			.010
q14		.648								
q15	.539									
q16				.754		.697				.003
q17										
q18		.770						.322		.201
q19										
q20				.609		.757				.022
q21										
q22					.811				.687	.015
q23					.636		.706			.005
q24		.606						.699		.009

Table 6. Principal Axis Factor Structure Matrixes for CMCQ Factor Replication

Item	External Sample Factor Loadings					CMCQ Factor Loadings				Squared Difference
	F1	F2	F3	F4	F5	F1	F2	F3	F4	
q1	.611					.523				.007
q2										
q3			.464			.604				.020
q4							.343			
q5										
q6				.355						
q7	.698									
q8			.327					.338		.000
q9										
q10									.612	
q11			.600							
q12			.382							
q13					.334			.616		.080
q14		.551								

q15	.412				
q16		.624	.740		.013
q17		.343			
q18	.684			.554	.017
q19					
q20		.453	.818		.133
q21					
q22			.571	.621	.003
q23			.658	.546	.013
q24	.396			.392	.000

In the replication analysis the five factor solution was comprised of the Social Form of Communication, Privacy, Intimacy, Social Context, and Interactivity factors, see Table 7. The Social Form of Communication factor was related to how CMC is viewed as a pleasant, social form of communication. The Privacy factor was related to the confidentiality of the CMC medium. The Intimacy factor describes the degree to which students express their personal stories and feelings. The Social Context factor related to the ability of CMC to build trusting, caring, social relationships. Finally, the Interactivity factor related to one’s CMC skill set and communication style.

Table 7. External Replication: Factors, Items, and Item Stems

Factor	Item	Item Stem
Social form of Communication	7	CMC is a pleasant means of communicating with others
	1	CMC are social forms of communication
	15	CMC allows realtionships to be based upon sharing and exchanging information
Privavcy	18	It is unlikely that someone might obtain personal information about you in CMC
	14	CMC is technically reliable (e.g. free from reliability errors).
	24	It is unlikely that someone might redirect your messages.
Intimacy	11	The language used to express oneself in CMC is meaningful
	3	CMC messages convey feeling and emotion
	12	The language used to express oneself in CMC is easily understood.
	17	The aggressive over-participation of others in CMC may cause to participate less.
Social Context	16	CMC allows me to build more caring and social relationships.
	20	CMC permits the building of trust relationships
	6	CMC is a sensitive means of communicating with others.
Interactivity	22	My keyboarding skills allow me to be comfortable participating in CMC.
	23	I am comfortable with the communication styles employed by CMC users.
	13	I am comfortable participating, even if not familiar with the topics.

**DISCUSSION**

The external replication failed the most basic tenant of internal consistency reliability analysis with the determination of outlier test items (i.e. items 4 & 19). Nevertheless the reaming 22 items did meet the threshold for retention with the Cronbach alpha reliability coefficient deemed as moderately good and the sample determined to be appropriate for factor analysis.

In assessing the congruence of the factor matrixes presented it is clear that the communalities and item loadings for the factors do not meet structural equivalence in the strictest sense. Although the factor structure and pattern for item loadings across factors differs from the original EFA validation study conducted by Tu and Yen (2006) there are some similarities that are worthy noting. As such, there is some evidence that the Privacy, Social Context and Interactivity factors and item loadings have moderate item equivalency.

In both studies the Privacy factor was the second factor retained in the factor matrix. In the original EFA oblique rotation it accounted for 7.241 percent of the variance (Tu, 2005). In the replication analysis the amount of variance accounted for was slightly higher at 8.106 percent. Items 24 and 18 in both studies loaded above the .320 correlation with the squared difference for the item loadings being low. The Social Context factor (i.e. factor 1), in the original EFA, accounted for the most variance at 24.042 percent (Tu, 2005) whereas in the replication analysis the amount of variance explained was 5.625 percent. For comparative purposes in the original EFA items 1, 3, 16, and 20 had item loadings that warranted retention. In the replication analysis the item loadings were mixed with 16 and 20 loading on the Social Context factor (i.e. factor 4), item 1 loading on the Social Communication factor (i.e. 1) and item 3 loading on the Intimacy factor (i.e. 3). In all instances the squared differences for the four items is low, however, only items 16 and 20 loaded on a single factor. The last remaining factor that exhibited moderate equivalency was the Interactivity factor. In the original EFA the Interactivity factor (i.e. factor 3) accounted for 6.839 percent of the variance (Tu, 2005) whereas in the replication analysis it accounted for 5.469 percent of the variance (i.e. factor 5). For comparative purposes two out of the three items from the original EFA loaded in the replication analysis. In this respect, items 13 and 23 had low squared differences whereas item 8 which loaded on the Interactivity factor in the original EFA validation did not in the replication analysis and instead had a relatively low item loading on Factor 3.

Further similarities between the original and replication EFA studies can be found among those items that did have component loadings above the absolute value. (i.e. .320). In this instance items 2, 5, 19 and 21 did not load on any communality in either study. Based then on the results of the replication analysis there is some evidence of moderate to low equivalency with the original EFA, however it would be erroneous to assume the strong structural equivalence exists. As to why this might have occurred it is important to understand that an EFA external replication analysis is specific to the context and sample from which the study originated. In the present replication analysis, limitations exist in the sample as it was drawn from those individuals who had little or no previous experience with online learning.

As well, the sample is from a different cultural context. In this respect, cultural mindset and native language communication patterns have been shown to shape computer mediated communication patterns (Yildiz, 2009). The degree of media integration in online learning has also been found to be a significant predictor of social presence (Kim, Kwon & Cho, 2011) and could be another possible explanation as to potential differences within the online programs from which the sample was drawn. Although the sample for the replication analysis is not heterogeneous in gender composition, research on the construct social presence has found that gender differences do not exist (Kim, Kwon & Cho, 2011; Tu, Yen & Blocher, 2011).

## CONCLUSION

In the external replication analysis convenience sampling was used, however, this is viewed as a limitation as it affects the generalizability of the results. Future research should continue external replication of social presence instrumentation in varying contexts and with more systematic sampling procedures. As well, refinement of the item pool and item construction of the CMCQ might alleviate some of the issues surrounding structural equivalence and the generalization of findings.

## REFERENCES

- 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.
- Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American Journal of Distance Education*, 11(3), 8-26.
- Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks, CA: Sage Publications.
- Kim, J., Kwon, Y., & Cho, D. (2011). Investigating factors that influence social presence and learning outcomes in distance higher education. *Computers and Education*, 57(2), 1512-1520.
- Lobry de Bruyn, L. (2004). Monitoring online communication: Can the development of convergence and social presence indicate an interactive learning environment? *Distance Education*, 25(1), 67-81.
- McIsaac, M. S., & Gunawardena, C. N. (1996). Distant education. In D. Jonassen (Ed.), *Handbook for research on educational communications and technology* (pp. 403- 437). New York: Scholastic Press.
- Mykota, D., & Duncan, R. (2007). Learner Characteristics as Predictors of Online Social Presence. *Canadian Journal of Education*, 30(1), 157-170.
- Osborne, J. W., & Fitzpatrick, D. C. (2012). Replication Analysis in Exploratory Factor Analysis: What it is and why it makes your analysis better. *Practical Assessment, Research & Evaluation*, 17(15), 1-8.
- Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks*, 7(1), 68-88.

- Rifkind, L. J. (1992). Immediacy as a predictor of teacher effectiveness in the instructional television. *Journal of Interactive Television, 1*(1), 31-38.
- Rourke, L., Anderson, T., Garrison, R. D., & Archer, W. (2001). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 4*(21), 50-71.
- So, H-J. (2008). When groups decide to use asynchronous online discussions: collaborative learning and social presence under a voluntary participation structure. *Journal of Computer Assisted Learning, 25*, 143–160.
- So, H-J., & Brush, T.A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education, 51*, 318-336.
- Stein, D.S., & Wanstreet, C.E. (2003). *Role of social presence, choice of online or face-to-face group format, and satisfaction with perceived knowledge gained in a distance learning environment*. Paper presented at 2003 Midwest Research to Practice Conference in Adult, Continuing, and Community Education, Columbus, OH.
- Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior, 28*, 1738–1747.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Upper Saddle River, NJ: Pearson Allyn & Bacon.
- Thurston, L. L. (1947). *Multiple-factor Analysis*. Chicago, IL.: University of Chicago Press.
- Tu, C. H. (2005). *Computer-mediated communication (CMC) questionnaire*. Flagstaff, AZ: Author.
- Tu, C. H., & McIssac, M. S. (2002). An examination of social presence to increase interaction in online classes. *American Journal of Distance Education, 16*(2), 131-150.
- Tu, C. H., & Yen, C. J. (2006). A study of multidimensional online social presence. In L. W.Cooke (Ed.), *Frontiers in higher education* (pp.77-104). Hauppauge, NY: Nova Science.
- Tu, C., & Yen, C. (2007). *A study of the relationship between computer-mediated communication proficiency and online social presence*. Presented at the annual conference of AERA, Chicago, IL.
- Tu, C-H., Yen, C-J., & Blocher, M. (2011). A study of the relationship between gender and online social presence. *International Journal of Online Pedagogy and Course Design, 1*(3), 33-49.
- Yildiz, S. (2009). Social presence in the web-based classroom: implications for intercultural education. *Journal of Studies in International Education, 13*(46), 46-65.
- Wise, A., Chang, J., Duffy, T., & del Valle, R. (2004). The effects of teacher social presence on student satisfaction, engagement, and learning. *Journal of Educational Computing Research, 31*(3), 247-271.