

Latent Semantic Similarity in Initial Computer-Mediated Interactions

Vivian P. Ta

University of Texas at Arlington

Abstract

The initial interactions of same-sex strangers were investigated to examine the development of latent semantic similarity (LSS; that is, how interaction partners come to use words in the same way) in computer-mediated interactions. A previous study by Ta, Babcock, and Ickes (2016) found that verbal behaviors, rather than nonverbal behaviors, predicted the development of LSS in the initial face-to-face interactions of same-sex strangers. This suggested that LSS might develop similarly and as efficiently in conversations in which only words are exchanged (i.e., computer-mediated interactions). As such, using a sample of 120 same-sex dyads (all strangers) who interacted with each other using AOL Instant Messenger for 18 minutes, it was hypothesized that (1) dyad-level behaviors that introduce more words into the conversation will be essential and unique predictors of dyad-level LSS; (2) higher dyad-level LSS would predict higher dyad-level perceived interaction quality; (3) higher dyad-levels of agreeableness, extraversion, and openness to experience would predict higher dyad-level LSS; (4) dyad-level LSS would mediate the relationship between dyad-level personality measures of agreeableness, extraversion, and openness to experience and dyad-level perceived interaction quality; and (5) dyad-level LSS would increase over time. The results revealed that (1) higher number of messages sent, but not the number of questions asked, significantly predicted higher LSS; (2) LSS did not significantly predict perceived interaction quality, but did predict certain factors of perceived interaction quality; (3) agreeableness, extraversion, and openness to experience did not significantly predict LSS; (4) LSS was not a significant mediator of the relationship between these 3 personality measures and perceived interaction quality; and (5) LSS decreases, rather than increases, over time, which may be attributed to a compensation effect, or to differences in dyad-level extraversion.

How much of our daily communication consists of computer-mediated communication? As an aggregated estimate, researchers have reported that 182.9 billion emails were sent and received *per day* in 2013 (Radicadi, 2013). The same researchers predict that this number will increase to 206.6 billion emails sent and received per day by 2017. Given numbers as staggering as those, it should not come as a surprise that people all over the planet now use computer-mediated communication as their primary form of communication. Computer-mediated communication has become an increasingly preferred method of communication since the internet was made available to the general public in the early 1990s.

Today it is quite common for friends, family members, classmates, and co-workers to communicate via computer-mediated communication (e.g., through email, instant messaging, texting) on a daily basis using various desktop, laptop, and handheld devices. Research has shown that a growing number of people today prefer to do most of their communication this way rather than in face-to-face interactions, and this trend is expected to continue as the supporting technology progresses and the time required to sustain face-to-face interactions becomes less available (Keller, 2013; Onishi & Steger, 2012; Tardanico, 2012; Thayer & Ray, 2006).

A significant (though as-yet-unestimated) percentage of computer-mediated communication takes place between pairs of strangers. Based on the statistics reported above, it is safe to presume that hundreds of millions of computer-mediated communications are exchanged every day between people who have never interacted with each other before. This situation is common on dating websites, such as Match.com and eHarmony, which allow individuals to send and receive messages from new romantic prospects. In addition, this situation is common in online forums, such as Reddit or Quora, in which users can easily take part in conversations or answer questions that were initiated by other users. However, computer-

mediated communication between strangers is not limited only to dating websites or online forums: Google announced in 2014 that it will start allowing Gmail users to contact anyone on Google+ (which is Google's social networking website) without having to know their email address. Considering Gmail's 425 million users, Google Plus's 300 million monthly users, and the recent push for a more technologically connected society, computer-mediated communication between strangers is likely to become far easier and far more commonplace than ever.

Computer-mediated communication differs from face-to-face communication in several important ways. For example, it provides little or no information about the many nonverbal indicators that are available in face-to-face communication such as eye contact, gestures, body orientation, etc. In itself, it also provides little or no information about other important characteristics that people use to create inferences about their interaction partners, such as the partner's physical attractiveness and tone of voice. With all of these missing components, is computer-mediated communication still sufficient to enable pairs of strangers to understand each other?

To assess how—and how well—interaction partners come to understand each other through computer-mediated communication, we must first consider how interaction partners manage to do this in a face-to-face interaction. Many writers have argued that interaction partners rely heavily on language to develop a “common-ground understanding” (e.g., Abbeduto, Short-Meyerson, Benson, Dolish, & Weissman, 1998; Kecskes & Zhang, 2009; Krauss & Fussell, 1991; Schober & Clark, 1989; Wilkes-Gibbs & Clark, 1992) or an “intersubjective meaning context” (Gesn & Ickes, 1999; Morganti, 2008). In other words, the development of a “common-ground” understanding depends on the interaction partners first getting on the same

page in a linguistic sense, so that they can use same words in essentially the same way (i.e., to have the same intended meanings). But how are people able to do this?

How interaction partners manage to do this might be less of a mystery than it first appears, as suggested by Babcock, Ta, and Ickes (2014, p. 2)

Even a casual observation of everyday social interaction suggests that this process depends on (1) *how much the interaction partners talk to each other* (a behavior that is essential for the sampling and eventual mutual alignment of what words are used and how they are used in relation to other words); (2) *how much the interaction partners look at each other* (a behavior that is essential for the detection of those nonverbal cues—including the emotional ones—that help the partners achieve a more qualified and nuanced understanding of what their partner’s words mean in context); and (3) *how much the interaction partners acknowledge each other, both verbally and nonverbally* (behaviors that often, though not always, signal that one’s partner feels that he or she has grasped one’s intended meaning and wishes to express that perceived understanding, see Abbeduto et al., 1998).

This perspective suggests that the talking, looking, and acknowledgments (both verbal and nonverbal) that occur between previously-unacquainted dyad members are essential behaviors for the development of a “common-ground” understanding or “intersubjective meaning context” (Babcock et al., 2014). However, the ability of researchers to test these claims empirically had to wait for the development of a measure of the implied outcome variable —“the extent to which two people use the same words in essentially the same way.” Fortunately, such a measure was developed by Landauer and his colleagues and is now available in the form of the *latent semantic similarity* (LSS) index.

Latent Semantic Similarity (LSS)

Latent semantic similarity is assessed using Latent Semantic Analysis (LSA), an automated statistical method that establishes the contextual meaning of any text by analyzing the relationship among the words that are used (Landauer & Dumais, 1997; Landauer et al., 1998). Among other applications, LSA can be used to compare two blocks of text and assess the degree of their *latent semantic similarity* (LSS), i.e., the degree to which the same words are used in the same way in both blocks of text. Depending on whether or not a word does or does not appear within a corpus in various contexts, the LSA program combines the contexts in which a word is used and then implements a set of rules that determines the similarity of meanings of words and groups of words. After selecting one of the options in Landauer's online software program, the LSA Pairwise Comparison program (Laham et al., 1998, <http://lsa.colorado.edu>), the user is prompted to input two blocks of texts into the program's high-dimensional semantic space in which computations can be made using up to 150 dimensions. The program then computes the cosine of the angles between the two resulting vectors to estimate the overall degree of latent semantic similarity between the two blocks of text based on the words that are used and how those words are used in relation to other words. If, within a given corpus, two words that are similar in meaning and context occur within the same text, LSA identifies them as similar even if the words rarely occur within the same passage or sentence.

It is important to characterize the LSS index as a measure of *latent semantic similarity* (i.e., the extent of shared meaning) rather than a measure of *lexical similarity* (i.e., the extent to which the same exact words are used, which could be computed simply as a ratio of the words used by both partners divided by the total words used). An example of how the LSA program identifies words that are similar in meaning and/or context is provided by Arnulf, Larsen,

Martinsen, and Bong (2014): A LSS index of .80 was generated when the LSA program performed a pairwise comparison between the sentences “Doctors operate on patients” and “Physicians do surgery”. The LSA program recognized a shared meaning between the two sentences even though these two sentences do not have any words in common. Programs that assess only lexical similarity do not take into account the specific meanings of the words being used (Arnulf et al. 2014), which can be quite troublesome when a word has multiple meanings (i.e., homographs) and is used in several different contexts. Hence, assessing latent semantic similarity, rather than lexical similarity, is more appropriate for analyzing the conversation that occurs in unstructured dyadic interactions.

The LSA program itself differs from other text-analysis approaches in a number of important ways. First, LSS does not only use the summed contiguous pairwise-occurrences of words in its initial data; it also uses the detailed patterns of occurrences of words over the entire available set of local meaning-bearing contexts (i.e., contexts that are treated as unitary wholes, such as sentences or paragraphs). Second, LSA ascribes high significance to dimensionality (i.e., to a simultaneous representation all of the local word-context relations), such that reducing dimensionality¹ of the observed data from the number of initial contexts to a smaller (but still large) number will generally produce results that are more similar to human cognitive relations. Third, LSS takes into account the overall distribution of words over usage contexts that is evaluated apart from their correlations. Fourth, the computational procedure by which the LSS index is generated depends not only on how often similar words are used or on the contingencies of the occurrence of words in a given corpus, but also on a statistical analysis that is able to accurately interpret the associations that go beyond first order co-occurrences (e.g., the context in which a given word is used). This computational procedure therefore produces broad and

¹ The parameters by which a word or passage is described.

inclusive measures of semantic similarity that have been shown to relate in predictable ways to a number of human cognitive phenomena involving association or semantic similarity (Landauer & Dumais, 1996; Landauer et al., 1997). These aspects enhance the degree to which LSS captures “shared meaning” over an entire interaction rather than in more isolated parts of it (Landauer et al., 1998).

Previous research using this index was primarily limited to comparing the LSS between a “target” text and a single, standard “criterion” text (for examples, see Caspar, Berger, & Hautle, 2004; Landauer, Laham, & Foltz, 2003; Lautenschlager, Dunn, Bonney, Flicker, & Almeida, 2006). More recently, interest in the LSS construct has extended to comparing the similarity between pairs of texts that appear on the Internet, for example, the similarity between the words that two friends use, or the similarity between the words that advertisers plan to use and the words used by their target audience. (For more specific examples, see Katsanos, Tselios, & Avouris, 2008; Mihalcea, Corley, & Strapparava, 2006; Recchia, & Jones, 2009).

Babcock et al. (2014) applied the LSS construct to the study of the conversations that occur in initial dyadic interactions. Their goal was to correlate the LSS index score for each dyad with various dyad-level behaviors, such as the frequency and duration of mutual gazes, the number of verbal acknowledgments, and the number of conversation sequences initiated, and then see if the resulting pattern of correlations would provide insights into how interaction partners come to understand each other. They found that the level of LSS within each dyad was higher when the partners (1) exchanged a lot of verbal information during their conversation; (2) displayed more individual and mutual gazing; and (3) displayed more verbal and nonverbal acknowledgements.

However, the results of a follow-up study by Ta, Babcock, and Ickes (2015) revealed a rather different picture. In this study, the authors sought to (1) replicate the results from Babcock et al. (2014), (2) extend these previous findings by examining other potential correlates of LSS index, and (3) determine which of these behaviors, when represented as relatively independent factors, make unique and independent contributions to each dyad's LSS index. In addition to successfully replicating the results from Babcock et al. (2014), this study found a few additional correlates of LSS (see Appendix A). At the level of zero-order correlations, the results suggested that the dyad members achieved higher levels of LSS when they exchanged more words with each other and elicited additional information by asking a large number of questions. Also, the dyad members achieved higher levels of LSS when they used more gestures to qualify and/or emphasize the meaning of what they said, and when they smiled and laughed more to signal moments of emotional rapport. Lastly, higher levels of LSS were also related with more positive thoughts and feelings, but with fewer thoughts and feelings about third-party others, which may have reflected greater enjoyment of the interaction and a focus on each other rather than on other people.

In order to determine if these behaviors, when represented as relatively independent factors, make unique and independent contributions to each dyad's LSS index, a factor analysis was conducted on the 13 dyad-level behavioral measures. This analysis resulted in four relatively independent factors: "Looking and Acknowledging," "Gesturing," "Talking and Asking Questions," and "Smiling and Laughing" (See Appendix B for the behavioral measures that make up these factors). The composite score for each of the four factors, along with a dummy-coded variable that represented the dyad's gender composition (MM and FF), were then entered into a multiple regression model to predict LSS. The results revealed that the factor representing

“Talking and Asking Questions” behaviors was the only unique predictor of LSS. This finding suggests that the development of LSS between first-time interaction partners depends primarily on the exchange of words, and that all other dyad-level behaviors are nonessential predictors when the effect of the Talking and Asking Questions factor is statistically controlled. Although these other behaviors (e.g., looking and acknowledging, gesturing, smiling and laughing) may play an essential role in the nonverbal exchange that occurs in dyadic interactions (Bruder, Doslukhambetova, Nerb, & Manstead, 2012; Manusov & Patterson, 2006; Patterson, 2014), they do not appear to play an essential role in helping interaction partners achieve high LSS, or, in other words, “to get on the same page” linguistically.

The major conclusion suggested by the Ta et al. (2015) results—that “the words may be all you need” for the development of LSS in the initial interactions of strangers—further suggests that LSS might develop as readily in computer-mediated communication as it does in face-to-face interactions. (Ta et al., 2015). Because an increasing amount of initial interactions now occur online in the form of computer-mediated communication, this implication warrants an empirical investigation of its own.

The Present Study

The goal of the present study was to examine the development of latent semantic similarity in written online interactions between first-time interaction partners. In addition to determining whether the amount of conversation underlies the development of LSS in online interactions, the current study also examined whether dyad-level personality measures could be used to predict the level of LSS that develops between first-time interaction partners. Finally, the current study tracked the development of LSS across different stages of the dyads’ online

conversations to explore the issue of how LSS changes over time. The following five hypotheses were tested:

Hypothesis 1. Dyad-level behaviors that introduce more words into the conversation (i.e., the number of questions asked and number of messages sent) will be essential and unique predictors of the dyads' LSS. Ta et al. (2015) found that in initial face-to-face interactions, only talking-related behaviors appear to play an essential role in the development of LSS. Hypothesis 1 proposes that this finding will also apply to first-time conversation partners who are interacting through computer-mediated communication. Indeed, this finding is virtually guaranteed based almost entirely on the words the partners use in their online computer-mediated communication interaction.

Hypothesis 2. The dyad-level LSS indices will significantly predict a global measure of perceived interaction quality across the set of dyads. That is, there should be a positive association between the dyads' LSS index score and a general factor derived from the dyad-level measures of (1) how smooth, natural, and relaxed the interaction was; (2) how much the interaction partners felt accepted and respected by each other; (3) how much of a connection the interaction partners felt with each other; (4) how much the interaction partners liked each other; (5) how much the interaction partners would like to interact with each other in the future; (6) how much the interaction partners felt they understood each other; (7) how much the interaction partners enjoyed interacting with each other; and (8) how comfortable interaction the partners felt when interacting with each other. In other words, being able to get on the same page linguistically and establish a "common-ground understanding" should elicit a positive evaluation from dyad members regarding the perceived quality of their interaction experience.

Conversely, there should be a significant negative association between the dyads' LSS index scores and the partners' perception of (1) how awkward, forced, and strained the interaction was; and (2) how much the interaction partners felt put down, patronized, or rejected by each other. In other words, not being able to get on the same page linguistically and establish a "common-ground understanding" should elicit a negative evaluation of their interaction experience from the dyad members.

Hypothesis 3. The dyad-level personality measures of agreeableness, extraversion, and openness to experience should significantly predict dyad-level LSS. The rationale for these predictions is as follows. A high level of extraversion, which is characterized by the tendency to seek stimulation in the company of others (Matthews, Deary, & Whiteman, 2003), should facilitate behaviors that add more words into the conversation, a crucial aspect in the development of LSS. Openness to experience, which is characterized by the tendency to be curious rather than cautious (Ashton, 2013), and agreeableness, which is characterized by the tendency to be cooperative rather than antagonistic (Ryckman, 2012), should also facilitate the dyads' development of LSS. Greater openness to experience should lead to a greater exchange of words, whereas greater agreeableness should motivate a more active and responsive conversational style in which attempts to accommodate the partner's manner of speaking play a greater role.

Hypothesis 4. Dyad-level LSS should mediate the relationship between the dyad-level personality measures of agreeableness, extraversion, and openness to experience and the dyads' overall perception of the quality of their interaction. Within the dyadic interactions, high levels of agreeableness, extraversion, and openness to experience should lead to a higher level of LSS, which should in turn lead to a higher overall quality of interaction perception rating.

Hypothesis 5. Dyad-level LSS is expected to increase over time in the initial online conversations of strangers. Each dyad's interaction will be divided into three equal stages of interaction (i.e., the first 6 minutes, middle 6 minutes, and final 6 minutes) and an LSS index score will be generated for each of these three stages. These LSS index scores are expected to increase as the interaction partners pass through the three stages of their initial interaction. As they do so, they should (1) engage more in behaviors that add words to the conversation (i.e., "talk" and asking questions more), and (2) take advantage of accumulating opportunities to sample each other's word choices and align their own word choices and intended meanings to achieve higher LSS with their interaction partner, thus increasing their level of LSS across time.

Method

120 dyads same-sex dyads (43 male-male dyads and 77 female-female) were recruited using the departmental subject pool. Two same-sex participants were randomly assigned to come into the lab at the same time but were seated in different rooms, in separate locations, that were each equipped with a computer (i.e., they did not see each other during their online interaction).

The experimenter began by randomly assigning each participant their instant messenger screen name: one participant was assigned "Participant #1" and the other was assigned "Participant #2". The experimenter then gave a brief outline of the study to each participant individually: (1) this study concerns how interaction partners become acquainted with each other in "get to know you" chat sessions; (2) participants will first complete a survey that includes demographic questions and personality measures; (3) upon their completion of the survey, the experimenter will prompt both participants to start chatting with each other using AOL Instant Messenger for a total of 18 minutes; and (4) after the 18-minute interaction period

is over, the experimenter will prompt both participants to complete the post-interaction questionnaire, which was designed to assess each participant's interaction experience with his or her partner.

Once the experimenter had obtained informed consent from both participants, the experimenter prompted the participants to complete a short pre-interaction survey that included demographic questions and the Big 5 Personality Inventory (see Appendix C). Before the participants began interacting with each other, the experimenter read the "chat procedure" which informed the participants that (1) they would be chatting with their interaction partner for a total of 18 minutes; (2) they should chat with their partner in order to get acquainted with them, just as if they were getting to know someone in real life; and (3) they were free to discuss any topics of their choosing (see Appendix D). The participants then began the online conversation with each other.

Once the 18 minutes of online conversation had passed, the experimenter returned to administer the post-interaction questionnaire which assessed each participant's interaction experience (see Appendix E). Upon the completion of this measure by both participants, the experimenter administered the debriefing statement to each participant separately (see Appendix F). The debriefing statement informed the participants of the true nature of the study and the fact that their chat logs would be saved for data analysis. Written permission to use each participant's chat log was requested at this time. Once their verbal permission was granted, the participants were given an authorization form (see Appendix G) to sign. All participants authorized the use of their chat logs for data analysis.

Latent Semantic Similarity

The speaking turns within each chat log were divided into two electronic text files, each of which contained only one dyad member's portion of the conversation. Each electronic text file was edited down into a solid block of text by deleting all indentions and line breaks, thereby enabling the submission of the text files for both participants as input into the Latent Semantic Analysis program (Laham et al., 1998, <http://lsa.colorado.edu>).

To compute the LSS index for each dyad, the two solid blocks of text that represented each dyad member's portion of the conversation were copied into the input box provided on the LSA website (<http://lsa.colorado.edu>). A blank line was inserted between each block of text to indicate the end of the first block of text and the beginning of the second. The program options: "Pairwise Comparison" (which allows one block of text to be semantically compared to another), "Document-to-Document" (indicating that the two blocks of text are to be semantically compared to each other as documents rather than terms), "Maximum Factors Available" (to determine LSS using the maximum number of dimensions possible), and "General Reading up to 1st Year College" were selected for each computation.

When the two blocks of text for each dyad were submitted for analysis, the LSA program generated an LSS index that ranged between -1 and 1, with a higher positive score indicating greater semantic similarity between the dyad members. In other words, the LSS index represents how similar two writing samples are in the words that are used and in how those words are used in relation to other words (Laham et al., 1997).

To test Hypothesis #5, the same procedure was conducted to generate LSS indices except that the electronic text files that contained each dyad member's portion of the conversation across the entire 18-minute interaction were further divided into three sections of equal time

length (6 minutes each). Specifically, using the timestamps that appeared in all of the AOL Instant Messenger chat logs, each dyad member's electronic text file was divided into three sections based on when the content occurred in the interaction: the first 6 minutes (i.e., minutes 0-6), the middle 6 minutes (i.e., minutes 7-12), and the final 6 minutes (i.e., minutes 13-18). To put it simply, three LSS indices were generated for each dyad that represented their LSS in the first 6 minutes, middle 6 minutes, and last 6 minutes of their initial interaction.

Results

Before conducting the tests of the research hypotheses, a factor analysis was conducted on the post-interaction questionnaire items to determine if the items loaded on one, or more than one, factor. The results of this analysis are reported in Table 1.

Two factors emerged from the analysis. The following items loaded on the first factor, which accounted for 68.49% of the variance: *The interaction seemed smooth, natural, and relaxed to me; The interaction seemed smooth, natural, and relaxed to my partner; The interaction seemed awkward, forced, and strained to me; The interaction seemed awkward, forced, and strained to my partner; I felt accepted and respected by my partner; My partner felt accepted and respected by me; I felt put down, patronized, or rejected by my partner; and My partner felt put down, patronized, or rejected by me.* Viewed collectively, the contents of these items all focused on the perceived interaction quality of each dyad's interaction. For this reason, Factor 1 was labeled *Perceived Interaction Quality*.

The following items loaded on the second factor, which was a much smaller factor that accounted for only 5.75% of the variance: *I felt a connection with my partner; My partner felt a connection with me; I would like to interact more with my partner in the future; My partner would like to interact more with me in the future; I understood my partner; My partner*

understood me; I enjoyed the interaction with my partner; My partner enjoyed the interaction with me; I felt comfortable interacting with my partner; My partner felt comfortable interacting with me; I liked my partner; and My partner liked me. In contrast to the items on Factor 1, which concern the perceived quality of the interaction, the items on Factor 2 concern the degree to which the partners perceived a positive emotional connection to each other. For this reason, Factor 2 was labeled *Emotional Connection with Partner*.

Factor scores for these two factors were then computed. Only Factor 1 was treated as an outcome measure in the analyses for Hypothesis 2 and 4 because these hypotheses specifically examine the influence of LSS on the perceived quality of the dyads' interactions. However, exploratory analyses that examined Hypothesis 2 and 4 with Factor 2 as the outcome measure were also run and are reported in Appendix H.

Tests of the Research Hypotheses

Hypothesis 1

A multiple regression analysis was used to test the hypothesis that the average number of questions asked and the average number of messages sent would be significant predictors of dyad-level LSS, computed over the entire 18-minute interaction period. Replicating the earlier findings of Ta et al. (2015), the present findings indicated that these two predictors together accounted for a significant proportion of the variance in LSS, $R^2 = .06$, $F(2, 118) = 3.98$, $p = .02$. However, the average number of messages sent was the only significant predictor of LSS, $b = .002$, $SE = .001$, $\beta = .28$, $t(118) = 2.70$, $p = .01$, $sr^2 = .06$.

Hypothesis 2

A multiple regression analysis was used to test the hypothesis that dyad-level LSS, computed over the entire interaction period, should be a significant predictor of the global

measure of dyad-level perceived interaction quality. The scores on all of the Factor 1 items in the post-interaction questionnaire were summed and then averaged to create a global measure of perceived interaction quality measure for each individual dyad member. These individual measures were then averaged to obtain the global dyad-level measure of perceived interaction quality.

The results of the regression model used to test Hypothesis 2 were not significant, $F(1, 119) = .94, p = .34, R^2 = .008$. This outcome suggests that LSS does not predict a global measure of perceived interaction quality in the initial, online interactions of same-sex strangers. It is still possible, however, that LSS predicts certain specific facets that contribute to the overall measure of perceived interaction quality. To test this possibility, regression analyses were used to determine if (1) LSS significantly predicted any of the Factor 1 items individually, and (2) if the gender composition of the dyad (M-M or F-F) was a significant moderating variable. The results of these analyses showed that LSS significantly predicted dyad-level perceptions on the following items: *The interaction seemed smooth, natural, and relaxed to me*, $F(3, 116) = 3.85, p = .01, R^2 = .09$), *The interaction seemed smooth, natural, and relaxed to my partner*, $F(3, 116) = 3.73, p = .01, R^2 = .09$.), and *My partner felt accepted and respected by me*, $F(3, 116) = 4.08, p = .01, R^2 = .10$.

For the item *The interaction seemed smooth, natural, and relaxed to me*, LSS was a significant predictor, $b = 1.93, SE = .90, \beta = .20, t(116) = 2.15, p = .03, sr^2 = .03$. The LSS X gender interaction effect was also significant, $b = -2.78, SE = .90, \beta = -.29, t(116) = -3.09, p = .003, sr^2 = .07$. After probing the interaction by gender, the results showed that LSS was a significant predictor of the item *The interaction seemed smooth, natural, and relaxed to me* in

the M-M interactions, $b = 4.71$, $SE = 1.44$, $\beta = .49$, $t(116) = 3.26$, $p = .001$, $sr^2 = .08$, but not in the F-F dyads, $b = -.84$, $SE = 1.07$, $\beta = -.09$, $t(116) = -.79$, $p = .43$, $sr^2 = .01$.

The same pattern of results was found for the dyad members' ratings on the item *The interaction seemed smooth, natural, and relaxed to my partner*, where the interaction between LSS and the dyads' gender composition was again significant, $b = -2.73$, $SE = .82$, $\beta = -.31$, $t(116) = -3.31$, $p = .001$, $sr^2 = .08$. As before, the effect of LSS was significant for the M-M dyads, $b = 3.83$, $SE = 1.32$, $\beta = .44$, $t(116) = 2.90$, $p = .004$, $sr^2 = .07$, but not for the F-F dyads, $b = -1.62$, $SE = .98$, $\beta = -.19$, $t(116) = -1.65$, $p = .10$, $sr^2 = .02$.

Further evidence of this pattern was found for the item *My partner felt accepted and respected by me*. Once again, the interaction between LSS and gender composition was significant, $b = -2.60$, $SE = .82$, $\beta = -.29$, $t(116) = -3.16$, $p = .002$, $sr^2 = .08$. And, once again, LSS was a significant predictor of the post-interaction ratings of perceived acceptance and respect in the M-M dyads, $b = 3.86$, $SE = 1.31$, $\beta = .44$, $t(116) = 2.92$, $p = .004$, $sr^2 = .07$, but not in the F-F dyads, $b = -1.32$, $SE = .98$, $\beta = -.15$, $t(116) = -1.35$, $p = .18$, $sr^2 = .01$.

The consistency of the interaction effect across the three items suggests that one should take them seriously rather than regard them as chance-based effects. On the other hand, they raise the obvious question of why LSS emerged as significant predictor of these items for only the male-male dyads and not for the female-female dyads. A speculative discussion of this effect, and its implications, is included in the discussion section.

Hypothesis 3

A multiple regression analysis was used to test the hypothesis that the dyad-level measures of openness to experience, extraversion, and agreeableness would significantly predict dyad-level LSS. They did not. The results indicated that these predictors together accounted for

a non-significant proportion of the variance in LSS, $F(3, 117) = .94, p = .43, R^2 = .02$. A speculative discussion of these null findings is also included in the discussion section.

Hypothesis 4

With respect to Hypothesis 4, a mediation analysis was used to determine if dyad-level LSS mediated the relationship between dyad-level agreeableness and dyad-level global measure of perceived interaction quality (i.e., Factor 1). The total effect of agreeableness on the global measure of perceived interaction quality was significant, $b = .27, SE = .06, t(119) = 4.17, p = .0001, 95\% CI [.14, .39]$. However, the dyad-level effect of agreeableness on LSS was not significant, $b = .003, SE = .01, t(119) = .28, p = .78, 95\% CI [-.02, .02]$. The effect of LSS on the global measure of perceived interaction quality was not significant, $b = .60, SE = .65, t(119) = .92, p = .36, 95\% CI [-.69, 1.89]$. Not surprisingly, therefore, the direct effect of agreeableness on the global measure of perceived interaction quality, after controlling for all the mediator, was significant, $b = .26, SE = .06, t(119) = 4.14, p < .001, 95\% CI [.14, .39]$.

To determine whether the indirect effect of LSS on agreeableness and the global measure of perceived interaction quality was significant, a bias-corrected bootstrapping procedure was used with 1000 samples. The results showed that LSS was not a significant mediator of the relationship between agreeableness and the global measure of perceived interaction, Effect = .002, $SE = .01, 95\% CI [-.01, .03]$. The Sobel test of the indirect effect generally matched the result that was obtained using the bootstrapping procedure. That is, LSS was not a significant mediator, Effect = .002, $SE = .01, z = .19, p = .85$. It seems clear that LSS did not mediate the relationship between the dyad-level measures of agreeableness and the global measure of perceived interaction quality

Another mediation analysis was conducted to determine whether dyad-level LSS mediated the relationship between dyad-level extraversion and the dyad-level global measure of perceived interaction quality. The total effect of extraversion on the dyad-level global measure of perceived interaction quality was significant, $b = .22$, $SE = .09$, $t(119) = 2.45$, $p = .02$, 95% $CI[.04, .39]$. However, the effect of extraversion on LSS—the proposed mediator—was not significant, $b = .02$, $SE = .01$, $t(119) = 1.34$, $p = .18$, 95% $CI[-.01, .04]$. The effect of LSS on the global measure of perceived interaction quality was not significant, $b = .47$, $SE = .68$, $t(119) = .69$, $p = .49$, 95% $CI[-.88, 1.83]$. Not surprisingly, therefore, the direct effect of extraversion on quality of interaction, after controlling for the mediator, was still significant, $b = .21$, $SE = .09$, $t(119) = 2.34$, $p = .02$, 95% $CI[.03, .38]$.

To determine whether the indirect effect of LSS on extraversion and the global measure of perceived interaction quality was significant, a bias-corrected bootstrapping procedure was used with 1000 samples. The results revealed that LSS was not a significant mediator of the relationship between extraversion and the global measure of perceived interaction quality, Effect = $.01$, $SE = .02$, 95% $CI[-.01, .07]$. The Sobel test of the indirect effect generally matched the results that were obtained using the bootstrapping procedure. It also revealed that LSS was not a significant mediator of the relationship between extraversion and the global measure of perceived interaction quality, Effect = $.001$, $SE = .01$, $z = .51$, $p = .61$.

A third and final mediation analysis was conducted to determine whether dyad-level LSS mediated the relationship between dyad-level openness to experience and dyad-level perceived interaction quality. The total effect of openness to experience on perceived interaction quality was not significant, $b = .06$, $SE = .07$, $t(119) = .88$, $p = .38$, 95% $CI[-.08, .21]$. And, with regard to the relationship between the independent variable and the mediator, the effect of openness to

experience on LSS was not significant either, $b = -.002$, $SE = .01$, $t(119) = -.15$, $p = .88$, 95% $CI[-.02, .02]$. The relationship between the mediator and Factor 1 was not significant, $b = .68$, $SE = .69$, $t(119) = .98$, $p = .33$, 95% $CI[-.69, 2.05]$. Finally, the direct effect of openness to experience on quality of interaction, after controlling for the mediator, was not significant, $b = .07$, $SE = .07$, $t(119) = .89$, $p = .38$, 95% $CI[-.08, .21]$.

To determine whether the indirect effect of LSS on openness to experience and the global measure of perceived interaction quality was significant, a bias-corrected bootstrapping procedure was used with 1000 samples. The results revealed that LSS was not a significant mediator of the relationship between openness to experience and the global measure of perceived interaction quality, Effect = $-.001$, $SE = .01$, 95% $CI[-.03, .02]$. The Sobel test of the indirect effect generally matched the results that were obtained using the bootstrapping procedure. That is, LSS was not a significant mediator of the relationship between openness to experience and the global measure of perceived interaction quality, Effect = $-.001$, $SE = .01$, $z = -.11$, $p = .92$.

Hypothesis 5

A mixed model repeated-measures analysis of covariance was used to examine the trajectory of dyad-level LSS across the three 6-minute time periods. In this model, the dyad's gender composition (male-male vs. female-female) was analyzed as a between-dyad factor and the dyad's LSS score for each of the three time periods was analyzed as a within-dyad factor. The main effects of time period and gender composition, as well as the interaction between time period X gender composition, were tested in this model, in which I controlled for differences in the number of words used in each interaction period by using total word count per period as a covariate.

There was a significant main effect of gender, $F(1, 145.80) = 4.34, p = .04$, such that female-female dyads ($M = .51, SE = .02$) scored significantly higher in LSS than male-male dyads ($M = .48, SE = .02$). There was also a significant main effect of time, $F(2, 257.49) = 14.06, p < .001$. According to the results of Bonferroni post-hoc analyses, LSS was significantly different between all three time periods (see Figure 1). LSS was significantly higher in Time 1 (i.e., the first 6-minutes of interaction) than Time 2 (i.e., the second 6-minutes of interaction; $p = .02$) and Time 3 (i.e., the third 6-minutes of interaction; $p < .001$), and LSS in Time 2 was significantly higher than Time 3 ($p = .047$). At odds with hypothesis 5, this pattern of results revealed that LSS decreased, rather than increased, over time. The covariate was also significant, $F(1, 244.46) = 51.86, p < .001$, indicating that total word count per time period does have an effect on the outcome variable, dyad-level LSS over the 3 time periods. However, even after controlling for the covariate, the main effects of gender and time still emerge as significant. The gender X time interaction was not significant,

Why Did LSS Decrease, Rather Than Increase, Over Time?

Why did the dyad-level LSS scores decrease, rather than increase, across the three interaction periods? One possibility is that strangers work harder to achieve an acceptable level of LSS at the beginning of their online interaction. However, as soon as they feel that their level of semantic similarity is sufficient to sustain the conversation, their efforts to achieve an acceptable level of LSS are relaxed and LSS then decreases.

Previous research and theory provide evidence that is consistent with this interpretation. In one study that examined initial interactions between same-sex strangers (Ickes, Patterson, Rajecki, & Tanford, 1982), perceivers who were led to expect that their partner would exhibit unfriendly behavior compensated for this expected unfriendly behavior by expressing a higher

than usual level of positive affect in order to activate the targets' "friendly" behavior. The use of this compensatory strategy was presumably a way to mitigate the potential costs of an expected unpleasant interaction. And, not surprisingly, the compensatory behavior was most evident during the earliest phase of these interactions, and was "relaxed" later on, once it was clear that the interaction was not going to be unpleasant (Ickes, Miles, Rajcecki, & Tanford, 1982).

The need to "work harder" at first to ensure a pleasant interaction experience is also emphasized by Berger and Calabrese's (1975) *uncertainty reduction theory*. According to this theory, when strangers interact for the first time, they are especially motivated to reduce their uncertainty (that is, to increase their mutual predictability) early in their interaction because of their recognition that they do not yet possess relevant knowledge about the beliefs and attitudes of the other party. This uncertainty is reduced through interpersonal communication, specifically when both interaction partners ask for and give the same kinds of information at the same rate of exchange. However, as soon as their initial uncertainty has been substantially reduced, this rapid exchange of information is no longer needed, and they settle into a more relaxed level of mutual information exchange.

The logic of this reasoning suggests a way to test it: examine the content of what the dyad members say to each other during the first 6-minute interaction period and see if there is evidence of their trying harder to get in sync linguistically than they are in the two subsequent interaction periods. Fortunately, a software tool for conducting such an exploration is available in the Language Inquiry and Word Count (LIWC) program, which generates the percentage of words that represent different linguistic categories (e.g., parts of speech, personal pronouns, social concerns, and emotions) that are present in any corpora of text (Tausczik & Pennebaker, 2010). Previous studies have shown that the words individuals use to talk and write are

correlated with their physical and mental health and can provide a glimpse into their current cognitive state (Gottschalk & Glaser, 1969; Rosenberg & Tucker, 1978; Stiles, 1992). For a broader view of how the LIWC software has been applied, see Campbell & Pennebaker, 2002; Fratteroli, 2007; Lepore & Smyth, 2002; Pennebaker, 1997; Pennebaker, Mayne, & Francis, 1997)

If dyad members are trying harder to get in sync linguistically at the beginning of their interaction than in the two subsequent interaction periods, I predicted that there would be a higher frequency of “you” category words in the first 6-minutes of interaction than in the subsequent interaction periods. A higher frequency of “you” category words in the first 6-minutes of interaction suggests that dyad members are directly involving their interaction partner in their conversation and are asking each other more questions initially in order to maximize their level of LSS and to reduce the level of uncertainty in their interaction. The transcripts of each dyad’s chat interaction by time period were run through the LIWC software. The percentages of each word category, along with each dyad’s LSS indices for each 6-minute interaction period, were entered into an analysis of variance model to determine if there was a greater frequency of “you” category words between any of the three 6-minute interaction periods.

As predicted, there was a significant main effect of “you” category words, $F(2, 333) = 24.52, p < .001$. Bonferroni post hoc analyses indicated that the first 6-minutes contained a significantly higher amount of “you” category words ($M = 3.68, SE = .19$) than both the second ($M = 2.41, SE = .13; p < .001$) and the third 6-minutes ($M = 2.30, SE = .13; p < .001$) of interaction. On the other hand, there was not a significant difference of “you” category words between the second and third 6-minutes of interaction ($p = 1.00$).

There was also a main effect of “they” category words, $F(2, 333) = 7.86, p < .001$. Bonferroni post hoc analyses indicated that the first 6-minutes contained a significantly lower amount of “they” category words ($M = .17, SE = .03; p = .002$) than the second 6-minutes ($M = .51, SE = .08; p = .002$) and the third 6-minutes ($M = .50, SE = .08; p = .002$) of interaction. There was not a significant difference between the second and third 6-minutes of interaction ($p = 1.00$). This suggested that dyads switch from directly involving their interaction partner in their conversation to talking about others once they have achieved an acceptable level of mutual understanding and LSS at the beginning of their online interaction.

Providing further evidence for a “compensation effect” during the earliest stage of the dyads initial interactions, there were also significant main effects of “social” category words (i.e., words that denote social processes such as verbs that suggest human interaction, like *talking* and *sharing*), $F(2, 333) = 8.01, p < .001$; “negative emotion” category words, $F(2, 333) = 3.15, p = .04$; and “sad” category words, $F(2, 333) = 4.04, p = .01$. According to the results of Bonferroni post-hoc analyses, there was a higher percentage of “social” category words used during the first 6 minutes of interaction ($M = 10.94, SD = 3.65$) than during the second 6-minutes ($M = 9.02, SD = 3.55; p = .001$) or third 6-minutes ($M = 9.40, SD = 4.16; p = .01$) of interaction.

This pattern suggests that dyad members are more focused on the need to make an effective social connection in the earliest phase of their initial interaction than they are later on. Interestingly, there were higher percentages of “negative emotion” category words in the third 6-minutes ($M = 1.09, SD = 1.17$) than the first 6-minutes ($M = .77, SD = .76; p = .047$) of interaction and higher instances of “sad” category words in the third 6-minutes ($M = .18, SD = .44$) than both the second 6-minutes ($M = .08, SD = .19; p = .045$) and first 6-minutes ($M = .07, SD = .21; p = .02$) of interaction. These findings would follow from the compensation hypothesis

of one assumes that dyad members withhold these “negative emotion” and “sad” category words when they are actively trying to establish high levels of positive affect, mutual rapport, and latent semantic similarity (i.e., LSS) during the earliest phase of their initial interaction.

Does Dyad-Level Extraversion Influence the Trajectory of LSS Over Time?

Additional analyses were conducted to further explore why LSS decreased, rather than increased, over time. In these analyses, I investigated whether any dyad-level characteristics influenced this effect. I suspected that another variable, perhaps a variable related to the tendency to seek stimulation in the company of others, could have moderated this effect. I therefore tested to see if dyad-level differences in extraversion level could have influenced the development of LSS across the three time periods.

To investigate this possibility, dyads were classified as representing one of the following three levels of extraversion: high extraversion, medium extraversion, and low extraversion. Dyad-level extraversion scores that were .5 standard deviations above and below the mean were identified as dyads with high extraversion ($n = 35$) and low extraversion ($n = 34$), respectively. The remaining dyads were identified as medium extraversion ($n = 51$; See Appendix J for the frequency distribution of the number of cases within the ranges represented in a normal distribution). Multi-level modeling was then used to determine if gender and extraversion level moderated the trajectory of dyad-level LSS across the three time periods.

Multilevel modeling is a statistical method that is commonly used to analyze hierarchical data; it permits the computation of residual components at each level in the hierarchy (Howell, 2013). This method was used to account for the variance in the three LSS indices for each dyad that pertain to the three time periods of their interaction. All factors were fixed, with the dyad identifier treated as a random factor. An intercepts-only model and a model with all variables

were run using a maximum likelihood estimation method, a variance components covariance matrix for random effects, and a Toeplitz pattern for repeated effects (see Table 2 for fit indices and improvement). As expected, there was a significant main effect of time, $F(2, 230.03) = 10.44, p < .001$. The three-way interaction between gender, extraversion level, and time period was also significant, $F(4, 213.27) = 3.77, p = .01$ (see Figure 2). There were no other significant main effects or interactions (Table 3 reports the degrees of freedom, F values, and p values for this analysis).

Probing the 3-way Interaction

Dyad-level LSS for the female-female (FF) low extraversion dyads was significantly higher in the first 6-minutes than in the second 6-minutes of interaction ($p = .01$) and the third 6-minutes of interaction ($p < .001$); in addition, LSS in the second 6-minutes was significantly higher than LSS in the third 6-minutes of interaction ($p = .02$). For the male-male (MM) low extraversion dyads, however, there were no significant differences in LSS across time. Further, the MM dyads had significantly lower LSS scores than the FF dyads during the first 6-minutes of interaction, $p = .02$. However, in the third 6-minutes of interaction, the MM dyads had significantly higher LSS scores than the FF dyads, $p = .01$. Overall, this pattern of results revealed that the decline of LSS across time was most evident in FF dyads with low extraversion. Although LSS also increased for MM dyads with low extraversion, this increase was not significant. Low extraversion MM dyads began their interaction with a lower level of LSS than low extraversion FF dyads did, but then finished their interaction with a higher level of LSS because it did not decline as much as it did in the FF dyads.

Dyad-level LSS for medium extraversion FF dyads in the first 6-minutes was significantly higher than in both the second 6-minutes ($p = .02$) and the third 6-minutes ($p =$

.002) of interaction. There was not a significant difference in LSS between the second 6-minutes and the third 6-minutes of interaction. Thus, LSS seems to decline significantly after the first 6-minutes of interaction for medium extraversion FF dyads. Similarly, for medium extraversion MM dyads, LSS in the first 6-minutes was significantly higher than LSS in the third 6-minutes of interaction ($p < .001$). There was not a significant difference of LSS between the first 6-minutes and the second 6-minutes of interaction.

The reader should note that LSS for medium extraversion MM dyads did not significantly decline until the last 6-minutes of interaction, whereas it significantly declined after the first 6-minutes of interaction for medium extraversion FF dyads. As in the case of the data for the low extraversion dyads, the FF dyads showed an early decline whereas the MM dyads did not. However, the later decline within the moderate extraversion MM dyads was a substantial one, as indicated by the fact that the MM dyads had significantly lower LSS than FF dyads in the last 6-minutes of interaction, $p = .01$. There was not a significant difference in LSS between MM and FF moderate extraversion dyads for the first twelve minutes of interaction.

Unlike dyads with low or medium extraversion, LSS did not differ across time for either MM dyads nor FF dyads with high extraversion. Instead, LSS remained stable over time and did not significantly decline across time. There were also no gender differences in the trajectory of LSS across time. Table 4 provides the relevant means and standard errors.

These results indicate that dyads who score high in extraversion experience a more stable trajectory of LSS across time than dyads with low or medium extraversion, and this may be due to their natural tendency to seek stimulation in the company of others. After establishing a high level of LSS at the beginning of their interaction, dyads with low or medium levels of extraversion do not maintain the same level of LSS as dyads with high extraversion do—instead,

their LSS declines significantly by the end of their interaction. This decline is consistent with the compensatory behavior that occurs early, but not later, in an initial interaction, and the only extraversion range in which it is not observed is in high extraversion dyads who presumably seek to maintain a high level of stimulation in the company of others.

The Slope of LSS across Time

In addition to looking at LSS across time as three separate time points, the slope of dyad-level LSS across time (i.e., the rate of change of LSS across time) was generated for each dyad and used for analysis (Yin, Schmidt, & Besag, 2006; for more detailed information, see Appendix I). The slope represents the best-fit line across the three time points and allows for the consolidation of the three LSS indices across time into one observation. This single measure illustrates the trajectory of LSS across time in a different, but complementary, form.

A mixed-model analysis of variance was used to determine if gender and extraversion level moderated the slope of LSS. In this model, the dyad's gender composition and extraversion level were treated as between-dyads factors, whereas the slope of LSS across time was treated as a within-subjects factor. The main effects of extraversion level and gender composition, as well as the interaction between extraversion level X gender composition, were tested in this model, in which I controlled for differences in the number of words used in each interaction period by using total word count per period as a covariate.

The interaction between extraversion level and gender was significant, $F(2, 113) = 5.82$, $p = .004$, $\eta^2 = .09$ (See Figure 3). After probing this interaction, the results indicated that low extraversion FF dyads showed a larger decrease of LSS (i.e., had a steeper slope of LSS) across time than low extraversion MM dyads did ($p = .004$). This finding seems to capture the “gist” or essence of the gender difference that was revealed in the earlier analysis that examined LSS

during all three time periods. The decreasing slope of LSS across time for low extraversion FF dyads was significantly steeper than for high extraversion FF dyads ($p = .001$). This finding suggests that the low extraversion FF dyads may have lost interest in participating in the online interaction at a faster rate than the high extraversion FF dyads did, and therefore stopped working so hard to maintain a high level of LSS. The decreasing slope of LSS across time for medium extraversion MM dyads was steeper than that of the low extraversion MM dyads ($p = .03$), but there is no obvious reason for this unexpected difference.

Discussion

The present findings identify important factors that influence the development of LSS in the initial interactions of same-sex partners who are using computer-mediated communication (i.e., AOL Instant Messenger) to “talk” with each other. First, based on previous findings reported by Ta et al. (2015), I predicted that the dyad-level behaviors that introduced more words into the conversation (i.e., the number of questions asked and number of messages sent) would be significant behavioral predictors of the dyads’ LSS scores. The results showed, however, that it was the total number of messages that were sent by the dyad members, but not the total number of questions they asked each other, that significantly predicted their dyad-level LSS scores. Why was the total number of questions asked a non-unique predictor of LSS? The most likely reason is that this variable is a subset of the total number of messages sent, and does not contribute anything unique to LSS that is not already captured by the total number of messages exchanged.

Second, because getting on the same page linguistically and coming to “use the same words in the same way” should elicit a positive reaction from pairs of strangers who are interacting for the first time (Babcock et al, 2014; Ta et al., 2015) I predicted that higher LSS

scores, computed across the entire 18-minute interaction, would be associated with higher levels of perceived interaction quality. The results showed that dyad-level LSS did not predict the overall index of perceived interaction quality; however, it did predict certain aspects of perceived interaction quality, but for the male-male dyads only. These components included (1) the extent to which the interaction seemed smooth, natural, and relaxed to the dyad members, (2) the extent to which they thought their partner also perceived the interaction as smooth, natural, and relaxed; and (3) the extent to which they thought that their interaction partners felt accepted and respected by them.

Simply put, a higher level of LSS contributed to the perception of higher-quality initial interactions in MM dyads, but not in FF dyads. But what is the reason for this difference? A possible reason could be the range differences of LSS between genders. The range of LSS for FF dyads was narrower (i.e., .5 to .9) than MM dyads (i.e., .4 to .9). This may not seem like a large difference, but considering that LSS among all dyads ranges from .4 to .9, this amount accounts for 20% of the difference between genders. In addition, female-female dyads were overall rated as having higher quality interactions ($M = 4.10$, $SD = .58$) than male-male dyads ($M = 3.98$, $SD = .40$). Although this difference was not significant, it may have also contributed to a range-restriction-based gender difference in the correlation between LSS and certain post-interaction measures of perceived interaction quality.

Third, because agreeableness, extraversion, and openness to experience are personality characteristics that should lead to behaviors that facilitate a “common-ground understanding”, I predicted that they should account for some of the variance in dyad-level LSS. Despite the intuitive plausibility of this hypothesis, the results revealed no support for it. Instead, they showed that the dyad level measure of the partners’ agreeableness, extraversion, and openness to

experience did not predict dyad-level LSS, nor did dyad-level LSS mediate the relationship between the dyad-level measures of agreeableness, extraversion, and openness to experience and a global measure of dyad members' interaction experience (i.e., an averaged measure of the items of the post-interaction questionnaire).

The reason for these unexpected result null findings is unclear, but it is possible that they can be traced to the way that personality is expressed over the computer and on the internet versus how it is expressed in face-to-face interactions. Previous studies have shown that the Big Five personality traits, including extraversion, agreeableness, and openness to experience, tend to be expressed more *weakly in* computer-mediated communication than in face-to-face communication. This is due to the unique characteristics that are created through computer-mediated communication, such as (1) the relative ease of exerting greater social control within these interactions (for example, if someone is usually shy when interacting with people who are very physically attractive, he or she might be able to control his or her shyness through computer-mediated interactions because his or her interaction partner's physical attractiveness is not affecting the interaction); (2) the greater perceived anonymity between interaction partners; and (3) the greater perceived "distance" between interaction partners. These characteristics can lead to more inconsistent and unpredictable behavior that is generally more reflective of the limitations of online communication than of the interaction partners' respective personalities (Blumer & Doering, 2012; Stritzke, Nguyen, & Dirkin, 2004).

Fourth, I predicted that the trajectory of dyad-level LSS should increase over the course of each dyad's 18-minute interaction. Instead, dyad-level LSS actually *decreased* over time. These results are contrary to the commonsense expectation that led Ta et al. (2015) to argue that, over the course of their initial interaction, strangers should engage in behaviors (specifically,

talking and asking questions) that add more and more words into the conversation and therefore enable them to increasingly sample each other's word choices and mutually align their word choices and intended meanings, which should result in higher LSS.

Why was dyad-level LSS higher at the beginning of the interaction and subsequently decreased over time? I have proposed two possible explanations for this effect. First, it is possible that strangers work harder to achieve an acceptable level of LSS at the beginning of their online interaction and, when they feel that their level of semantic similarity is sufficient enough to sustain the conversation, their efforts to achieve an acceptable level of LSS are relaxed. The results from the LIWC analysis do support a compensation effect: a higher frequency of "you" and "social" category words in the first period of interaction compared to subsequent periods of interaction suggest that interaction partners are spending the beginning of their interaction getting to know each other by talking and asking each other questions as reflected by the increased usage of "you" category words (e.g., How are *you*? What are *you* majoring in? What classes are *you* taking?). This behavior, along with the avoidance of using "negative emotion" and "sad" category words during this time, provide a way for interaction partners to not only align their own word choices and intended meanings with each other when they are less familiar with each other (i.e., during the beginning of their interaction), but is also used to reduce the potential costs of an unpleasant interaction. Once dyads reach an acceptable level of LSS, they no longer need to provide this kind of information at the same rate of exchange and can shift the focus to third-party others (as indicated by the higher frequency of "they" category words in the second and third 6-minutes of interaction) in their interaction.

A second factor that may play a role in the declining trajectory of LSS across time is dyad-level extraversion level. Overall, the pattern of results reveals that dyads with high

extraversion levels tend to have the most stable trajectory of LSS across time, whereas dyads with low or medium extraversion levels tend to experience a significant decline of LSS over time. These pattern of results are consistent with the results of previous research examining the influence of extraversion level in interpersonal communication. For instance, Funder and Sneed (1993) reported evidence that an increase in extraversion is related to observers' ratings of more talking, disclosing more information about oneself, and enjoying an initial interaction more. These findings suggest that highly extraverted dyad members should be the least likely to scale back their involvement in an initial interactions, and might therefore be expected to maintain a more stable level of LSS over time. However, to be confident about the reliability and correct interpretation of the present findings, additional research will be required.

Summary of the Major Findings

In summary, there were several major findings that emerged in the data for this study. First, the total number of messages that were sent by the dyad members significantly predicted their dyad-level LSS scores. Second, a higher level of LSS contributed to the perception of higher-quality initial interactions in male-male dyads, but not in female-female dyads. Third, perhaps because of the weaker expression of personality in computer-mediated interactions than in face-to-face interactions, the personality traits of agreeableness, extraversion, and openness to experience did not predict LSS in the present study; nor did LSS mediate the relationship between these three personality traits and the perceived quality of their initial interactions. Finally, the data revealed that dyad-level LSS decreased across time, and that this decline was attributable primarily to dyads with lower levels of extraversion.

Recent technological developments have provided us with tools such as email, instant messaging, and text messaging (just to name a few) that allow for easy communication with

people all over the world at any time we choose. Whether it is between colleagues, friends, family, or strangers, a large portion of our daily communication now takes place through computer-mediated interactions. Although these tools allow people to be more well-connected with each other, it is still a relatively new type of communication; as such, a lot is yet to be uncovered regarding the processes and the resulting effects of computer-mediated interaction. This study provides a first glimpse into how interaction partners develop a mutual understanding for each other (i.e., develop LSS) during their initial interactions through computer-mediated interaction, and how personality characteristics, especially extraversion, influences the development of LSS.

Strengths and Limitations

With regard to its strengths, the current study is, to my knowledge, the first to investigate how LSS is achieved by first-time interaction partners using computer-mediated communication (i.e., AOL Instant Messenger). The total interaction period in this study was considerably larger longer than in previous studies that investigated LSS within initial dyadic interactions (Babcock et al., 2014; Ta et al., 2015). This feature of the study enabled me to compute an LSS index for each dyad for the entire 18-minute interaction and also provided sufficient time to break the interaction down into three 6-minute periods and track LSS across successive stages of the interaction. Although the participants were given instructions and prompted by the experimenter to begin interacting with their partner, they were also informed by the experimenter during the chat procedure that they could interact with their chat partner as much or as little as they chose. This aspect of the procedure gave the dyad members the latitude to regulate their own level of interactional involvement over time.

The young-adult age range of the participants suggests that they have had considerable exposure to and experience with computers. Possessing computer skills and having some level of familiarity with instant messaging programs (or a lack thereof) should influence the way they react in online initial interactions, and thus, might also influence LSS. As computer-mediated interactions become even more commonplace in society (as they are projected to be), the number of people with considerable experience and exposure with computers and computer-mediated interactions will presumably be higher than ever before. Thus, the sample in the present study is likely to be representative of the majority of people in the near future. But what about the people who do not have considerable experience with computers? In this regard, the limited age-range of the present sample might be viewed as a limitation of the study because it most likely “undersamples” individuals with limited computer skills. Future studies should therefore include a more diverse age range within their samples and question the participants about their familiarity with online communication to capture any age-related nuances in online communication and related technological skills.

Directions for Future Research

The current study examines the trajectory of LSS across an interaction period of 18 minutes. Due to the accessibility and relative ease of use of computer-mediated interaction, these interactions between interaction partners can persist for long periods of time. Thus, future research should employ longitudinal study designs that investigate LSS across longer periods of time. Doing so would answer several questions: for instance, at what point during the interaction does LSS cease to decline and begin to stabilize? Does LSS return to its initial levels after a certain amount of interaction? Or, is the high level of LSS that is achieved in the initial stages of interaction unique to only that period?

Second, although extraversion was the only personality characteristic that had any significant effect on LSS in the current study, this isolated finding does not mean that the other personality characteristics have absolutely no influence on LSS. As noted above, personality is expressed differently in computer-mediated interactions than in face-to-face interactions; accordingly, future research should investigate how and in what situations personality is expressed online and the extent to which such expression occurs. Not only can such research determine what kinds of effects personality can have on the latent semantic similarity that develops between interaction partners, but it can potentially answer other questions pertaining to the role of personality in online and face-to-face interactions more generally.

Third, future research should consider if the amount of information that dyads have about each other prior to their interaction can affect the development of LSS. For instance, if dyad members were provided with their interaction partner's picture, demographic information, or interests/hobbies prior to their interaction (similar to a situation on online dating websites), would this further facilitate the development of LSS? Would more attractive interaction partners maintain higher LSS throughout their interaction than less attractive interaction partners? Or, if the initial interaction were between student and teacher, colleagues, family members, or others, could these different relationship dynamics also influence how interaction partners develop a basis of understanding with each other?

Fourth, because the LSS index is calculated using the words that interaction partners use in their interaction, future research should also investigate if the interaction partners' "language community" influences their dyad-level LSS. For example, there are 24 recognized dialects, or "language communities," of American English and each of these dialects have their own grammar, vocabulary, and common phrases that distinguishes them from other dialects of

American English (Delaney, 2010). Would strangers have higher levels of LSS if they share the same, versus different, dialect of American English? Similarly, would strangers have higher levels of LSS if their dialects come from regions that are geographically closer to each other? Or would differences in “language community” lead to higher compensation effects that last for more than the first 6-minutes of interaction?

Finally, researchers should explore how motivation for the interaction can influence the development of LSS. In the current study, participants interacted with each other simply because the experimenter requested that the participants chat with their partner in order to get acquainted with them, and said that they were free to discuss any topics they would like. Would a different prompt for interaction influence their development of LSS? For instance, if strangers were assigned to work together interdependently to successfully complete a project or to solve an important problem, would they develop higher levels of LSS and/or develop them more rapidly? Additionally, if incentives were tied to the outcome of their interaction (e.g., each interaction partner would receive a \$100 prize if the interaction partners successfully completed a project together), would LSS remain high throughout the interaction? In other words, would greater or more desirable incentives result in higher levels of LSS and/or a more rapid development of LSS? These are just some of the many interesting questions that researchers could address to further our understanding of the antecedents and consequences of shared semantic similarity (LSS) in computer-mediated communication.

References

- Abbeduto, L., Short-Meyerson, K., Benson, G., Dolish, J., & Weissman, M. (1998). Understanding referential expressions in context: Use of common ground by children and adolescents with mental retardation. *Journal of Speech, Language, and Hearing Research, 41*, 1348-1362.
- Arnulf, J. K., Larsen, K. R., Martinsen, O. L., & Bong, C. H. (2014). Predicting survey responses: How and why semantics shape survey statistics on organizational behaviour. *PLoS ONE, 9*(9): e106361.
- Ashton, M. (2013). *Individual differences and personality* (2nd ed.). Burlington, MA: Elsevier Academic Press.
- Babcock, M. J., Ta, V. P., & Ickes, W. (2014). Latent semantic similarity and language style matching in initial dyadic interactions. *Journal of Language and Social Psychology, 33*(1), 77-88.
- Berger, C. R., Calabrese, R. J. (1975). "Some Exploration in Initial Interaction and Beyond: Toward a Developmental Theory of Communication". *Human Communication Research, 1*, 99-112.
- Blumer, T., & Doering, N. (2012). Are we the same online? The expression of the five factor personality traits on the computer and the internet. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 6*(3), article 5.

- Bruder, M., Dosmukhambetova, D., Nerb, J., & Manstead, A. R. (2012). Emotional signals in nonverbal interaction: Dyadic facilitation and convergence in expressions, appraisals, and feelings. *Cognition and Emotion, 26*(3), 480-502.
- Campbell, R.S. & Pennebaker, J.W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*, 60-65.
- Caspar, F., Berger, T., & Hautle, I. (2004). The right view of your patient: A computer-assisted, individualized module for psychotherapy training. *Psychotherapy: Theory, Research, Practice, Training, 41*, 125-35.
- Delaney, R. (2010). A Dialect Map of American English.
<http://www15.uta.fi/FAST/US1/REF/dial-map.html>
- Funder, D.C., & Sneed, C.D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64*, 479-490.
- Gesn, P.R., & Ickes, W. (1999). The development of meaning contexts for empathic accuracy: Channel and sequence effects. *Journal of Personality and Social Psychology, 77*, 746-761.
- Gottschalk, L.A., & Gleser, G.C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. Berkeley: University of California Press.
- Katsanos, C., Tselios, N., & Avouris, N. (2008). Automated semantic elaboration of website information architecture. *Interacting with Computers, 20*(6), 535-544.
- Keller, M. (2013). Social media and interpersonal communication. *Social Work Today, 13*(3), 10.
- Keckes, I., & Zhang, F. (2009). Activating, seeking, and creating common ground: A sociocognitive approach. *Pragmatics & Cognition, 17*, 331-355.

- Krauss, R. M., & Fussell, S. R. (1991). Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition, 9*, 2-24.
- Laham, D., Steinhart, D., Jones, M., Stahl, G., dePaula, R., Landauer, T., & Rehder, B. (1998). What is LSA? *Latent Semantic Analysis @ CU Boulder*. Retrieved from <http://lsa.colorado.edu>.
- Landauer, T. K., Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211-240.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284.
- Landauer, T. K., Laham, R. D., & Foltz, P. W. (2003). Automated scoring and annotation of essays with the intelligent essay assessor. *Assessment in Education, 10*, 295-308.
- Lautenschlager, N. T., Dunn, J. C., Bonney, K., Flicker, L., & Almeida, O. P. (2006) Latent semantic analysis: An improved method to measure cognitive performance in subjects of non-English-speaking background. *Journal of Clinical and Experimental Neuropsychology, 28*, 1381-1387.
- Lepore, S. J., & Smyth, J. M. (2002). *The Writing Cure: How Expressive Writing Promotes Health and Emotional Well-Being*. Washington, DC: American Psychological Association.
- Manusov, V., & Patterson, M. L. (2006). *The sage handbook of nonverbal communication*. Thousand Oaks, CA: Sage Publications, Inc.
- Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality Traits*. New York, NY: Cambridge University Press.

- Mihalcea, R., Corley, C., & Strapparava, C. (2006). Corpus-based and knowledge-based measures of text semantic similarity. *Association for the Advancement of Artificial Intelligence*, 6, 775-780.
- Morganti, F. (2008). What intersubjectivity affords: Paving the way for a dialogue between cognitive science, social cognition and neuroscience. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions*. Amsterdam, Netherlands: IOS Press, 10, 3-14.
- Onishi, A.T., & Steger, A.P. (2012). Our information: Availability, communication, and perception. In D.G. Nemeth, R.B. Hamilton, J. Kuriansky (Eds.), *Living in an environmentally traumatized world: Healing ourselves and our planet* (pp. 199-206). Santa Barbara, CA, US: Praeger/ABC-CLIO.
- Patterson, M. L. (2014). Reflections on historical trends and prospects in contemporary nonverbal research. *Journal of Nonverbal Behavior*, 38, 171-180.
- Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, 8, 162-166.
- Pennebaker, J. W., Mayne, T., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology*, 72, 863-871.
- Radicati, S. Email Statistics Report, 2013-2017. The Radicati Group, Inc.
- Recchia, G., & Jones, M. N. (2009). More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods*, 41(3), 647-656. HarperCollins Publishers.
- Rosenberg, S.D. & Tucker, G.J. (1978). Verbal behavior and schizophrenia: The semantic dimension. *Archives of General Psychiatry*, 36, 1331-1337.

- Ryckman, R.M. (2012). *Theories of Personality* (10th ed.). Belmont, CA: Wadsworth.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211-232.
- Stiles, W.B. (1992). *Describing talk: A taxonomy of verbal response modes*. Newbury Park, CA: Sage.
- Ta, V.P., Babcock, M.J., & Ickes, W. (2014). *The words may be all you need: Latent semantic similarity in initial, unstructured interactions*. Manuscript submitted for publication.
- Tausczik, Y.R., & Pennebaker, J.W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24-54.
- Tardanico, S. (2012, April). Is social media sabotaging real communication? *Forbes Magazine*. Retrieved from <http://www.forbes.com/sites/susantardanico/2012/04/30/is-social-media-sabotaging-real-communication/>
- Thorne, A. (1987). The press of personality: A study of conversations between introverts and extraverts. *Journal of Personality and Social Psychology*, 53(4), 718-728.
- Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, 31, 183-194.
- Yin, R., Schmidt, R., & Besag, F. (2006). Aggregating student achievement trends across states with different tests: Using standardized slopes as effect sizes. *Peabody Journal of Education*, 81(2), 47-61.

Table 1. Factor Loadings of the Post-Interaction Questionnaire Items

Factor	Items	Factor loadings	% of variance
1: Perceived Interaction Quality			68.49%
	The interaction seemed smooth, natural, and relaxed to me.	0.66	
	The interaction seemed smooth, natural, and relaxed to my partner.	0.65	
	The interaction seemed awkward, forced, and strained to me.	0.60	
	The interaction seemed awkward, forced, and strained to my partner.	0.85	
	I felt accepted and respected by my partner,	0.72	
	My partner felt accepted and respected by me.	0.72	
	I felt put down, patronized, or rejected by my partner.	0.87	
	My partner felt put down, patronized, or rejected by me.	0.86	
2: Emotional Connection with Partner			5.75%
	I felt a connection with my partner.	0.81	
	My partner felt a connection with me.	0.88	
	I liked my partner.	0.66	
	My partner liked me.	0.78	
	I would like to interact more with my partner in the future.	0.71	
	My partner would like to interact more with me in the future.	0.78	
	I understood my partner.		
	My partner understood me.	0.60	
	I enjoyed the interaction with my partner.	0.73	
	My partner enjoyed the interaction with me.	0.76	
	I felt comfortable with my partner.	0.64	
	My partner felt comfortable interacting with me.	0.65	
Total factor variance			74.24%

Note. These results are from a varimax rotation factor analysis. All factors had eigenvalues greater than 0.99.

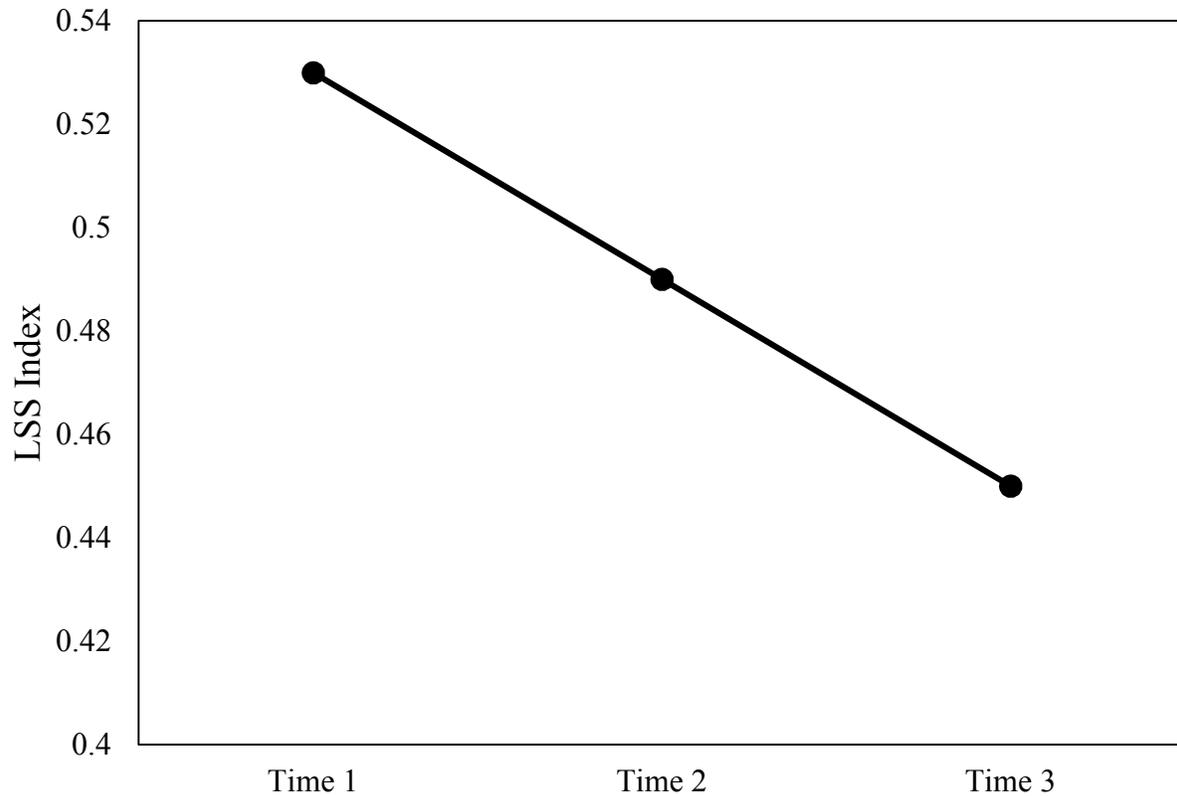


Figure 1. The trajectory of dyad-level LSS across the 18-minute interaction. Time 1 represents the first 6 minutes of interaction; Time 2 represents the middle 6 minutes of interaction; Time 3 represents the last 6 minutes of interaction. Time 1 and Time 2 were significantly different from each other ($p < .05$); Time 2 and Time 3 were significantly different from each other ($p < .05$); Time 1 and Time 3 were significantly different from each other ($p < .001$).

Table 2. Fit Indices for Model I and Model II

	Model I: Unadjusted	Model II: Adjusted for extraversion level, gender, time
-2 Log Likelihood	-446.78	-575.92
Akaike's Information Criterion (AIC)	-438.78	-529.92
Hurvich and Tsai's Criterion (AICC)	-438.67	-526.63
Bozdogan's Criterion (CAIC)	-419.20	-417.54
Schwartz's Bayesian Criterion (BIC)	-423.20	-440.54

Table 3. Statistics for Multilevel Model

Fixed effect	Model I: unadjusted		Model II: adjusted for extraversion level, gender, time	
	<i>df</i>	<i>F</i>	<i>df</i>	<i>F</i>
Intercept	113.58	6593.87**		
Gender			157.90	3.61
Extraversion level			122.13	1.09
Time period			230.03	10.44**
Gender*Extraversion level			159.54	1.36
Gender*Time period			213.64	0.97
Time period*Extraversion level			197.96	1.64
Gender*Extraversion level*Time period			213.27	3.77*

Note: * $p < .05$. ** $p < .001$.

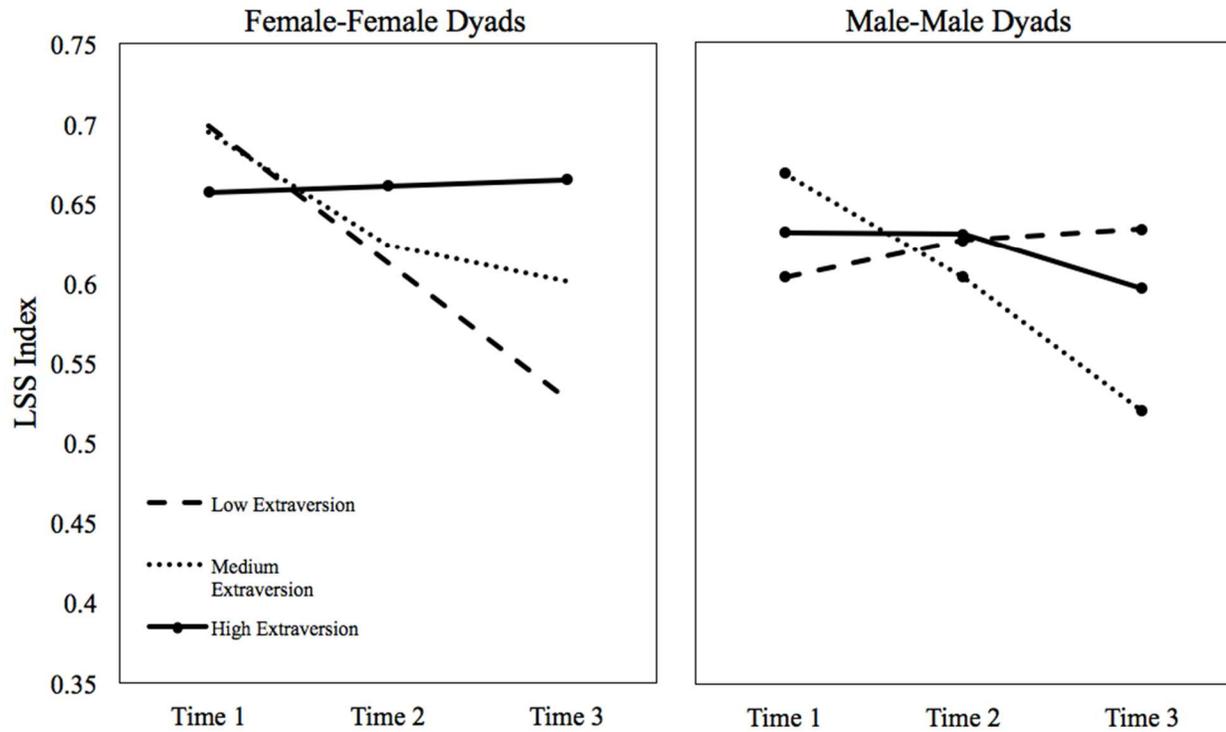


Figure 2. The significant 3-way interaction between gender, time period, and extraversion level on the outcome measure, LSS.

Table 4. Means and Standard Errors for 3-Way Interaction

		<i>M (SE)</i>		
		Time 1	Time 2	Time 3
FF Dyads				
	Low Extraversion	0.70 (0.02)	0.61 (0.02)	0.53 (0.02)
	Medium Extraversion	0.70 (0.02)	0.62 (0.02)	0.60 (0.02)
	High Extraversion	0.66 (0.02)	0.66 (0.02)	0.67 (0.02)
MM Dyads				
	Low Extraversion	0.60 (0.04)	0.63 (0.03)	0.63 (0.04)
	Medium Extraversion	0.67 (0.02)	0.60 (0.03)	0.52 (0.02)
	High Extraversion	0.63 (0.03)	0.63 (0.03)	0.60 (0.03)

Note. Means for the gender X extraversion level X time period interaction. Standard errors are in parentheses.

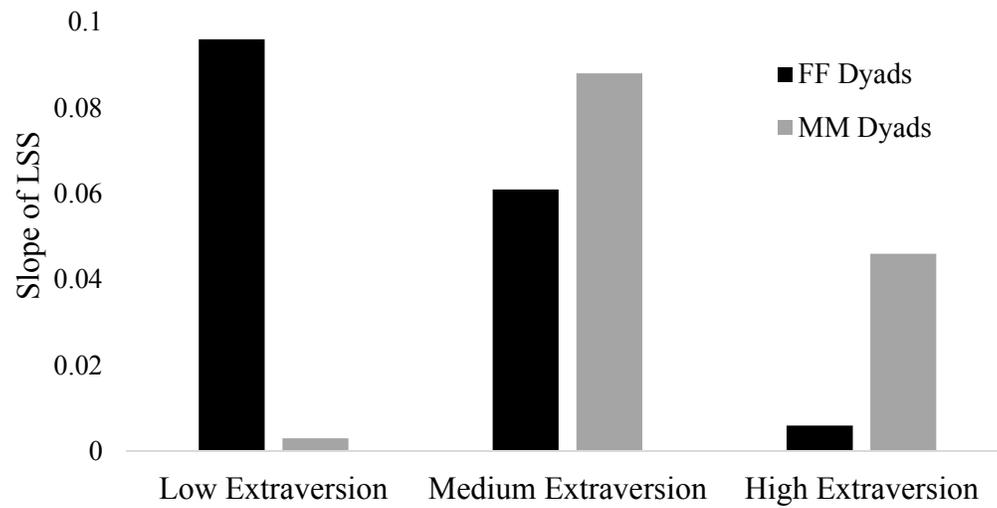


Figure 3. The slope of dyad-level LSS across time for low, medium, and high extraversion male-male and female-female dyads. A higher slope number denotes a steeper change in LSS across time.

Appendix A

Additional Correlations between LSS and Dyad-Level Behavioral and Thought/Feeling Content Measures that were Unique to the Present Study

<i>Measure</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>
LSS	—							
Word count	.61***	—						
Frequency of expressive gestures [^]	.54***	.67***	—					
Duration of expressive gestures [^]	.41**	.60***	.86***	—				
Frequency of positive affect	.47**	.41**	.59***	.49***	—			
Number of questions asked	.43**	.30	.23	.27	.22	—		
Percentage of other directed thought/feeling entries	-.42**	-.34*	-.40**	-.15	-.24	-.08	—	
Percentage of positive thought/feeling entries	.38**	.58***	.62***	.67***	.38**	.44**	-.27	—

Note. [^]Indicates that this variable was transformed using a square-root transformation (see the text for details).

* $p < .025$. ** $p < .01$. *** $p < .001$.

Appendix B

Factor Loadings of the Dyad-Level Behavioral Correlates of LSS in the Present Study

Factor	Correlates	Factor Loadings	Percent of Variance
1: Looking and acknowledging			25.28%
	Duration of mutual gazes [^]	0.94	
	Duration of directed gazes	0.91	
	Frequency of mutual gazes	0.73	
	Number of head nods	0.64	
	Frequency of directed gazes	0.46	
	Number of verbal reinforcers	0.43	
2: Gesturing			23.43%
	Duration of expressive gestures [^]	0.89	
	Frequency of expressive gestures [^]	0.85	
3: Talking and asking questions			16.72%
	Number of questions	0.87	
	Frequency of talking	0.79	
	Duration of talking	0.34	
4: Smiling and laughing			15.38%
	Frequency of positive affect	0.85	
	Duration of positive affect	0.84	
Total factor variance			80.81%

Appendix C
Pre-Interaction Questionnaire

1. In order to ensure that you receive credit for participation, please provide your first and last name. Any identifying information will be removed after credit is awarded and before data analysis begins.
2. What is your age? Please enter the number of years.
3. What is your gender?
 - a. Female
 - b. Male
 - c. Decline to Answer
4. What best reflects or represents your racial or ethnic background?
 - a. White/Anglo-American
 - b. Black/African-American
 - c. Hispanic/Latino
 - d. Asian
 - e. Native American or Alaskan Native
 - f. Native Hawaiian or Pacific Islander
 - g. Other/Multiracial
 - h. Decline to answer
5. Which of the following best describes your father's (or legal guardian's) level of education?
 - a. No high school diploma or GED
 - b. A high school diploma or GED
 - c. Some college or university education but no degree
 - d. A two-year degree from a community college or university
 - e. A four-year (bachelor's) degree from a college or university
 - f. A master's degree from a college or university
 - g. A doctoral (Ph.D.) degree from a college or university
 - h. Decline to answer
6. Which of the following best describes your mother's (or legal guardian's) level of education?
 - a. No high school diploma or GED
 - b. A high school diploma or GED
 - c. Some college or university education but no degree
 - d. A two-year degree from a community college or university
 - e. A four-year (bachelor's) degree from a college or university
 - f. A master's degree from a college or university
 - g. A doctoral (Ph.D.) degree from a college or university
 - h. Decline to answer
7. In which of the following ranges is your family's (or your legal guardian's) annual household income?
 - a. Less than \$30,000
 - b. \$30,000 to \$50,000
 - c. \$50,000 to \$70,000
 - d. \$70,000 to \$90,000

- e. \$90,000 to \$110,000
- f. \$110,000 to \$130,000
- g. More than \$130,000
- h. Decline to answer

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Using the scale below, please choose a response for each statement to indicate the extent to which you agree or disagree with that statement.

- A. Disagree strongly
- B. Disagree a little
- C. Neither agree nor disagree
- D. Agree a little
- E. Agree strongly

- 8. I see myself as someone who is talkative.
- 9. I see myself as someone who tends to find fault with others.
- 10. I see myself as someone who does a thorough job.
- 11. I see myself as a person who is depressed, blue.
- 12. I see myself as someone who is original, comes up with new ideas.
- 13. I see myself as someone who is reserved.
- 14. I see myself as someone who is helpful and unselfish with others.
- 15. I see myself as someone who can be somewhat careless.
- 16. I see myself as someone who is relaxed, handles stress well.
- 17. I see myself as someone who is curious about many different things.
- 18. I see myself as someone who is full of energy.
- 19. I see myself as someone who starts quarrels with others.
- 20. I see myself as someone who is a reliable worker.
- 21. I see myself as someone who can be tense.
- 22. I see myself as someone who is ingenious, a deep thinker.
- 23. I see myself as someone who generates a lot of enthusiasm.
- 24. I see myself as someone who has a forgiving nature.
- 25. I see myself as someone who tends to be organized.
- 26. I see myself as someone who worries a lot.
- 27. I see myself as someone who has an active imagination.
- 28. I see myself as someone who tends to be quiet.
- 29. I see myself as someone who is generally trusting.
- 30. I see myself as someone who tends to be lazy.
- 31. I see myself as someone who is emotionally stable, not easily upset.
- 32. I see myself as someone who is inventive.
- 33. I see myself as someone who has an assertive personality.
- 34. I see myself as someone who can be cold and aloof.
- 35. I see myself as someone who perseveres until the task is finished.
- 36. I see myself as someone who can be moody.
- 37. I see myself as someone who values artistic, aesthetic experiences.

38. I see myself as someone who is sometimes shy, inhibited.
39. I see myself as someone who is considerate and kind to almost everyone.
40. I see myself as someone who does things efficiently.
41. I see myself as someone who remains calm in tense situations.
42. I see myself as someone who prefers work that is routine.
43. I see myself as someone who is outgoing, sociable.
44. I see myself as someone who is sometimes rude to others.
45. I see myself as someone who makes plans and follows through with them.
46. I see myself as someone who gets nervous easily.
47. I see myself as someone who likes to reflect, play with ideas.
48. I see myself as someone who has few artistic interests.
49. I see myself as someone who likes to cooperate with others.
50. I see myself as someone who is easily distracted.
51. I see myself as someone who is sophisticated in art, music, or literature.

Appendix D

“Get to Know You” Session Procedure

You will be chatting with another participant, who is male (or female), via AOL Instant Messenger for a total of 18 minutes. Please chat with your partner in order to get acquainted with them, just as if you are getting to know someone in real life. You are free to discuss any topics you like. Once the 18 minutes is up, I will come back into the room and administer the Post-Interaction Questionnaire for you to complete regarding your experience chatting with your partner. Do you have any questions?

Appendix E

Please rate the following statements *based on the interaction you just had* with your partner using the scale below. Your partner is the person you were chatting with via instant messenger.

- A. Strongly disagree
 - B. Disagree
 - C. Neither disagree nor agree
 - D. Agree
 - E. Strongly agree
1. Please insert your name (this is for organizational purposes only and will be removed prior to data analysis).
 2. I am:
 - a. Participant #1
 - b. Participant #2
 3. My interaction partner was:
 - a. Participant #1
 - b. Participant #2
 4. The interaction seemed smooth, natural, and relaxed to me.
 5. The interaction seemed smooth, natural, and relaxed to my partner.
 6. The interaction seemed awkward, forced, and strained to me.
 7. The interaction seemed awkward, forced, and strained to my partner.
 8. I felt accepted and respected by my partner.
 9. My partner felt accepted and respected by me.
 10. I felt put down, patronized, or rejected by my partner.
 11. My partner felt put down, patronized, or rejected by me.
 12. I felt a connection with my partner.
 13. My partner felt a connection with me.
 14. I liked my partner.
 15. My partner liked me.
 16. I would like to interaction more with my partner in the future.
 17. My partner would like to interaction more with me in the future.
 18. I understood my partner.
 19. My partner understood me.
 20. I enjoyed the interaction with my partner.
 21. My partner enjoyed the interaction with me.
 22. I felt comfortable interacting with my partner.
 23. My partner felt comfortable interaction with me.

Appendix F

Debriefing Statement

Thank you for participating in this study. This study's main focus is actually about latent semantic similarity, or LSS. LSS is an overall measure of how similar two blocks of texts are to each other and is regarded as a measure to which interaction partners achieve a "common-ground understanding". In this study, we are interested in analyzing LSS among interaction partners who are involved in online "get to know you" sessions. In other words, we are interested in how interaction partners develop LSS through written correspondence (in this case, through online chatting) as opposed to face-to-face interactions.

We had you chat with another participant in order to obtain both you and your partner's writing samples. The writing samples will be analyzed using a program called Latent Semantic Analysis (<http://lsa.colorado.edu>) in order to assess the level of LSS between you and your partner. As a result, your chat log will be saved and used for analysis. All identifiers will be removed before data analysis, and all information will be kept completely confidential. If you are not willing to grant us permission to use your chat log, it will be deleted and will not be used in this study with no penalty. If you are willing to grant us permission to use your chat log, you will be given an authorization form to sign that will allow the researchers of this study to use your chat log. You also have the right to retract your consent to participate now that you have been informed of the study's true purpose. Please indicate your decision to the experimenter now.

If needed, you may contact UTA Counseling Services at 817-272-3671. Please refrain from discussing this study with other individuals. If you have any other questions, please feel free to contact Vivian Ta at Vivian.Ta@mavs.uta.edu. Again, thank you for your participation in this study.

Appendix G

Authorization Form

"I, _____, authorize the researchers for the study entitled "Study of Online "Get to Know You" Sessions" to use my writing samples (i.e., chat logs) for further analysis procedures using Latent Semantic Analysis. I understand that my identity will be protected during this process and that any identifiers will be removed."

Printed Name

Signature

Date

Appendix H

A multiple regression analysis was used to determine if dyad-level LSS, computed over the entire interaction period, is a significant predictor of the degree to which the partners perceived a positive emotional connection to each other (i.e., Factor 2). The results of the regression model used to test this hypothesis were not significant, $F(1, 119) = .00, p = 1.00, R^2 = .00$. Additional regression analyses were used to determine if (1) LSS significantly predicted any of the Factor 2 items individually, and (2) if the gender composition of the dyad (M-M or F-F) was a significant moderating variable. The results of these analyses showed that LSS did not significantly predict any of the individual items that make up Factor 2.

A mediation analysis was used to determine if dyad-level LSS mediated the relationship between dyad-level agreeableness and the dyad-level measure of the degree to which the partners perceived a positive emotional connection to each other (i.e., Factor 2). The total effect of agreeableness on the Factor 2 was not significant, $b = .07, SE = .09, t(119) = .80, p = .43, 95\% CI[-.11, .25]$. The dyad-level effect of Factor 2 on LSS was not significant, $b = -.01, SE = .65, t(119) = -.02, p = .98, 95\% CI[-1.31, 1.28]$. The effect of agreeableness on LSS was not significant, $b = .00, SE = .01, t(119) = .29, p = .77, 95\% CI[-.02, .03]$. The direct effect of agreeableness on Factor 2, after controlling for the mediator, was not significant, $b = .07, SE = .09, t(119) = .80, p = .43, 95\% CI[-.11, .25]$.

To determine whether the indirect effect of LSS on agreeableness and Factor 2 was significant, a bias-corrected bootstrapping procedure was used with 1000 samples. The results showed that LSS was not a significant mediator of the relationship between agreeableness and Factor 2, Effect = $-.0001, SE = .09, 95\% CI[-.02, .02]$. The Sobel test of the indirect effect generally matched the result that was obtained using the bootstrapping procedure. That is, LSS

was not a significant mediator, Effect = $-.0001$, $SE = .01$, $z = -.01$, $p = .99$. It seems clear that LSS did not mediate the relationship between the dyad-level measures of agreeableness and Factor 2.

Another mediation analysis was conducted to determine whether dyad-level LSS mediated the relationship between dyad-level extraversion and Factor 2. The total effect of extraversion on Factor 2 was significant, $b = .20$, $SE = .08$, $t(119) = 2.40$, $p = .02$, 95% $CI[.03, .36]$. However, the effect of extraversion on LSS was not significant, $b = .02$, $SE = .01$, $t(119) = 1.34$, $p = .18$, 95% $CI[-.01, .04]$. The effect of LSS on Factor 2 was not significant, $b = -.19$, $SE = .64$, $t(119) = -.29$, $p = .77$, 95% $CI[-1.47, 1.09]$. The direct effect of extraversion on Factor 2, after controlling for the mediator, was significant, $b = .20$, $SE = .08$, $t(119) = 2.41$, $p = .02$, 95% $CI[.04, .37]$.

To determine whether the indirect effect of LSS on extraversion and Factor 2 was significant, a bias-corrected bootstrapping procedure was used with 1000 samples. The results revealed that LSS was not a significant mediator of the relationship between extraversion and Factor 2, Effect = $-.003$, $SE = .01$, 95% $CI[-.06, .01]$. The Sobel test of the indirect effect generally matched the results that were obtained using the bootstrapping procedure. It also revealed that LSS was not a significant mediator of the relationship between extraversion and Factor 2, Effect = $-.003$, $SE = .01$, $z = -.23$, $p = .82$.

A third and final mediation analysis was conducted to determine whether dyad-level LSS mediated the relationship between dyad-level openness to experience and Factor 2. The total effect of openness to experience on Factor 2 was not significant, $b = -.11$, $SE = .10$, $t(119) = -1.02$, $p = .31$, 95% $CI[-.31, .10]$. The effect of openness to experience on LSS was not significant either, $b = -.004$, $SE = .01$, $t(119) = -.28$, $p = .78$, 95% $CI[-.03, .03]$. The relationship between the

mediator and Factor 2 was not significant, $b = -.10$, $SE = .11$, $t(119) = -1.02$, $p = .31$, 95% $CI[-.32, .10]$. Finally, the direct effect of openness to experience on Factor 2, after controlling for the mediator, was not significant, $b = -.10$, $SE = .11$, $t(119) = -1.02$, $p = .31$, 95% $CI[-.32, .10]$.

To determine whether the indirect effect of LSS on openness to experience and Factor 2 was significant, a bias-corrected bootstrapping procedure was used with 1000 samples. The results revealed that LSS was not a significant mediator of the relationship between openness to experience and Factor 2, Effect = .001, $SE = .01$, 95% $CI[-.02, .02]$. The Sobel test of the indirect effect generally matched the results that were obtained using the bootstrapping procedure. That is, LSS was not a significant mediator of the relationship between openness to experience and Factor 2, Effect = .0001, $SE = .01$, $z = .01$, $p = .99$.

Appendix I

The slope of each dyad's LSS scores across time was calculated by solving the equation of the regression line, $y = a + bx$, for each dyad, where b is the slope. For example, if a dyad scored an LSS of .60 in Time 1, .59 in Time 2, and .58 in Time 3:

	X	Y	X*Y	X*X
	1	.6	.6	1
	2	.59	1.18	4
	3	.58	1.74	9
Σ	6	1.77	3.52	14

Where the values of X represent the time periods of each dyad's interaction.

Next, use each of the following equations to find a and b :

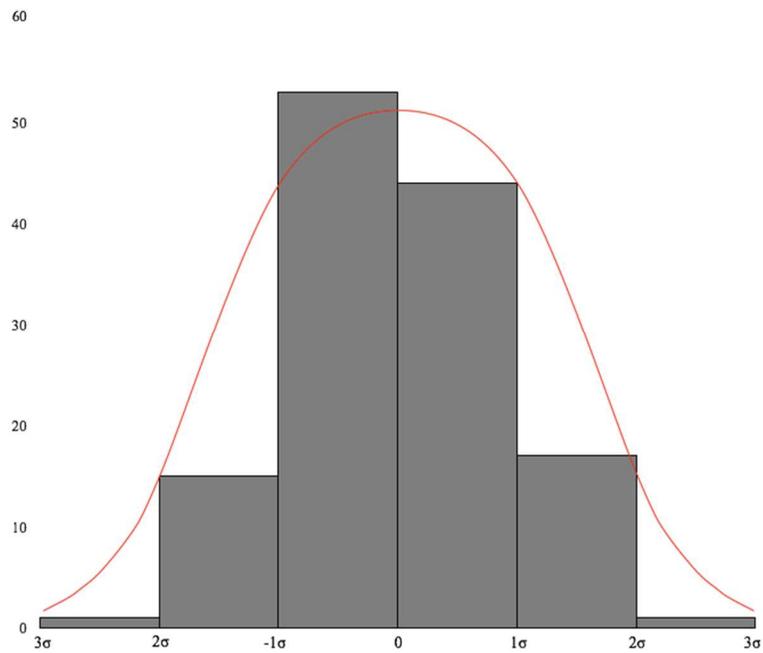
$$a = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{n \cdot \sum X^2 - (\sum X)^2} = \frac{1.77 * 14 - 6 * 3.52}{3 * 14 - 36} = .61$$

$$b = \frac{n \cdot \sum XY - \sum X \cdot \sum Y}{n \cdot \sum X^2 - (\sum X)^2} = \frac{3 * 3.52 - 6 * 1.77}{3 * 14 - 36} = -.01$$

Then, substitute a and b in the regression equation formula, $y = a + bx$

$y = .61 - .01x$, where $-.01$ is the slope of this dyad's LSS across time.

Appendix J



The frequency breakdown of dyad-level extraversion level within the ranges of a normal distribution. $M = 3.20$. $SD = .56$