Comparison of Social Presence in Voice-based and Text-based Asynchronous Computer Conferences

Karen King, Ph.D.
East Tennessee State University
kingk@etsu.edu

Timothy J. Ellis, Ph.D.
Nova Southeastern University
ellist@nova.edu

Abstract

The significance of social presence in asynchronous computer conferencing has become an increasingly important factor in establishing high-quality online learning environments. Levels of social presence exhibited in asynchronous computer conferences influence students’ perceptions of learning and satisfaction levels in Web-based courses. Recently, faculty teaching online courses have begun to use voice-based asynchronous conferencing tools with little research to support the appropriateness of the media. Deployment of voice-based technology as a pedagogical tool is delivered at a considerable cost to higher education institutions. These tools are often marketed based on the effectiveness of the technology in a learning environment. However, according to this study, there is no apparent benefit in using voice-based rather than text-based technology tools to facilitate asynchronous computer conferences in a Web-based learning environment.

1. Introduction

Course delivery via the Internet continues to grow at a rapid rate [1] and text-based asynchronous computer conferencing remains an essential tool for promoting student-to-student and student-to-teacher interaction and for promoting enhanced social presence and learning [2, 3, 4]. Recently, the use of voice-based technologies has been introduced in Web-based courses as an enhancement to the asynchronous computer conferencing tool [5, 6, 7, 8, 9, 10]. Evidence in the literature regarding the use of voice in online environments focuses on the potential of voice tools to increase social presence by the delivery of communication cues such as paralanguage [11, 12]. However, there has been little empirical evidence to indicate that the use of voice in asynchronous conferences will improve levels of social presence and, ultimately, learning. This study compared social presence levels as measured by interaction patterns in voice-based and text-based asynchronous computer conferences. It should be noted that this study was limited to the evaluation of voice-enhanced discussion forums as a computer mediated communication technology available in an asynchronous learning environment.

1.1. Problem Statement

A review of higher education institutions indicates a growing trend in the adoption of asynchronous voice tools in Web-based learning environments [5, 6, 7, 8, 9, 10]. Bishop and Cates [13] reported that the use of short speech recordings increased comprehension. Additional studies indicate that voice has the potential to deliver nonverbal communication cues, such as intonation, pace, timing, and periods of silence [11, 15, 29]. This nonverbal communication, or paralanguage, has the potential to improve the social presence in an asynchronous computer conference [11, 14, 15].

However, there is conflicting evidence in the literature regarding the benefits for inclusion of voice in asynchronous computer conferencing in Web-based courses. A case study for the use of asynchronous voice in language instruction reported improvement in imprecise data such as student enthusiasm, comfort levels, and attitude, when using voice tools for asynchronous communication [16]. An additional case study reported student dissatisfaction with asynchronous voice tools and further reported that students abandoned the communication method early in the Web-based course for alternate communication methods [17]. In addition, there are no studies that have specifically examined social presence in voice-based asynchronous computer conferences.

1.2. Goals

The goal of this study was to compare social presence in voice-based and text-based discussion forums to determine if voice discussion demonstrates levels of social presence equal or superior to text discussion. Social presence was measured by three basic categories of interaction described by Rourke et al. [2]: affective, communicative reinforcement, and cohesive interaction patterns.
2. Background

Although oral articulation is an important educational component [18], effective implementation of voice to enhance asynchronous computer conferencing must be linked to well-established learning theory [19]. A fundamental element in establishing a link from learning technologies, such as voice-based asynchronous computer conferencing to theory, is the pedagogical model. Pedagogical models are based on theoretical constructs regarding cognition and are the foundation of learning theory. The result is development of learning strategies that influence the use of specific learning technologies [19]. Examination protocols for the effectiveness of learning technologies must also be grounded in well-established theory to accurately determine if specific learning technologies such as voice-based asynchronous computer conferences actually improve learning [20].

Constructivist learning theory posits that cognitive development or learning requires social presence as an integral part of a learning environment. The theory of social presence was further developed [15] to incorporate the concept of immediacy, which can be defined as “those communication behaviors that enhance closeness to and nonverbal interaction with another” (p. 203). Subsequent research [21] defined social presence as “the salience of the other in a mediated communication and the consequent salience of their interpersonal interactions” (p. 65). Furthermore, the lack of nonverbal cues would result in a negative perception of the communication experience. Contemporary research has established a direct link between social presence and the improvement of instruction [2, 22, 23].

Social presence theory establishes that communication occurs between people and not between individuals and inanimate objects such as computers [23]. Low social presence results in a disconnection of group members and poor group dynamics, while high social presence indicates members are more engaged and involved in group tasks. Social presence is potentially impacted by the ability of individuals to collaborate effectively through technologies capable of transmitting vocal inflections, paraverbal utterances, and ambient sounds, indicating the use of voice could enhance social presence in computer-mediated conferencing environments. The collective research of these authors concludes that social presence is correlated with student satisfaction levels and increased success in learning outcomes.

A conversational voice narration as opposed to a formal narrative style has been identified [24] as a factor that has a significant impact on learning in computer-mediated communication. This personalization principle creates a higher sense of social presence and encourages learners to actively engage in the instructional environment. Voice has also been shown to be an important factor in creating a sense of social presence in which learners will consider the narrator a reciprocal collaborator. The use of voice has been shown [18] to promote relationship building in an online learning community.

Despite this foundation of research, there is, however, a paucity of research available to determine if voice-based asynchronous computer conferencing tools will increase the level of social presence over that evidenced in text-based asynchronous computer conferencing. In addition, the integration of technology into a course is often deployed with little or no research evidence regarding the effectiveness of the learning technology. The use of technology as a pedagogical tool is delivered at a considerable cost to higher education institutions, an environment which typically operates on limited financial resources [25]. There are numerous tools available that are marketed based on the effectiveness of the technology in a learning environment. Since the type of technology used for synchronous communication in an online learning environment can have a profound effect on the communicative behaviors of the participants, it is critical to determine the appropriateness of specific technologies in online learning environments [26, 27].

It is also important to determine if there are unacceptable detrimental effects on learning associated with specific technologies. Multimedia environments, for example, present instructional material using multiple forms of communication. However, the multimedia-enriched learning environment has been shown to potentially increase the cognitive load of the learner resulting in a decrease in the cognitive processing capacity of the learner [28]. For example, an examination of the simultaneous delivery of text with narration to determine the impact on learning revealed a negative impact on learning when textual information is added to a self-explanatory diagram or graph [29]. According to the cognitive load theory, when learners must split their attention between simultaneous text and graphics, it overburdens the working memory capacity of the learner [30]. Computer-based instructional materials consisting of concurrent diagrams and auditory verbal information, on the other hand, have been shown to be more efficient than those consisting of concurrent diagrams and on-screen text with redundant auditory information. Additional studies [28, 29] indicate that the working memory of the learner can process only a few elements at any one time. When words are presented as spoken text, they do not overload the visual working memory and allow for more active cognitive processing by the learner. Also, if information is presented via graphics and narration, learners are able to select both with no
cognitive overload. However, a cognitive overload results if the information is presented via graphics, narration, and on-screen text [14].

The human voice can be a critical factor in creating a sense of social presence in computer-mediated conferencing. Modest social cues such as voice accent, intonation, and rhythm of speech can influence the learning environment. These social cues affect how much cognitive processing is required to understand the content and how much of that processing ability is diverted away from meaningful processing to interpret the narration [29].

The impact of personalized messages on cognitive load has also been examined [28]. Elimination of extraneous sounds, presentation of words as narration rather than on-screen text, and presentation of narration with graphics rather than following the presentation of the graphics, were techniques used to decrease the cognitive load in a computer-based learning environment. Results indicated that cognitive load is reduced when a human voice is used for narration rather than a synthesized voice. In addition, the use of a conversational tone of voice rather than a formal speech style enabled learners to use their available cognitive capacity for active learning as opposed to using their cognitive capacity to process the media delivery. Personalizing messages in computer-based environments presents a friendly or helpful atmosphere, increases the sense of social presence for the learner, and encourages the learner to engage in active cognitive processing. However, large amounts of verbal information should not be presented without corresponding visual information [29].

Technological advancements have enabled the use of voice-based asynchronous computer conferencing tools in Web-based courses as an alternative to text-based discussions [16]. However, using voice-based asynchronous computer conferencing tools requires considerable financial resources from university administration and significant support resources for faculty and students [25]. To justify the costs associated with deployment of voice-based asynchronous computer conferencing in online learning environments, it is essential to determine if social presence as measured by interaction is improved when voice-based asynchronous computer conferencing tools replace text-based computer conferencing tools. To date, there is little empirical evidence in the literature to indicate that the increased cost required to deliver voice-based asynchronous computer conferencing in Web-based courses will result in an environment superior to text-based asynchronous tools [16, 17].

However, evidence does exist regarding the benefits of conversational voice, paralanguage, voice dynamics, and voice style to increase social presence and reduce cognitive load in computer-based learning environments [28, 29]. The use of the Wimba voice tool enables students and faculty to record discussion entries in their own voice. Therefore, the conversational tone, paralanguage, and voice style have the potential to enhance social presence in the discussion without producing a cognitive overload for the listener. Educators must use empirical evidence to determine which technology to use for certain types of tasks in an online learning environment [26].

3. Methodology

This investigation examined the social presence as measured by interaction patterns in the voice-based and text-based asynchronous computer conferencing content of a Web-based course in a southeastern United States university. The following fundamental research question was explored in an attempt to achieve the goal of this study: How do social presence levels in voice-based asynchronous computer conferencing compare to social presence in text-based asynchronous computer conferencing as measured by interaction? The following hypotheses were tested to develop an answer to that research question:

- **Hypothesis 1** – There is no difference in the affective interactions of voice-based asynchronous computer conferencing when compared to affective interactions in text-based asynchronous computer conferencing.
- **Hypothesis 2** – There is no difference in the communicative reinforcement interactions of voice-based asynchronous computer conferencing when compared to communicative reinforcement interactions in text-based asynchronous computer conferencing.
- **Hypothesis 3** – There is no difference in the cohesive interactions of voice-based asynchronous computer conferencing when compared to cohesive interactions in text-based asynchronous computer conferencing.

The quasi-experimental study design included 86 undergraduate students. Typical compositions of students in this course included pre-medical, nursing, and health science majors at the junior or senior college level. The sample selected was a course delivered in Spring 2006. Therefore, the data was archived as part of a normal educational experience, and no demographic information was collected from the students. Participants in this course were required to make at least one text-based entry and one voice-based entry for each course topic. The course selected for this study contained seven discussion topics, which provided a total of 4527 computer conference entries. Use of voice and text was required for the duration of the course.
cause the use of text tools did not influence the use of voice tools in this environment, students were not required to use the tools in a particular order but to use both tools to participate in each discussion topic. Therefore, the asynchronous computer conference transcripts generated in this course were consistent with the characteristics identified in the literature to represent an adequate sample for this study.

3.1. Variables

The independent variable identified for this study was computer conferencing input/output, with two levels: voice-based input/output and text-based input/output. The three dependent variables associated with the levels of social presence interaction in asynchronous computer conferencing were identified:

1. Affective (APD) presence density [2]
2. Communicative reinforcement (IPD) presence density [2]
3. Cohesive (CPD) presence density [2]

Scores for each dependent variable were presented as ratio data. Each score represented the number of times the indicator was present per 1000 words [2].

3.2. Coding Procedures

For the purpose of this investigation, all text and voice messages were transcribed into a text format and saved in chronological order. Audio files were also saved in chronological order on a laptop computer to allow coders to hear as well as read transcripts of the audio conferences. A hard copy transcript of all text and audio files was provided to the coders.

Coding instructions defined rules regarding the maximum amount of time allowed for each coding session. Coders read the text file transcripts and underlined the instance of social presence indicators on the hard copy transcript and indicated the category by noting the letter “A” for affective indicators, the letter “I” for communicative reinforcement indicators, and the letter “C” for cohesive indicators above the underlined text. Coders also listened to the audio file while reading the text transcription of the audio file to ensure all social presence indicators were identified and marked the instances of social presence on the hard copy transcription with the same notations used for the text transcriptions.

Coders worked independently but came together for the purpose of negotiated coding sessions via telephone and in person throughout the coding procedures. Coders met in person for negotiated sessions once a week during the first two weeks of coding. During the in person sessions, coders would discuss any questions regarding categorization of social presence indicators. When coders disagreed on a category the coding decisions were discussed and if an agreement was reached the coders had the opportunity to change their coding decisions [27]. If the coders could not agree on the appropriate category no changes were made by the coders. It should be noted that the coders for this study were not involved with the research outside their coding responsibilities. The coding process required a significant amount of time and was completed over a lengthy four month process. In addition, social presence indicators were easily identifiable due to the technical nature of the discussion entries. The negotiated coding process over such a lengthy period of time enhanced coder agreement. During the last six weeks of coding procedures the coders discussed coding decisions via the telephone weekly. The process of negotiating coding decisions facilitated coder agreement and inter-rater reliability.

3.3. Level of Measurement

Because individual messages in both voice and text conference transcripts can differ in length as defined by the number of words each message contains, a simple count of the number of social presence indicators present in a message could not accurately represent the data. For example, one message may have contained 500 words with 15 indicators of social presence, while

<table>
<thead>
<tr>
<th>Categories</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective Responses (A)</td>
<td>• Expression of emotions</td>
</tr>
<tr>
<td></td>
<td>• Use of humor</td>
</tr>
<tr>
<td></td>
<td>• Self-disclosure</td>
</tr>
<tr>
<td>Communicative Reinforcement</td>
<td>• Continuing a thread</td>
</tr>
<tr>
<td>Interactive Responses (I)</td>
<td>• Quoting from others’ messages</td>
</tr>
<tr>
<td></td>
<td>• Referring explicitly to others’ messages</td>
</tr>
<tr>
<td></td>
<td>• Asking questions</td>
</tr>
<tr>
<td></td>
<td>• Complimenting, expressing appreciation, expressing agreement</td>
</tr>
<tr>
<td>Cohesive Responses (C)</td>
<td>• Phatics, salutations</td>
</tr>
<tr>
<td></td>
<td>• Addresses or refers to the group using inclusive pronouns</td>
</tr>
</tbody>
</table>

Table 1. Social Presence Categories & Indicators [2]
another message may have contained 800 words with only 12 indicators of social presence. A more effective method was to calculate a score for each message that could be compared to the scores of other messages. This calculation was produced by counting the number of times an indicator was identified in each unit of analysis, dividing that number by the total number of words in each unit of analysis, then multiply by 1000. The number of times an indicator occurred per 1000 words was the social presence density score (SPD) \[2\]. A per 1000 word score made it possible to compare the level of social presence for each unit of analysis. A social presence density calculation was performed for each dependent variable providing a density per 1000 words for the overall social presence of each thread, as well as the density per 1000 words for the affective interaction (ADS), communicative reinforcement interaction (IDS), and cohesive interaction (CDS) in both text and voice. A total social presence density score and a density score for each category enabled the investigator to make comparisons from the voice and text transcripts for each dependent variable.

### 3.4. Inter-rater Reliability

The primary test of reliability in quantitative content analysis studies is inter-rater reliability. For the purpose of this investigation, the Holsti percentage of agreement statistic was calculated to determine the inter-rater reliability measurement for this investigation. During five coder training sessions, if reliability scores were not at least .80, coders were asked to review the items on which there was disagreement and discuss those items until a consensus based on the coding protocol was accomplished \[27\]. The Holsti’s percentage of agreement was .74 on the first coding sample, .82 on the second coding sample, .86 on the third coding sample, and .91 on the 2 final coding samples. Based upon the Holsti’s percentage of agreement of .91 on the final 2 coding samples, it was concluded that coder training was successful and coders could begin coding the data.

### 3.5. Data Analysis

Data analysis for this study used a multivariate analysis of variance (MANOVA) to compare differences within subjects and between groups. Because each participant in this study was exposed to both levels of the independent variable, measuring the differences within subjects was used to determine which portion of the within group variance was attributable to pre-existing individual differences or preferences. This portion could then be dropped from the analysis before comparisons were measured between groups. The Statistical Package for the Social Sciences (SPSS) was used to perform the MANOVA.

The MANOVA also tested if the variability between the voice-based and text-based computer conferences is greater than the variability that could be attributed to random error. If the hypotheses fail rejection, there are no differences in groups, and any difference between groups is the result of random variation. However, if the hypotheses can be rejected, a difference between the levels of social presence as measured by the categories of interaction in voice-based and text-based computer conference groups exists \[30\]. Descriptive statistical tests to identify the mean and standard deviation for both levels of the independent variable were also performed.

### 4. Results

A total of 4527 entries were generated during the 15-week semester course. Of those transcripts, 269 were removed due to irresolvable coder disagreement, and 20 additional transcripts with no social presence score were also discarded. An a priori significance level of .01 was set for all statistical analysis in this study due to the large number of asynchronous computer conference messages examined.

#### 4.1. Descriptive Statistics

There were 2812 (66.35%) text-based computer conference entries and 1425 (33.65%) voice-based entries (n=1425) among the transcripts examined.

<table>
<thead>
<tr>
<th>Table 2. Descriptive Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text/Voice</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td><strong>Affective (A)</strong></td>
</tr>
<tr>
<td>Text</td>
</tr>
<tr>
<td>Voice</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td><strong>Communicative Reinforcement (I)</strong></td>
</tr>
<tr>
<td>Text</td>
</tr>
<tr>
<td>Voice</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td><strong>Cohesive (C)</strong></td>
</tr>
<tr>
<td>Text</td>
</tr>
<tr>
<td>Voice</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td><strong>Total Social Presence</strong></td>
</tr>
<tr>
<td>Text</td>
</tr>
<tr>
<td>Voice</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Table 2 lists the means and standard deviations for each dependent variable. The total social presence score indicates the social presence density score for the sum of affective, communicative reinforcement, and cohesive density scores. Mean scores are higher for the text group across the affective and communicative reinforcement categories. The cohesive dependent variable reveals an opposite result with the mean score for the voice group being higher than the mean score for the text group.
4.2. Multivariate Analysis

MANOVA results and analyses are presented in two phases. First, the global effect between the two levels of the independent variable is tested for significance. The most common global assessment index is the Wilks’ Lambda analysis. Wilks’ Lambda shows the amount of variance in the dependent variables that cannot be explained by the independent variable [30]. Multivariate tests of differences between the two levels of the independent variable revealed the Wilks’ Lambda to be statistically significant (F=13.653; p<.000), indicating the differences between the groups are related to the independent variables and do not occur by chance alone (Table 3). Hotelling’s Trace is also commonly used as an indicator of global significance when two dependent groups are examined (30, 31). For the purpose of this investigation, the voice and text groups are considered dependent since all individuals participated in both text and voice asynchronous computer conferences throughout the semester. Therefore, the statistical significance of Hotelling’s Trace (F=13.653; p<.000) was also considered in this analysis. Pillai’s Trace (F=13.653; p<.000) and Roy’s Largest Root (F=13.653; p<.000) also produced statistically significant results, lending additional support to the significance of the Wilks’ Lambda and Hotelling’s Trace.

Table 3. Multivariate Tests

<table>
<thead>
<tr>
<th>Effect</th>
<th>F</th>
<th>df</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.02</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Wilks’ Lambda</td>
<td>1.00</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Hotelling’s Trace</td>
<td>1.02</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Roy’s Largest Root</td>
<td>1.02</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Test_Report</td>
<td>1.00</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Wilks’ Lambda</td>
<td>1.00</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Hotelling’s Trace</td>
<td>1.02</td>
<td>3.00</td>
<td>.999</td>
</tr>
<tr>
<td>Roy’s Largest Root</td>
<td>1.02</td>
<td>3.00</td>
<td>.999</td>
</tr>
</tbody>
</table>

Since the Wilks’ Lambda and Hotelling’s Trace revealed differences in the means as statistically significant, the second phase of MANOVA analysis interpretation for each related hypothesis was performed [29, 30]. Comparisons between the two levels of the independent variable for each of the social presence indicators were examined. Results indicated a statistical significance at the .01 level for the affective (p<.000) and communicative reinforcement (p<.000) dependent variables. Additionally, the total for all categories indicates statistical significance at the .01 level (p<.000). The cohesive dependent variable is not statistically significant at the .01 level (p=.014). Although the cohesive dependent variable did not indicate statistical significance at the .01 level, it should be noted that the value would be significant at the commonly accepted level of p<.05. The a priori significance level of .01 was set for this study because of the large number of transcripts included in the analysis. While the statistical results reported in this study indicate which dependent variables are statistically significant, these statistics alone require further analysis and interpretation to determine if the null hypotheses tested in this study can or cannot be rejected. Each research hypothesis will be addressed separately in the following sections.

4.3.1. Hypothesis 1

There is no difference in the affective interactions of voice-based asynchronous computer conferencing when compared to affective interactions in text-based asynchronous computer conferencing. To determine if the first null hypothesis can be rejected or not requires examination of the statistical significance of the MANOVA, as well as examination of the variance of the means for the affective dependent variable. MANOVA analysis indicated the text and voice groups were different at a significance level of .01 (Table 4). The MANOVA significance level for affective social presence indicators was p<.000, indicating a difference in the affective social presence levels in the text and voice asynchronous computer conference transcripts. The mean value was higher for the text group (23.215) than the voice group (20.710). Therefore, the null hypothesis can be rejected. Estimated marginal mean scores for the affective dependent variable are presented in Table 5. The affective indicators of social presence occur more often in text-based asynchronous computer conference entries and at a statistically significant level (p<.000).

Table 4. Dependent Variable Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type I Sums of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Affective (A)</td>
<td>1220.745</td>
<td>1</td>
<td>1220.745</td>
<td>.974</td>
</tr>
<tr>
<td>Wilks’ Lambda</td>
<td>1560.619</td>
<td>1</td>
<td>1560.619</td>
<td>.058</td>
<td></td>
</tr>
<tr>
<td>Hotelling’s Trace</td>
<td>1618.939</td>
<td>1</td>
<td>1618.939</td>
<td>.058</td>
<td></td>
</tr>
<tr>
<td>Total Social Presence</td>
<td>.850</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>Communicative Reinforcement (B)</td>
<td>0.000</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks’ Lambda</td>
<td>0.000</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hotelling’s Trace</td>
<td>0.000</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Social Presence</td>
<td>.000</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>Coherence (C)</td>
<td>0.000</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks’ Lambda</td>
<td>0.000</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hotelling’s Trace</td>
<td>0.000</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Social Presence</td>
<td>.000</td>
<td>.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Affective Dependent Variable

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Voice</th>
<th>Mean</th>
<th>SD</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective</td>
<td>23.215</td>
<td>540</td>
<td>19.303</td>
<td>22.116</td>
<td></td>
</tr>
</tbody>
</table>
4.3.2. Hypothesis 2

There is no difference in the communicative reinforcement interactions of voice-based asynchronous computer conferencing when compared to communicative reinforcement interactions in text-based asynchronous computer conferencing. Analysis indicated the text and voice groups were different at a significance level of .01 (Table 4). The MANOVA significance level for communicative reinforcement social presence indicators was $p<.000$, indicating a difference in the communicative reinforcement social presence levels in the text and voice asynchronous computer conference transcripts. The mean value was higher for the text group (23.806) than the voice group (20.214). Therefore the null hypothesis can be rejected. Estimated marginal mean scores for the communicative reinforcement dependent variable are presented in Table 6. The communicative reinforcement indicators of social presence occur more often in text-based asynchronous computer conference entries and at a statistically significant level ($p<.000$).

Table 6. Communicative Reinforcement Dependent Variable

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Text/Voice</th>
<th>Mean</th>
<th>St. Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicative</td>
<td>Reinforcement</td>
<td>Text</td>
<td>23.806</td>
<td>22.821 24.791</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Voice</td>
<td>20.214</td>
<td>18.821 21.597</td>
</tr>
</tbody>
</table>

4.3.3. Hypothesis 3

There is no difference in the cohesive interactions of voice-based asynchronous computer conferencing when compared to cohesive interactions in text-based asynchronous computer conferencing. Analysis indicated the text and voice groups were not significantly different at the .01 level (Table 4). The MANOVA significance level for cohesive social presence indicators was $p<.014$, indicating no statistically significant difference in the cohesive social presence levels between text and voice asynchronous computer conference transcripts. The mean value was higher for the voice group (7.175) than the text group (5.890). Therefore, the null hypothesis cannot be rejected. Estimated marginal mean scores for the interactive dependent variable are presented in Table 7. Cohesive indicators of social presence occurred more often in voice-based asynchronous computer conference entries, but not at a statistically significant level ($p<.014$).

Table 7. Cohesive Dependent Variable

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Text/Voice</th>
<th>Mean</th>
<th>St. Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohesive</td>
<td>Text</td>
<td>5.890</td>
<td>.302</td>
<td>5.512 6.660</td>
</tr>
<tr>
<td></td>
<td>Voice</td>
<td>7.175</td>
<td>.424</td>
<td>6.502 7.850</td>
</tr>
</tbody>
</table>

5. Conclusions and Recommendations

5.1. Strengths of the Study

Several items emerged during the course of this research project that improved the strength of the study. Predominant strengths of this study include the large sample of transcripts examined, the use of the negotiated coding model, and the use of previously established categories and indicators of social presence.

First, the large number of asynchronous computer conferences transcripts (n=4238) available for examination provided data in quantities sufficient to ensure a rigorous statistical analysis. Such a large set of data allowed the statistical significance factor to be set at a level of .01, thus reducing the possibility that the researcher would make a Type I error.

Second, the negotiated coding model utilized [26] increased the agreement of the coders in a substantial way. Only 269 message transcripts were discarded from the total number of transcripts due to coder disagreement. While coders performed the coding assignments independently, they had the opportunity to discuss coding decisions and make changes if the discussion resulted in one of the coders changing a coding decision. The process of the negotiation also provided reinforcement regarding coding decisions. Therefore, as the coding progressed, the number of negotiated coding decisions was reduced. In addition, the content of the asynchronous computer conference entries was related to human physiology and consistently contained a high degree of technical language related to course content. Therefore, when there was evidence of social presence in the message, coders were able to identify the indicator easily. This further supported the percent of coder agreement and inter-rater reliability.

Third, the use of previously established categories and coding indicators strengthened the reliability and validity of this study. Building on the foundation of existing empirical evidence [2] permits comparison with other studies that have used the same system. Consequently, the project is more likely to contribute to theory and knowledge in the discipline.

5.2. Limitations of the Study

To demonstrate the implications of this study, the limitations of the study must be examined. A significant limitation of this study was the procedures employed for use of text-based and voice-based asynchronous computer conferences in the human physiology course. The course instructor set the course policies and guidelines for the asynchronous communication in this course. This investigator had no control over the management or style with which the course instructor
facilitated asynchronous communication among students. The asynchronous conferences in this course were used as discussion forums to answer quiz questions related to course topics. In addition, the large volumes of messages generated in asynchronous computer conferences may inhibit effective discussion development and produce a cognitive overload for participants [3]. A different asynchronous pedagogy designed to encourage community building through social presence in asynchronous communications and creating small groups for discussion within the course may have produced different results. Therefore, the generalizability to other text-based and voice-based asynchronous computer conferences must be viewed with caution and generalized only to like groups in courses with similar pedagogy.

Another significant limitation of this study was the a priori establishment of the .01 significance level [30, 31]. The cohesive dependent variable was reported as no significant difference for this study at the .01. Therefore, the cohesive indicators of social presence in voice-based and text-based asynchronous computer conferences warrant further examination.

Finally, the fundamental assumption of this study was that the tools used to identify categories and indicators in text-based conferences were also valid in voice-based conferences. Social presence categories and indicators are well-established in the literature [2] with regard to text-based asynchronous computer conferences. However, there was no evidence in the literature that these categories and indicators have been applied to voice-based asynchronous computer conference content.

5.3. Conclusions

This study builds on prior research [2] into the social presence categories and indicators used to compare text-based and voice-based asynchronous computer conference transcripts. The use of standard coding categories permits comparison with other studies that have used the same system. Therefore, the project is more likely to contribute to theory and knowledge in the discipline. Although this study compares levels of social presence between text and voice conference entries, the procedures and strategies used in the study are grounded in theory and practice evidenced in the review of the literature. Results revealed that social presence levels are higher in text-based computer conferences when compared to voice-based conferences for the affective and communicative reinforcement categories. Results regarding cohesive indicators of social presence were higher in voice-based conferences, but not a statistically significant level and require further examination.

There is a paucity of evidence in the literature with regard to the levels of social presence in voice-based asynchronous communication [15, 18, 29, 32]. This study indicates that the use of voice as a delivery medium of asynchronous communication in Web-based courses does not produce the desired result of an increased sense of social presence among the participants. Therefore, faculty considering the use of voice in asynchronous computer conferences should consider the lower levels of social presence in voice-based conferences when compared to text-based conferences. A review of the literature indicated that social presence in asynchronous computer conferences is a critical factor related to positive learning outcomes achieved in such a discussion-based pedagogy [3, 4]. This study further indicates that when students had the option to generate either voice or text messages in an asynchronous environment, students elected to use text to enter computer conference messages at a ratio of 2:1.

This study did suggest that one of the social presence indicators might be higher in the voice-based conferences than text-based conferences. Further research regarding the comparison of cohesive levels of social presence in text and voice conferences would be beneficial. Cohesive indicators are related to level of group commitment from the participants. These indicators function to strengthen the social order of a community and include informal expressions, such as inquiries regarding personal matters and remarks that indicate inclusion, involvement, and association [2]. Additional research to determine if the levels of cohesive indicators observed in this study occur in other similar populations would provide additional contribution to the knowledge and professional practice of Web-based instruction. Further implications of this study are related to the quality of conference interactions and the relationship of the interactive quality to learning outcomes. In addition, implications regarding the levels of social presence in computer conferences warrants further investigation. Specific recommendations are presented in the following section.

5.4. Recommendations for Future Research

According to the findings presented in this study, there is no apparent benefit in using voice-based tools to facilitate asynchronous computer conferences as an instructional technology. The following recommendations are based on the results of this study.

1. Faculty who are considering the implementation of technology tools to facilitate asynchronous computer conferences in Web-based courses should examine the evidence in the literature to ensure tools selected for use during course delivery are appropriate for the desired learning outcomes.
2. Although the conclusion of this study indicates that voice-based asynchronous communication is not equal or superior to the use of text-based asynchronous communication, it should be noted that the cohesive indicators of social presence were higher in the voice group than the text group. Further investigation of this dependent variable would provide additional insight into the use of voice as a medium for asynchronous facilitation of group communication.

3. The technical nature of the subject matter in the course examined for this study enabled easy identification of social presence interaction patterns. Replication of this study using a course with content more conducive to social interaction may produce different results.

4. The pedagogy used for the course examined in this study was determined by the course instructor. The asynchronous computer conferences represented a graded quiz for each topic. Replication of this study using a course in which the pedagogy was designed to facilitate social presence interaction patterns may produce different results. Asynchronous computer conference pedagogies in which students develop discussions throughout the semester that are related to course content, but not graded quizzes may exhibit different social presence interaction patterns than those produced in this study.

5. Students participating in the asynchronous computer conferences in this study were enrolled in a Web-based course with no face-to-face interaction. Replication of this study in a blended course may produce different results. Students who have the opportunity to meet in face-to-face environments at least a few times during an academic semester may communicate differently in asynchronous computer conferences.

6. This study also indicated that when students were given a choice between using voice and text to generate asynchronous computer conference content, they most often selected text. Even though student preferences were not part of this study, the number of text versus voice messages generated throughout the semester indicate the predominant use of text to generate asynchronous communication. Replication of this study in a course where students are given the choice of the asynchronous communication tool may produce different results. In addition, student perceptions of both the voice and text tools would provide further information regarding the appropriate use of voice-based technologies as an asynchronous communication tool.

7. An important recommendation for future research relates to the measurement of the quality of interactions beyond the presence or absence of specific indicators. There is a theoretical basis in the literature indicating high levels of social presence have a positive effect on learning. However, it was beyond the scope of this study to make statements regarding the quality of asynchronous computer conference interactions. Therefore, further examination of interaction patterns with regard to the quality of the interaction is required to produce new evidence connecting the quality of the interaction with the quality of learning outcomes produced in asynchronous computer conferences.

8. Although the monetary aspects of implementing voice-based technologies to support online learning environments was beyond the scope of this study, a cost-benefit analysis regarding the use of audio versus text is an important recommendation for future research. Consideration of the financial resources required for an enterprise integration of voice-enabled asynchronous discussion forums would enhance the findings of this study. Finally, levels of social presence also relates to the quality of the asynchronous computer conference interaction. An important question for future research includes the examination of social presence levels and what high or low levels of social presence actually imply regarding the quality of interactions.

6. References


