Evaluating Online Social Presence: An Overview of Social Presence Assessment

Guoqiang Cui
Yantai University

Abstract: As an important variable in online learning environment, the construct of social presence has been widely studied by researchers in order to investigate students’ online communication behavior and their related performance. This study will provide an overview of the assessment of social presence throughout its historical development. In this review, both primary subjective and objectives measures of social presence will be introduced, followed by criticisms towards current social presence measures, and offer recommendations for future development of social presence measurement tools.

Keywords: social presence, measurement, subjective, objective

1. Introduction

Social presence is an important construct in investigating students’ online learning experience. As Short, William, and Christie (1976) claimed that social presence is “the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships” (p. 65), the quality of social presence is greatly related to communication media’s attributes and affects students’ communication behaviors. In the meantime, social presence has been widely studied especially in the online learning settings and it has been found to have a great impact on online learning students’ satisfaction (Richardson & Swan, 2003), development of online community (Garrison, Anderson, & Archer, 2000), online interaction behaviors (Tu & McIsaac, 2002), building of virtual world (Mennecke, Triplett, Hassall, Conde, & Heer, 2011), among others. In spite of the significant importance of the social presence construct, researchers (Blocher, Amato, & Storslee, 1996; Gunawardena, 1995; Lin, 2004) have been arguing about certain flaws in current assessment methods. Therefore, this study will provide an overall summary of the research and studies about the assessment of social presence, in order to provide a thorough implication for the design and development of social presence measurement in the future studies.

The development of social presence measurement is practically based upon the ever-evolving definitions of social presence (Cui, Lockee, & Meng, 2012). Systematic research into the measure of social presence began in the early 1970s, and is currently speeding up with the development of the conceptualization of social presence. Similar to the social presence concept research and
discussion, a commonly accepted paradigm for its assessment has yet to emerge (Biocca, Harms, & Burgoon, 2003); and far less frequent, is the discussion of the reliability and validity of those quantitative social presence measures (Biocca, Harms, & Gregg, 2001).

As foundational researchers in social presence, Short et al. (1976) initially developed the self-evident term and defined social presence as the degree of awareness of the other person in the interaction and their consequent interpersonal relationships. Based on their conceptualization of social presence, Short et al. (1976) also developed a social presence questionnaire using a semantic differential scale. Osgood, Suci, and Tannenbaum (1957) originally used this rating scale to measure the connotative meaning of concepts and asked respondents to rate conceptual meanings on a scale between two bipolar adjectives. Ever since then, semantic differential technique has been widely used in assessing physical presence (Darken, Bernatovich, Lawson, & Peterson, 1999; McCall, O’Neill, Carroll, & Benyon, 2004) however, Short et al. (1976) were the first to apply this technique in assessing social presence. Most of the current instrumentation in measuring social presence also adopted this semantic differential approach with minor variations (De Greef & IJsselsteijn, 2001; Gunawerda & Zittle, 1997; Lowenthal, 2012).

Currently, subjective and objective measures are the two general approaches to assessing social presence or its related concepts. Due to the nature of the specific research conducted, diverse backgrounds of researchers, and different conceptualizations of social presence, measurement selection can vary widely. Various measurements have been used to investigate people’s perceptions of presence in different environments, including self-reported questionnaires (Barfield & Weghorst, 1993; Garrison, Randy, Martha, & Tak, 2010; Nichols, Haldane, & Wilson, 2000; Witmer & Singer, 1998); qualitative measures (McGreevy, 1992; Murray, Arnold, & Thornton, 2000; Spagnolli, Varotto, & Mantovani, 2003); psychophysical measures (Snow & Williges, 1998; Welch, 1997); psychophysiological measures (Dillon, Keogh, Freeman, & Davidoff, 2000; Laarni, Ravaja, & Saari, 2003; Meehan, Insko, Whitton, & Brooks, 2001); behavioral measures (Ekman, 1982; Freeman, Avons, Meddis, Pearson, & IJsselsteijn, 2000; Prothero & Parker, 2003); and task performance measures (Basdogan, Ho, Srinivasan, & Slater, 2000; Slater, Linakis, Usoh, & Kooper, 1996).

Considering that social presence is relatively new and primarily a subjective experience, subjective measurements such as questionnaires, have mostly been adopted in assessing social presence and in social presence studies (De Greef & IJsselsteijn, 2001). A large majority of social presence research studies have adopted a subjective approach utilizing a questionnaire to assess participants’ perceptions of social presence. In this paper, the following key features of social presence measurement are discussed: (a) introduction and critical analysis of both subjective and objective measures, (b) major subjective measures used in assessing social presence, (c) discussions of the objective social presence measures, (d) the change of focus in social presence measurement throughout the development of social presence, and (e) certain criticisms on current measurement constructs of social presence.

2. Subjective Measures versus Objective Measures

Most of the approaches used in measuring social presence can generally be categorized into subjective measures and objective measures. According to Van Baren and IJsselsteijn (2004),
subjective measurement is an assessment approach in which participants are asked to pass their conscious judgment of their psychological state in relation to the mediated environment, while objective measurement is an approach assessing users’ automatic responses, without involving participants’ conscious deliberation. Subjective measures usually involve the use of self-administered questionnaires, continuous assessment of users’ sense fluctuations, content analysis of transcripts, and interviews. Concurrently, researchers also use more objective approaches to study people’s perceptions of their experience such as behavioral measures and the measures of users’ physiological responses including people’s change in heart rate, skin conductance, facial muscle tensions, respiration rate, blood pressure, reactions of the eyes, and muscular responses.

In a review of current social presence subjective measures and research, questionnaires are probably the most widely used tool in capturing users’ experience (Biocca et al., 2001; Biocca & Harms, 2003; De Greef & IJsselsteijn, 2001; Garrison et al., 2010; Gunawardena & Zittle, 1997; Kumar & Benbasat, 2002; Nowak & Biocca, 2003; Short et al., 1976; Thie & Wijk, 1998). Most researchers prefer using questionnaires in their studies because of the advantageous features such as high face validity that appears to measure the intended concept, inexpensive cost, great ease to administer, analyze and interpret, and no interference with the users’ experience (Insko, 2003; Van Baren & IJsselsteijn, 2004). However, Freeman, Avons, Pearson, and IJsselsteijn (1999) opposed that participants’ ratings on the questionnaire are sensitive to the effects of unrelated prior training sessions, and thus, creating the opportunity for unstable and inconsistent responses to be generated. Van Baren and IJsselsteijn (2004) also argued that questionnaires are usually retrospective and rely on users’ memories. Therefore, there could be invalid results if users cannot accurately reflect on their experiences. Insko (2003) also described that sometimes participants may not explicitly know the terms or concepts used in assessing their perceptions, and therefore, their responses could be varied depending on their personal interpretation.

Rooted in behavioral realism, objective measures are used variedly ranging from physiological measures, behavioral measures, task performance measures, to neural correlates (Van Baren & IJsselsteijn, 2004). There are certain discussions about the strengths and limitations of the objective measure used in assessing presence, and these discussions could also be applied in the social presence context. For example, Insko (2003) argued that measures such as the physiological approaches are more objective and they can capture the continuous and time-varying perceptions of users’ experience. Van Baren and IJsselsteijn (2004) also discussed that objective measures are relatively free from bias because they are generally not under users’ conscious control. However, as IJsselsteijn (2004) pointed out, objective measures such as behavioral approaches are prone to bias from the experimenter, who observes and interprets the behavior. He suggested that the risk could be minimized by having independent observers score the behavior and calculate the inter-rater reliability. In some cases, objective measures are difficult and complicated to manipulate. For example, Insko (2003) argued that some physiological differences such as skin temperature may change very slowly and thus it is difficult to track the difference caused by users’ mediated experience.

It is difficult to judge whether subjective measures are superior to objective measures, or otherwise, considering all the advantages and disadvantages of both types of measurement. However, when selecting the methodology to be utilized in research to assess social...
presence, researchers should be explicit about their own conceptualization of social presence and select appropriate strategies to effectively measure social presence in the specific research setting studied. It is also critical, if a research team develops an instrument, to report the validity and reliability of that instrument. As IJsselsteijn, De Ridder, Freeman, and Avons (2000) pointed out, the most fruitful approach is likely to combine both subjective and objective measures, thus producing different but complementary types of insights into the users’ mediated experience.

3. Subjective Measures of Social Presence

Though diversified subjective measures could be found in the research of presence, such as the use of questionnaire, interview, ethnographic observation, free format self-report, interaction analysis, and other means, the subjective measure of social presence is mostly limited to the use of the questionnaire. Many different types of questionnaires have been developed based on the developers’ different conceptualizations of social presence, their preferences of techniques for assessing users’ perceptions of social presence, and the context of media application. So far, the semantic differential approach has been the most widely adopted. This could be attributed to the influential work by Short et al. (1976) as well as the possible effectiveness this type of scale has in assessing social presence. In order to measure social presence through the use of telecommunications media, Short et al. (1976) introduced two experiments, using a series of seven-point bipolar scales, to assess the media capabilities including semantic differential items such as personal-impersonal, sensitive-insensitive, warm-cold, and sociable-unsociable. According to Short et al.’s (1976) research, the more personal, sensitive, warm, and sociable the medium is perceived to be, the higher social presence exists. Users were asked to assess the effect of medium they just experienced on the semantic differential scales, with an emphasis on the immediacy aspect of the communication. Short et al. (1976) introduced two different experiments, one with within-subjects design and the other with between-subjects design, in order to compare social presence from a variety of different systems such as face-to-face, closed-circuit television, and an audio system. In the first experiment, researchers found that 20 out of the 24 scales distinguished between audio medium and the two visual media to a statistically reliable degree.

Besides the introduction of several social presence questionnaires, qualitative and mixed methods measures of social presence such as content analysis and interview are introduced.

3.1 Short, Williams, and Christie’s Questionnaire

As initial investigators of social presence in the field, Short et al. (1976) proposed the concept of social presence in discussing social psychology of telecommunications. They also defined social presence as “the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships” (p. 65). Besides their influential definition, Short et al. (1976) also introduced a significant semantic differential method in measuring social presence based on the semantic differentiation work by Osgood et al. (1957).

In order to measure social presence through the use of telecommunications media, Short et al. (1976) introduced two experiments, using a series of seven-point bipolar scales, to assess the media capabilities including semantic differential items such as personal-impersonal, sensitive-insensitive, warm-cold, and sociable-unsociable. According to Short et al.’s (1976) research, the more personal, sensitive, warm, and sociable the medium is perceived to be, the higher social presence exists. Users were asked to assess the effect of medium they just experienced on the semantic differential scales, with an emphasis on the immediacy aspect of the communication. Short et al. (1976) introduced two different experiments, one with within-subjects design and the other with between-subjects design, in order to compare social presence from a variety of different systems such as face-to-face, closed-circuit television, and an audio system. In the first experiment, researchers found that 20 out of the 24 scales distinguished between audio medium and the two visual media to a statistically reliable degree,
with only four scales distinguished between the video medium and face-to-face, indicating in some way that social presence is a good discriminator between communications media. This is also consistent with Short et al.’s (1976) theory that social presence is the unidimensional quality of the medium and it varies significantly between different communications media. The similar bipolar semantic scales were also applied in the second study, to discriminate between two variations of the same telecommunications system. Researchers also found consistent results with the hypothesis that there is higher level of social presence in video with close-up picture than with small images. Though Short et al. (1976) described that their semantic differential questionnaire could discriminate and capture social and emotional capabilities of the medium, no further detailed validity and reliability report could be found from their work.

Short et al.’s (1976) initial and influential work has been referred to in many social presence studies and the use of certain variations of their semantic differential technique to assess students’ perceptions of social presence has been recognized. However, some scholars argue that this approach may not lead to proper measures of social presence. As Nicholls (1984) pointed out, respondents may not reliably identify the cause of attitudes, nor can they make an explicit judgment of how the medium causes their social presence. Biocca, Burgoon, Harms, and Stoner (2001) also argued against Short et al.’s (1976) measure, maintaining that rather than studying the direct attributes of the medium per se, the measurement of social presence should be based on the properties of the communication interaction. Biocca, Harms, and Burgoon (2003) contended that the semantic differential method was used to collect users’ social perceptions about the medium, instead of their judgment about the state within the medium. Nowak (2001) argued that Short et al.’s (1976) bipolar semantic social presence measurement is inadequate and inconsistent with their definition. However, she still adopted six items from Short et al.’s (1976) questionnaire without much change and applied them to investigate the effects of agency and anthropomorphism on social presence.

Tu (2002) also argued that there are difficulties in applying Short et al.’s (1976) measurement. On the one hand, the items are too general to measure user’s perception of social presence in computer-mediated communication (CMC) environment. Tu (2000b) argued that other variables such as topics, privacy, and task could also affect the degree of social presence, and they should also be considered in the measurement. On the other hand, Tu (2002) argued that the semantic differential method is not accurate because respondents may ascribe different meanings to the keywords.

### 3.2. Adapted GlobalEd Questionnaire

Gunawardena and Zittle (1997) introduced that CMC is generally believed to have low social presence, compared to face-to-face communication, because of its lack of nonverbal communication cues. However, they also argued that field researchers in CMC such as Walther (1992) and Baym (1995) also reported the development of online communities and warm friendship. Same to Short et al. (1976), Gunawardena and Zittle (1997) argued that social presence evolved from Argyle and Dean’s (1965) concept of intimacy and Wiener and Mehrabian’s (1968) concept of immediacy. Social presence studies in the traditional face-to-face environment indicated that teacher immediacy was a good predictor of students’ affective learning (Christophel, 1990; Gorham, 1988) however; Gunawardena and Zittle (1997) argued that few studies have been conducted to determine the effect of social presence on learner’s satisfaction in a CMC environment.
In order to investigate the influence of social presence in the CMC context, especially how effective social presence was for overall learner satisfaction in the computer conference setting, Gunawardena and Zittle (1997) conducted a follow-up study based on an inter-university “GlobalEd” computer conference in 1993 that provided a forum for graduate students in distance education to share and discuss research and experience in distance education by using CMC. They also developed their questionnaire based on the GlobalEd Questionnaire used at that time. A total of 52 five-point Likert-scale items were used from the original 61-item GlobalEd questionnaire with a focus on nine areas that included: social presence, active participation in the conference, attitude toward CMC, barriers to participation, confidence in mastering CMC, perception of having equal opportunity to participate in the conference, adequate training in CMC at participant’s site, technical skills and experience using CMC and overall satisfaction with the GlobalEd conference. Among the 52 items, 14 questionnaire items were used specifically to assess social presence and 10 items were used to assess students’ overall satisfaction. Gunawardena and Zittle (1997) also blended the semantic differential scales used by Short et al. (1976), with a focus on the immediacy construct. They adopted 17 five-point bi-polar scales including personal/impersonal, immediate/non-immediate, interactive/non-interactive, sensitive/insensitive, social/unsociable, and colorful/colorless that assess students’ feelings towards the use of CMC. A stepwise regression procedure was used to examine the relation between social presence and overall satisfaction.

In order to ensure the validity of this social presence measure, Gunawardena and Zittle (1997) only used six items from the original 17 five-point bi-polar instrument in order to specifically measure the social aspect of the medium. The new social measure was further validated with strong, positive correlations between bi-polar social indicators and social presence. To ensure the reliability of the measure, a stepwise regression procedure was used twice on different predictors, and social presence was consistently found to contribute to a large proportion of the variance.

Gunawardena and Zittle’s (1997) social presence measurement scale has been widely adopted by other researchers. For example, Richardson and Swan (2003) used a modified social presence scale by Gunawardena and Zittle (1997) to explore the relationship among students’ perceptions of online social presence and their perceived learning and satisfaction with the instructor. Though Gunawardena and Zittle’s (1997) measure used both pi-polar scales and Likert-scale items to measure social presence, Tu (2002) argued that the new instrument did not consider important social presence variables such as privacy, recipients, and topics; nor were the questions created for general students. Therefore, the new social presence instrument was not able to capture a thorough perception of social presence.

3.3. IPO Social Presence Questionnaire (IPO-SPQ)

In order to investigate the effects of video communication on social presence, especially through the use of PhotoShare tele-application, an advanced telecommunication platform, IPO Social Presence Questionnaire (IPO-SPQ) was developed to assess social presence with telecommunication applications. De Greef and Ijsselsteijn (2001) designed this questionnaire based upon the perception that social presence was quite distinct from physical presence, or the sense of “being there” in a mediated environment. They also agree with most scholars in social presence (Biocca et al., 2003; Gunawardena, 1995; Gunawardena & Zittle 1997; Reio & Crim, 2006; Rettie, 2003; Rice, 1993; Rourke et al., 2001; Sallnas, 2005;
Tu, 2001) by contending that intimacy and immediacy are particularly important factors for social presence.

De Greef and Ijsselsteijn (2001) constructed their final IPO-SPQ with a combination of two different approaches. The semantic differential technique, originally developed by Osgood et al. (1957) and popularized by Short et al. (1976) in the context of social presence, was applied in the first 12 items. Subjects were asked to rate the communication media on a series of bipolar scales such as insensitive-sensitive, cold-warm, impersonal-personal, and passive-active. Five other Likert-type 7-point agree–disagree scale items were created based on subjective attitude statements about system qualities. Besides the social presence items, the questionnaire also included a number of general items involving usability, communication in general, and audio communication.

De Greef and Ijsselsteijn (2001) tested the IPO-SPQ in an experiment to investigate the effects of video (audio only vs. audio and video), participant role, and pointing function of the application. After a reliability analysis of the items of responses, three items with a low item-total correlation were deleted. Cronbach’s alphas for social presence items were satisfactory, ranging from 0.72 for the attitude statements to 0.90 for the semantic differentials, indicting the items were consistently measuring the same quality. In addition, a substantial correlation of 0.58 was also achieved between the two social presence scales. To determine the validity of the questionnaire instrument, De Greef and Ijsselsteijn (2001) found that subjects were able to distinguish between media on a within-subjects comparison (audio only vs. audio and video), which was consistent with social presence theory prediction, but they could not reliably judge each medium individually on a between-subject comparison (male vs. female). However, De Greef and Ijsselsteijn (2001) cautioned against the generalization of their questionnaire in the physical presence study because of distinct features of the questionnaire in assessing social presence.

3.4. The Networked Minds Questionnaire

Researchers from the MIND (Media, Interface, and Network Design) labs (Biocca, Burgoon, Harms, & Stoner, 2001; Biocca et al., 2003; Biocca & Harms, 2002; Biocca, Harms, & Gregg, 2001) have been developing conceptualizations of social presence in the mediated environment and they call it as the Networked Minds Social Presence. At the same time, they have also been developing measurement of social presence based on their conceptualization of social presence.

Biocca, Harms, and Gregg (2001) provided a tentative definition of social presence by describing it as “the moment-by-moment awareness of the co-presence of another sentient being accompanied by a sense of engagement with the other” (p. 2). They described social presence as a composition of three underlying dimensions varying from co-presence, psychological involvement to behavioral engagement, from superficial to a deep sense. Deeper levels of social presence are activated based on the earlier layers of cognition. Biocca et al. (2001) thus created the Networked Minds Questionnaire based on their early perceptions of social presence.

In accordance with their preliminary research on social presence, Biocca et al. (2001) initially created a pool of 88 behavioral indicator items to assess social presence in its three dimensions: co-presence, psychological involvement to behavioral engagement. The questionnaire items were also paired to reflect both the participants’ own feelings and the participants’ perception of their communication partners’ feelings.
For example, item “I hardly noticed another individual” is matched by the item “the other individual didn’t notice me in the room.” A total of 69 items were retained after an initial analysis of face validity and content validity. In order to examine the factor structure and concurrent validity of their Networked Minds measure of social presence, Biocca et al. (2001) applied the questionnaire in a within-subjects experiment comparing face-to-face interaction with mediated teleconferencing. Participants were assigned to a purely verbal, non-emotional task of ranking the importance for survival in the desert. A factor analysis was carried out and internal consistency was calculated based on the obtained questionnaire scores. Items were removed when their correlations failed tests of internal consistency and only 38 out of the original items were retained. All questions utilized a 7-point Likert scale. For the final emerging factors and items, a concurrent validity of the scales was supported because differences indicated by scores through an analysis of variance were mostly coherent with the researchers’ predictions based on the social presence concepts and task nature. The scales also achieved a satisfactory average reliability of 0.77 through an analysis of the internal consistency data.

Besides Biocca et al.’s (2001) measurement, several other variations of the Networked Minds questionnaires have also been developed by researchers within the MIND labs such as Biocca and Harms (2003), Harms and Biocca (2004), and others. The importance of studying non-verbal cues in mediated social interaction has been iterated by many researchers (Burgoon et al., 2002; Walther, 1996; Walther & Burgoon, 1992), and the Networked Minds Social Presence Measurement Inventory was exemplary because it utilizes using behavioral indicators to support the self-report measure. However, as Biocca and Harms (2003) argued, the inventory could be applied in many settings such as the use of traditional media including picture and film, but it may not be suitable to measure the social presence of humans in non-mediated settings.

3.5. The Social Presence and Privacy Questionnaire (SPPQ)

According to Tu and McIsaac (2002), social presence is comprised of three dimensions that include social context, online communication, and interactivity, and they also emerge as important factors in establishing a sense of community among online learners; besides that, privacy is also an important element in the level of comfort for online students. Following a close review on the current social presence measures, Tu (2002) argued that current instruments are unable to measure the complicated issue of online social presence. Based on Steinfield’s (1986) CMC attitude instrument and Witmer’s (1997) perceived privacy instrument, as well as elements of social learning theory, Tu (2000b, 2002) redefined three dimensions of social presence and developed the Social Presence and Privacy Questionnaire (SPPQ) to evaluate CMC users’ perceptions of social presence and privacy. Steinfield’s (1986) instrument was originally designed to examine business users’ attitudes toward CMC. The instrument consisted of 16 items using a semantic differential scale. Witmer (1997) developed her instrument with 32 items using the Likert scale for an online newsgroup. While adopting these two instruments, Tu (2002) removed several items that were specific for groups for which they were initially developed. At the same time, he also added content based on his review on existing studies on social presence. Tu’s (2002) initial SPPQ instrument contains 59 items on a Likert scale format with demographic data.
Tu (2002) adopted content validation and construct validation to validate the instrument. For content validation, Tu (2002) categorized the content into eight different objectives including social presence, privacy, utility, ease of use, interactivity, language, CMC experience and competence, and demographics. Then he invited a panel of five qualified social presence content experts to evaluate the objectives by completing a questionnaire. The experts were asked to perform an item-matching task and select the best match for a specific objective. Judges’ matches on the items of the questions were collected and analyzed to determine item validity. Several items were either revised or discarded based on the statistical match data (lower than 60%) obtained from the content expert validation. The researcher also found no significant difference of item matching of the five content experts through chi-square test. The revised questionnaire includes 17 social presence items and 13 privacy items, with five-point Likert scale and demographic data.

Realizing that the content validation alone was not sufficient, Tu (2002) also conducted construct validation to further validate the instrument. In the construct validation, Tu (2002) invited 310 in-service and pre-service teachers to respond to the content validated questionnaire in both online and article-and-pencil formats. Exploratory factor analysis was used from responses on the 30 question-item questionnaire to determine the emergence of the dimensions of social presence. Five factors including social context, online communication, interactivity, system privacy, and feeling of privacy emerged and accounted for 82.33% of the variance, with Cronbach’s alpha values ranging from .74 to .85. Three items were removed from the loading, with a cutoff of .45. This result is consistent with literature review that social context, online communication, interactivity, and online privacy are important factors in impacting the degree of social presence.

Tu (2002) also checked reliability of the instrument by splitting factors into an online questionnaire, article-and-pencil questionnaire, and the three CMC systems (E-mail, bulletin board and real time discussion). In his research, Tu (2002) found that the coefficients and factor structures are almost the same and significant correlations were also found between all emerged factors. Though this instrument appeared to be powerful with the consideration of its validity and reliability, some researchers criticized Tu’s SPPQ instrument by stating that it had limited applications to other contexts. For example, Henniger and Viswanathan (2004) argued that the SPPQ model was only examined in the text-based context, and therefore, its scope was very limited due to the restriction of text-based computer-mediated communication.

3.6. Qualitative and Mixed Methods Measures

Although subjective measures usually include the use of questionnaires, continuous measurement, qualitative measures, psychophysical measures, and subjective corroborative measures, questionnaire by far is the largest category within the group in study of social presence (Van Baren & IJsselsteijn, 2004). Besides the wide adoption of questionnaires in measure of social presence, qualitative measures such as content analysis, observation, interview, and case study are also used by researchers in assessing social presence (Lowenthal, 2012).

One of the most cited qualitative social presence measure is the content analysis method used by Rourke, Anderson, Garrison, and Archer (2001) in coding transcripts from courses in text-based computer conferencing. Garrison, Anderson, and Archer (2000) originally developed the framework of social presence, model of community inquiry, which constitutes three overlapping key elements: cognitive presence, social presence, and
teaching presence. They argued that deep and meaningful learning occurs through the interaction of these three core elements within the community of inquiry. Through an iterative process of research and content analysis on social presence, Garrison et al. (2000) categorized social presence into open communication, emotional expression, and group cohesion. In order to better reflect the nature of emergent indicators of social presence, Rourke et al. (2001) relabeled the three categories as interactive responses, affective responses, and cohesive responses. Rourke et al. (2001) postulated that indicators developed within those categories could reveal the level of social presence in an online community of inquiry. Low scores would indicate a cold and impersonal social environment while high frequencies would indicate that the environment was warm and collegial.

Based on the construct of social presence they suggested, Rourke et al. (2001) developed 12 indicators to reflect the three categories of social presence including affective responses, interactive responses, and cohesive responses. Indicators included items such as expression of emotions, user of humor, self-disclosure, continuing a thread, quoting from others’ messages, asking questions, use of vocatives, and so on. Content analysis methodology was used to test the efficacy and reliability of those indicators. Rourke et al. (2001) applied their template of indicators in two graduate courses, supported primarily by computer conferencing, and then coded the selected transcripts from these two courses. Raw number of instances of social presence and the total number of words were calculated and used for reliability tests. A total of 90 posted messages from the fifth week were selected for analysis in one course and a total of 44 messages from the sixth week of the conference from another course were selected for analysis. Three researchers collaboratively worked together to code messages while two other coders were also invited to independently code the selections. Interrater reliability was calculated upon completion of the coding for percent agreement. A high average of coefficient of interrater reliability was found in two courses, ranging from 0.91 on first transcript to 0.95 on the second transcript. The social presence density in one transcript was considerably higher than that in another transcript, which confirmed their intuitive impressions based on their reading of the sociability and educational effectiveness from the transcripts, indicating that the instrument is valid, thus being able to expose and quantify important differences in social presence. Similar content analysis technique was also adopted by Ubon and Kimble (2004), who examined transcripts from the electronic bulletin boards in order to investigate the development of social presence among online members in text-based online learning environment.

Beside the use of content analysis method, interview is another frequently used method in measuring social presence and it helps to provide researches with a different perspective of students’ perceptions of social presence. In examining Chinese student’s interactions and perceptions of social presence in online learning environments, Tu (2001) conducted an in-depth interview with six Chinese graduate students studying at a university in the United States. They were all enrolled in three different distance learning courses in a system with e-mail, bulletin board, and real-time chat functions. At the 12th week of the semester, the researcher conducted six semi-structured in-depth interviews with students, along with direct observation of conversation and document analysis, to understand Chinese students’ views on three different CMC systems involving issues such as privacy, social relationships, task orientation, online communication, and social interaction. In
his earlier study, Tu (2000a) defined the three dimensions of social presence in the context of online environment that included social context, online communication, and interactivity. He also argued that privacy was an important issue related to social presence though not a significant dimension. Based on this framework, Tu (2001) explained the qualitative data in three dimensions of social presence (social context, online communication and interactivity) and issues of privacy. He found that the three dimensions of social presence can affect Chinese students’ perceptions of CMC. Their feeling of private/public was identified as an important factor related to the level of social presence. Tu (2001) also argued that the level of social presence of Chinese students was not only affected by attributes of different CMC systems, but also by their subjective perceptions. The results also indicated that it was important to consider students’ local culture, language skills, and keyboarding skills when integrating CMC into an online learning environment.

Besides Tu’s (2001) ethnographic approach in studying social presence, interview is also used by mixing with other methods. For example, in order to investigate the relationship between perceived social presence and projected presence in online discussions, Swan and Shih (2005) used the interview method, along with the use of content analysis and a questionnaire. The questionnaire they used was adopted from Richardson and Swan’s (2003) survey with 5-point Likert scales to find out students’ perceptions of social presence, satisfaction with instructors, perceived learning, and perceptions of interaction. Five respondents with the highest ratings and the five respondents with the lowest ratings of perceived social presence were identified and their online postings were coded based on Rourke et al.’s (2001) categories of social presence. An inter-rater reliability of 0.94 was found before consensus was made between two independent coders. High and low social presence students were also further interviewed to explore their perceptions and experiences in online discussions. The interview questions mainly focused on students’ message preparation, the ways in which they responded to others, how they formed impressions, and their feelings about instructors. Transcripts of interviews were discussed using thematic cross-case analysis to explore the ways in which they perceived their online discussion experiences and potential differences between the two groups. Swan and Shih (2005) found that students who perceived the most presence of others in online discussion also presented more social presence indicators in their messages. This indicated that the perception of social presence was related to its presentation.

4. Objective Measures of Social Presence

From a review of current social presence measures, it is obvious that much emphasis has been placed on the subjective measures, especially with the widespread use of questionnaires in assessing users’ perceptions of social presence. At the same time, there is also a growing interest in objective measures that focus on behavioral or physiological responses to media, considering the potential instability of subjective measures (De Greef & Ijsselsteijn, 2001). According to Blascovich (2000), we could use some psychophysiological indicators such as heart rate, blood flow, skin conductance, or fMRI to measure social psychological responses.

To date, there is a large body of research conducted to measure presence using the objective approach such as observation of facial expression (Ekman, 1982), postural responses (Freeman et al., 2000), pointing behavior (Slater, Usoh, & Chrysanthou, 1995),
and reflex responses (Nichols et al., 2000). Though supportive of the use of objective measures in assessing social presence, Biocca et al. (2003) argued that we were still unaware of their use explicitly to measure mediated social presence. This could explain the limited amount of objective research approaches used in measuring social presence. Most often, objective measures are integrated into social presence research as a supplementary support to subjective measures.

Bailenson, Blascovich, Beall, and Loomis (2001) conducted their experiment to explore interpersonal distance in a virtual environment. According to Argyle and Dean’s (1965) equilibrium model, mutual gaze and personal space are inversely related to each other; mutual gaze could non-verbally promote intimacy and it will be decreased by increases in personal space. Based on this equilibrium model, Bailenson et al. (2001) wanted to test the inverse relationship of mutual gaze with personal space by varying the degree of mutual gaze between virtual agent and the participant. The experiment was conducted in which participants were involved in a memory task with virtual agents. Besides administering virtual social presence questionnaire, they also captured participants’ social responses by tracking the position and orientation of participants in interaction with a precision tracking system. The head mounted display was used to render the virtual environment. According to the researchers, the orientation of the participant's head was tracked by a three axis orientation sensing system, and the location of the participants head was tracked three dimensionally by a passive optical position sensing system. With this hybrid tracking system, researchers could record accurate sensory input when participants turn head or walk. For each participant, they objectively recorded the minimum distance between the center point of the participant's head and the center-point of the agent’s head, and the amount of time participants spent inside the agent’s intimate space during the experiment. They found that the equilibrium model can also possibly be applied in the virtual environment. Though objective measures can be found in literature in assessing boundary concepts such as behavioral presence (Meehan et al., 2001), spatial presence (Laarni et al., 2003), and telepresence (Sheridan, 1992), limited studies address the objective measures of social presence.

As an emergent area of interest in educational technology ventures, methods such as social network analysis has been used to understand human relationship (Shea et al., 2010) and optimize learning and its environment (Buckingham & Ferguson, 2012). Though still at its initial development stage, measuring social presence using social network analysis has been explored by researchers (Choi & Strobel, 2012; Mika, 2007; Shea et al., 2010). For example, in order to rigorously capture the nuances of social presence in the online setting, Choi and Strobel (2012) combined Social Network Analysis with the assessment of social presence and they found that this modified instrument was a meaningful extension to existing measures. In their study, besides using the social presence measure, researchers adopted the social network analysis data mining method to analyze the level of “betweenness” of students through their discussion board postings. The five units in their coding analysis included sentence, paragraph, message, thematic unit, and illocutionary unit and proper value was also assigned to different coding themes. Choi and Strobel (2012) found in their study that social network analysis was a good way to show different aspects of social presence and gave greater insights into students’ online communication behaviors.
5. Criticism of Current Measures of Social Presence

With the development of social presence research, one fundamental debate on the measurement of social presence varies around what we are supposed to measure. Should we measure media properties or other variables such as users’ perceptions including their attitude, feelings towards media, or degree of interpersonal interactions? Many researchers such as Short et al. (1976) developed their measurement instrument to study the properties of a medium. As initial investigators of social presence, they postulated that social presence was an important factor in a communication medium and they considered social presence as the unidimensional quality of the medium itself (Short et al., 1976). Therefore, they created their measure by asking respondents to directly rate the properties of medium for social presence. However, with the development of CMC technologies and the expansion of emerging features of new media, more researchers have realized that social presence is a phenomenon instead of the attribute of a specific medium (Biocca et al., 2003; Gunawardena, 1995; Kreijns, Kirschner, Jochems, & Van Buuren, 2004; Tu, 2002). Thus, they started to measure social presence with a different emphasis. As Biocca et al. (2003) pointed out, social presence measure should assess the phenomenological state that varies with the medium, knowledge of the other, content of the communication, and social context. They further elaborated that though the medium may still affect a fluctuating level of social presence, measure of social presence should not specifically direct attributions about the medium per se. Kreijns et al. (2004) constructed a self-reporting social presence scale that tried to capture the psychological sensation associated with social presence. They designed their instrument to measure the degree that individuals feel about their transporting experience. This difference in measurement emphasis seems to be caused by researchers’ different conceptualizations of social presence. Most researchers now have moved away from the simple media attribute study and begin to explore users’ overall perceptions of their mediated experience. As Kehrwald (2008) pointed out, social presence studies have moved on to reflect the communication experience.

Another criticism towards the measure of social presence is that there has not been a widely accepted, validated, and generalized measure across different media or situations. Most instruments developed by researchers are usually tailored for a specific technology or environment such as F2F interaction, email system, or virtual environments (Biocca et al., 2003). However, as Blocher, Amato, and Storslee (1996) pointed out, although these systems share some common features, they vary in their operations and functions and the difference could lead to the different degrees of social presence. For example, Henniger and Viswanathan (2004) found that the Tu’s SPPQ model (2002) was only examined in the text-based context, and therefore, it cannot be generalized in other research settings due to the restriction of text-based computer-mediated communication. Biocca et al. (2003) also elucidated that social presence is a phenomenon that is independent of a specific technology and a useable measure should be able to measure social presence across most media.

Using questionnaires as a measurement tool is the target of another wave of criticism. Research has indicated that subjective measures such as questionnaires are potentially unstable or ineffective in measuring presence or co-presence (Bailenson et al., 2004; Bente, Ruggenburg, Tietz, & Wortberg, 2004; Ellis, 1996; Freeman et al., 1999). Though this statement was made referring to the measures of presence or co-presence, it can also be
applied to social presence context because of its similar dynamic nature. Hostetter and
Busch (2006) also noted that solely relying on questionnaire data could be misleading
because respondents tend to provide socially desirable and acceptable answers. Bailenson
et al. (2004) also argued that one of greatest limitations in using questionnaires is that
participants may not always judge their own thoughts or feelings accurately and may
misreport affective or cognitive responses to stimuli. Researchers also argue that some
questionnaires are not adequately validated. For example, Lin (2004) argued that though
Gunawardena and Zittle’s (1997) instrument provided concurrent validity, the scale itself
was not validated however, the instrument had been adapted to measure social presence in
spite this potential flaw.

6. Summary

The development of valid and reliable social presence measure is dependent on
its solid theoretical framework (Cui et al., 2012). A comprehensive understanding
of the concept and deep insight into the phenomenon will lead to the development of
valid and reliable measures of social presence. However, the current body of literature
surrounding social presence provides diverse concepts of the term and can be categorized
into different dimensions and levels. This leads to the diversified measures of social
presence in different studies. Therefore, an explicit and widely accepted theory of
social presence should be produced before precise measures of social presence can be
achieved. Social presence study initially began with an emphasis placed on the media
attribute, and now researchers have gradually moved on to consider user’s perceptions of
interpersonal relationships. Social presence measure is undergoing the same changes with
the influence of this foundational change. Lowenthal and Dunlap (2013) reiterated that
the development of social presence instrument should also align with foundational framework
and its studies.

Both subjective measures and objective measures have been used in assessing
users’ perceptions of social presence, and they are equally effective considering their
distinguished features. Questionnaire is currently the most frequently used method
in assessing social presence considering the flexibility and ease in administering, but this
method also has some potential flaws and may not be very reliable when used independently.
In order to avoid certain inner drawbacks in measuring social presence, IJsselsteijn et al.
(2000) suggested combining both measures thereby capturing all types of information.
Currently, there are also few objective measures used in assessing social presence,
and more objective approaches should be explored in order to provide a comprehensive
evaluation of social presence. Social presence is a social and psychological phenomenon that
is independent of specific media or technology (Biocca et al., 2001). Therefore, the measures
should be constructed without the constraint of a specific technology and should be able to be
replicated across different media and research settings. For future developers of social
presence measures, they should construct items and indicators of social presence measure in
consideration of their generalization across different media. In the meantime, researchers
should also consider applying various approaches to assess social presence such as
behavioral tracking, learning analytics, and social network analysis. Recommendations
for future research on social presence and the development of effective instrumentation to
measure the construct should (a) have solid conceptualization of social presence, (b)
clearly report instrument validity and reliability in published research to allow for the use and
selection of effective, valid, and reliable social presence instruments in the field, (c) ensure
that media concerns and generalization are taken into consideration, and (d) assess social presence with diverse approaches.

References


Evaluating Online Social Presence: An Overview of Social Presence Assessment


Contact the Author

Guoqiang Cui
Yantai University
Email: andyguoqiang@hotmail.com