A Multimodal Approach to Measuring Behavior in Social Interactions

Kent C. Halverson, Ian Yohai, Jonathan Lansey
Shawn Weil
Aptima, Inc.
Woburn, MA

Brian Lande
University of California at Santa Cruz,
Department of Computer Science
Santa Cruz, CA

Daniel Olguin Olguin, Maegen Demko
Sociometric Solutions, Inc.
Boston, MA

ABSTRACT

Members of the Armed Forces are routinely faced with situations in which social interaction (e.g., cross-cultural, cross-linguistic) has both strategic and tactical implications. In these situations, anecdotal evidence indicates that some military members have greater outcome success (e.g., fewer subsequent fire-fights or IED events) when interacting with local citizenry than others, regardless of their individual language capabilities. In these increasingly common social contexts, where language barriers are often present, other modes of communication rise in importance. Nonverbal communication modes (e.g., body movement, physical proximity, gestures) and paralinguistic speech features (e.g., volume, pitch, turn-taking behavior), provide cues that carry significant meaning which can enhance cooperative interaction and build trust and rapport. This suggests that social interaction inherently involves latent communication features that are not easily discernible by an observer, and are similarly difficult for researchers to measure. However, these subtle "honest signals" (Pentland, 2010) are now measureable through the use of wearable sensors that capture body movement, proximity data, and speech features such as volume and pitch. The purpose of this paper is to introduce and test a conceptual model that captures the multimodal components of social interactions. An empirically validated model of social interactions will provide critical social skills content that will enable both training and evaluation. Data was collected using wearable sensors and observer assessments in a military training program designed to teach soldier leaders to be adaptive in unfamiliar environments. Training scenarios included interactions among soldiers and ethnic role players (who spoke languages other than English) acting as local citizens. Multimodal data were used to identify key features of social interaction that correlate with outcome measures. For example, body movement measures from the wearable sensors were found to be correlated with observer ratings of engagement. .

ABOUT THE AUTHORS

Kent C. Halverson, PhD is a Senior Scientist at Aptima, Inc. with experience in organizational analysis, leadership development, training evaluation, and social network analysis.

Ian Yohai, PhD is a Senior Quantitative Social Scientist with Aptima, Inc. with experience in statistical methods for language analysis, missing data, causal inference, regression modeling, and the analysis of survey data.

Jonathan Lansey is a Scientist with Aptima, Inc. with experience in quantitative analysis and modeling.

Shawn Weil, PhD is the Director of Aptima's Analysis, Modeling and Simulation Division. Research interests include individual and organizational performance, C2, and advanced training methods.

Brian Lande, PhD is a sociologist at the Center for Games and Playable Media, at University of California Santa Cruz, with experience in social psychology, social theory, ethnographic methods, and criminology.

Daniel Olguin Olguin, PhD is a Co-Founder and the Chief Technology Officer of Sociometric Solutions, Inc. with experience in wearable computing, human behavior sensing, pattern recognition, and organizational engineering.

Maegen Demko is a data analyst with Sociometric Solutions, Inc. with experience in quantitative analysis.

A Multimodal Approach to Measuring Behavior in Social Interactions

Kent C. Halverson, Ian Yohai, Jonathan Lansey, Shawn Weil
Aptima, Inc.
Woburn, MA

Brian Lande
University of California at Santa Cruz, Department of Computer Science
Santa Cruz, CA

Daniel Olguin Olguin, Maegen Demko,
Sociometric Solutions, Inc.
Boston, MA

INTRODUCTION

Members of the armed forces are stationed in countries around the world and interact with citizens from different cultures who speak different languages, making communication between military members and local citizens challenging. However, military member experiences and anecdotes have indicated that some military members have greater outcome success (e.g., success in intelligence gathering, advising, fewer subsequent fire-fights or IEDs) in interacting with local citizenry than others, regardless of their individual verbal second language capabilities. In these unique social interaction contexts, where language barriers are present, other modes of communication become much more important. Nonverbal communication modes, such as facial expressions, body position, gaze, and gestures, and voice volume and intonation, provide subtle cues that carry information for coordinating attention, activity, and signaling emotional states and relational states such as potential for trust. The purpose of this paper is to introduce and test a conceptual measurement model that captures the salient components of social interactions between strangers. We want to know how and why some members of the armed services succeed and others fail during intercultural contacts. We believe that part of the answer lays in styles of engagement – how people manage themselves and others during intercultural encounters. Data was collected in a military training program designed to teach soldiers to be adaptive in unfamiliar environments, which inherently includes conversations with strangers. The data provided partial support for hypotheses suggesting a relationship between soldier engagement and dyadic rapport. In addition, the data suggest that a soldier's activity and mirroring are related to their engagement level.

LITERATURE REVIEW

People are frequently put in situations in which they are required to engage in a social interaction with unfamiliar people, whether it is a shopper talking with a sales clerk or a soldier negotiating a civilian's passage through a roadblock. These interactions are unique in that both individuals have little or no past experience with one another, may not share any common ground (similar expectations, historical experiences, stories, etc.), and have access only to situationally perceptible physical appearance, clothing, and location (i.e., context of the interaction). In other words, in many operational contexts, members of the armed services do not come into situations with a "pre-coordinated" stock of decontextualized linguistic and cultural knowledge (e.g., rules, beliefs, scripts, norms, values) to draw on. In fact, it is not at all clear that service members require such a stock of knowledge to be successful. Rapid deployments of service members to linguistically and culturally diverse areas means that it is not feasible to turn service members into lay anthropologists. Rather, the need to collaborate with different cultural groups requires the ability to learn rapidly about unfamiliar social settings and the agility to adapt. Until recently, the military was slow to recognize that successful cross cultural encounters required more than just second language competence and anthropological knowledge of "target" societies (Selmeski, 2007; Caliguiuri, Noe, Nolan, Ryan, & Drasgow, 2011); Rasmussen & Sieck, 2012).

That second language or cultural diversity instruction has not solved the difficulties in intercultural social interaction should not be a surprise since there is good reason to believe that the principles of social interaction are independent and indeed ontogenetically and phylogenetically developed prior to language (Levinson, 2006; Tomesello, 2008). Researchers (Gumprez, 1982; Tannen, 1985; Lakoff, 1976; Enfield & Levinson, 2006) have also long shown that having linguistic competence does not mean that cross-cultural encounters will go well since encounters require successful *interaction* and not just successful *speech*. Hence in Gumprez's classic studies of intercultural interaction, he was able to show that native English speakers who grew up in different social environments (England and India) enter into cross-cultural conflict because of *how* language is used, and its interactive context, goals, and task constraints lead to misunderstandings. Common ground (Clark, 1996), contextualization cues (Gumprez, 1982), shared interaction frames, participant structures (Goffman, 1981), concepts of social self and other, socialized postures, and joint perceptions all figure into the understanding or misunderstanding of face-to-face encounters.

What are participants to do if they share little in common? Considering the lack of interpretable information available in intercultural or zero-acquaintance interactions, the participants are likely to rely on initial impressions and swift trust (e.g., Robert, Dennis, & Hung, 2009), followed by a continuous stream of real-time cues that emerge and accumulate as the social interaction progresses. In these zero-acquaintance encounters, common ground is rapidly built, contextualization cues are identified, and expectancies of self and other are established (Gibbs & Clark, 2012). But these are accomplished in and through the interaction, not by reference to external knowledge. The behavioral cues, verbal and nonverbal, displayed by both parties are used to identify patterns from the social interaction, inform expectations, and guide subsequent behavioral responses. While some behaviors (e.g., a smile) are language-like and can be easily understood and coded by the receiver, other behaviors may only offer a "probabilistic association between behavior and meaning" (Grammer, Honda & Schmitt, 1999, p. 489). During the course of a social interaction, these probabilistic behaviors can augment or contradict the message conveyed by other more objective communication units. For example, if someone says she is really, really excited about an idea, but her posture is hunched, her face is blank, and she displays few, if any, dynamic hand gestures, it is likely that the other person might question the authenticity of the verbal communication since the accompanying nonverbal signals emitted by the sender contradict the verbal content. In these types of cases, it is relatively easy for the receiver, or even an external observer, to notice the mixed messages and question the authenticity of the communication. However, communication signals can often be much more subtle or seemingly undetectable, such as in the case of when people get an uncomfortable feeling from a person, despite the lack of any apparent questionable behavioral displays. This suggests that communication inherently involves hidden communication modes that are not easily discernible by an observer and similarly difficult for researchers to measure.

There are several practical measurement issues that increase the difficulty of accurately observing social interactions. First, as previously mentioned, communication cues can be very subtle and difficult to detect using traditional measurement instruments. Ethnographies provide rich qualitative data that provide valuable heuristic perspectives. However, this type of data collection relies on third-party observation and is inherently prone to rater bias. Although the use of multiple raters can increase reliability of ethnographies, there is still the possibility that observer training encourages specific cognitive frames that could result in priming, or bias toward coding specific communication modes, perhaps missing other relevant cues. In addition, some subtle communication cues, such as slight verbal intonations or facial gestures, can be so fleeting that it if the observer is not listening closely or looking directly at the subject, the behavior could be missed. The process used by human raters usually involves observing behavior and then suspending observation to record the information. It is during this recording period when human raters are likely to miss important social behaviors demonstrated by the subjects. However, ethnographers frequently use video recording that allows them to rate behavior in a post hoc manner when they can stop or rewind the video, which reduces the likelihood of missing fleeting cues. While this procedure can increase the reliability of ethnographic data, it is a labor and time intensive process that cannot provide reasonably quick feedback for analysis. Finally, social interactions inherently involve a continuous, reciprocating stream of large amounts of information that could exceed an observer's processing capacity. Even with multiple raters and video technology, ethnographies are still limited in the amount of total data that can be collected and coded. Considering this data challenge, it is necessary to leverage technology and use measurement instruments designed to collect large quantities of continuous data flows streaming from social interactions.

Computer science researchers and experts in social signal processing have shown that it is possible to detect conversational engagement from speech and body movement features. For instance, Yu, Aoki, and Woodruff (2004) used machine learning techniques (i.e., support vector machines [SVM] and hidden Markov models [HMM]) to classify user emotions and estimate engagement levels. Their approach employed SVM classifiers that used acoustic features as inputs to predict discrete emotional states of users. Those emotional states were then utilized as inputs to an HMM to model user emotional states and engagement in conversation as a dynamic, continuous process. The authors used acoustic features, such as volume and pitch, which are similar to those captured by the Sociometric Badges (Olguin Olguin, 2007) used in our study and described below. The emotional state recognition accuracy was up to 75% and the continuous engagement detection accuracy (on a 1 to 5 scale) was 61% when incorporating the HMM (Yu et al., 2004). To our knowledge, this is probably one of the earlier works that attempts to estimate continuous user engagement in spoken conversations.

Computational linguistic researchers have also studied the *mirroring* or *entrainment* phenomenon that occurs in dialog partners. Levitan, Gravano, and Hirschberg (2001) show that speaking partners tend to mirror back-channel preceding cues as evidence of continued engagement, that this similarity increases over the course of a dialogue; and that this similarity is associated with measures of dialogue coordination and task success. They identified five cues that tend to be present in speech preceding backchannels: pitch, derivative of pitch, volume, speaking segment length and voice quality (i.e., noise-to-harmonics ratio). Their mirroring features were negatively correlated with latency and the number of interruptions, both used as a measure of dialogue coordination (Levitan et al., 2001).

Sun, Nijholt, Truong, and Pantic (2011) discuss how mimicry behavior is an important indicator of cooperativeness and empathy during conversation and how previous studies in social psychology report that the more mimicry is observed, the more smoothly the interaction is perceived: mimicry enhances resonance, creates rapport and affiliation suggesting that it serves to strengthen social bonds. The authors suggest using non-verbal acoustic features such as pitch, vocal intensity (i.e. volume), and speech rate (Sun et al., 2011).

HYPOTHESES

Up to this point, the focus of the literature review has been on communication processes, presenting an argument that humans emit subtle or hidden verbal and non-verbal communication cues. While understanding communication processes is important, it is also important to understand how those processes impact meaningful outcomes of social interactions. One construct that has been suggested as being related to favorable social outcomes is rapport. Rapport can be defined as a "relation marked by harmony, conformity, accord, or affinity," (Merriam-Webster, 2013). The relationship between rapport and favorable social outcomes seems intuitive, and there is evidence of the benefits of rapport such as witness recall accuracy (Vallano & Compo, 2011), child performance (Feldman & Sullivan, 1971; Exner, 1966), customer satisfaction and loyalty (Delcourt, Gremler, van Riel, & van Brigelen, 2013), trust (Ross & Wieland, 1996), respect (Beach et al, 2006), and cooperation (Drolet & Morris, 2000).

Tickle-Degnen and Rosenthal (1990) suggest that rapport has three components: mutual attentiveness, described as "focused and cohesive interaction" (p. 286); positivity, described as "friendliness and caring" (p. 286); and coordination, described as "balance, harmony, and in sync" (p.286). Mutual attentiveness seems very similar to engagement or active listening. Those who are more engaged in a social interaction and display active listening behaviors might be more likely to develop rapport. This leads to the first hypothesis:

H1: Soldier engagement is positively related to rapport between soldier and role player.

As discussed previously, the subtlety of communication signals can make it difficult for external observers to determine the extent to which individuals are truly engaged in a social interaction. Overt active listening techniques such as acknowledging responses with "uh huh" or asking follow-up questions (Fisher & Geiselman, 1992) can be easily coded. However, other honest signals (Pentland, 2010) are present that might provide supplemental, and perhaps more valid, information about engagement. When an individual is engaged in a social interaction, it is likely the individual will maintain a focus on the other person by discontinuing other actions (e.g., movements) and positioning their body (i.e., posture) in a way that maximizes sensory perception. This leads to hypothesis 2a:

H2a: Soldier engagement is negatively correlated with soldier activity levels (observed in body movement, posture, volume and tone of voice).

Research has shown that behavioral mirroring enhances social affiliation and relationship building (Lakin, Jefferis, Cheng, & Chartrand, 2003; Wang & Hamilton, 2012). Lakin et al. (2003) suggest that mirroring is an evolutionary unconscious behavior that allowed access to more information useful for survival. Similarly, in order for soldiers to obtain useful information from a stranger, the soldier is likely to mirror the behaviors of the other interlocutor to put them into a deeper state of engagement. This leads to hypothesis 2b:

H2b: Soldier engagement is positively correlated with soldier mirroring of the role players (observed in body movement, posture, volume and tone of voice).

METHODS

Sample description

This study took place at the Asymmetric Warfare Group's Adaptive Warfare Leadership Program that offers adaptability training designed for soldiers to confront asymmetric threats. The soldiers are trained in a variety of advanced war-fighting skills (e.g., repelling, weapons) as well as adaptability (e.g., Pulakos, Arad, Donovan & Plamondon, 2000). The evaluation phase of the training involves social interactions with citizens (i.e., ethnic role players). The soldiers are required to interact with the role players in order to meet their objective. The role players are first-generation immigrants from India who speak their native languages, thus requiring the soldiers to use interpreters in the social interactions. Although the interpreter role may seem to mediate the communication between a soldier and citizen, all role players are also fluent in English, allowing them to hear the soldier's verbal communication and process it without any interpretation effect.

The primary participants in this study were U.S. Army soldiers in a training course. The soldiers ranged from E-3 through E-7 in the enlisted ranks and O-1 through O-4 in the officer ranks. For each scenario, one soldier was assigned the role of the patrol leader who was the primary soldier interacting with the role players. Social interaction data were collected from 12 AWALP training scenarios. The scenarios involved interviews with role players in the following roles: village elders, police, hotel owners, farmers, construction crew (i.e., IED hole diggers). In all of the scenarios, the social interactions between the soldiers and role players were calm and civil, meaning that social interactions did not involve overt coercion or use of force.

Procedure

All participants in the study, including the soldiers, the interpreters, and the citizen role players, wore Sociometric Badges (Sociometric Solutions, 2013) to collect voice, motion and proximity data. Sociometric Badges are a relatively new type of measurement instrument that can capture individual behavior (e.g., body movement and speech patterns) as well as social behavior (e.g., face-to-face interaction, proximity and conversational turn-taking patterns). This behavioral measurement tool was originally developed at MIT (Olguin-Olguin, 2007) with the goal of automatically capturing and quantifying subtle "honest signals" (Pentland, 2010), innate and difficult-to-fake non-verbal signals unconsciously displayed during social interactions. Using this sensor platform, Olguin-Olguin and Pentland (2010) further developed a sensor-based organizational design and engineering framework that combines behavioral sensor data with e-mail, surveys, and performance data—among other types of information—with the goal of providing feedback and designing organizational interventions that have a positive impact on relevant organizational outcomes. Although the badges don't record the content of the communication, they do record such things as volume, pitch, and turn-taking between participants. The badges are also equipped with accelerometers and Bluetooth to measure motion and proximity, respectively. The researchers issued the badges to the participants each morning prior to the start of their scenario exercise.

The latest version of the Sociometric Badges used in this study is capable of measuring the following social interaction components: face-to-face interaction detection: proximity detection to other people wearing badges and to base stations; speech social signal analysis: volume and pitch; conversational social signal analysis (e.g., turn-taking statistics); and, body movement social signal analysis (energy and posture). These components form the basis for measuring the following variables (activity and mimicry) that will be described in more detail later in the paper.

During the social interactions, researchers followed the soldier assigned to be the patrol leader and observed the social interaction using a tablet-based evaluation tool called SPOTLITE. Originally developed under contract to the United States Air Force Research Laboratory (AFRL), SPOTLITE is a handheld observer-based performance assessment tool with video recording capability as well as touch-screen assessment measures that allow external observers to assess performance and other behaviors in real-time and well as post hoc. SPOTLITE automatically timestamps all assessments, facilitating integration with the information collected by the Sociometric Badges and providing a multimodal measurement system for observing social interaction.

Measures

A wide variety of performance measures were used during and surrounding the social interactions to collect meaningful data on social interactions. In order to gain multiple perspectives on the interaction, the measures were collected by the researchers functioning as "impartial" observers, in addition to gathering self-report data from the soldiers and the role players immediately following the interaction.

Rapport. This construct was measured from three perspectives, by the external observer (two raters), the soldier (self-report), and the citizen (self-reports from one or two citizen role players), to allow triangulation and prevent common source variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). The two Likert items used were "The soldier and citizen(s) demonstrated mutual rapport (i.e., friendly social interaction)" and "The interaction between the soldier and citizen(s) was pleasant and smooth."

Soldier Engagement. This construct was measured by external observers (i.e., researchers) using two different instruments. The first instrument was a real-time engagement scale adjacent to the video display on the SPOTLITE evaluation tool. This scale consisted of a five-point scale with the top of the scale colored green to indicate full engagement and the bottom of the scale red to indicate full disengagement. The anchors used are described in Table 1 below.

Table 1. Soldier Engagement Rating Anchors

	Disengaged (1)	Passive (3)	Active (5)
Listening	• Provides no response cues to acknowledge role player communication	• Provides occasional response cues to acknowledge communication	• Provides consistent cues to acknowledge communication • Smiling, hand gestures, etc.
Observing	• Constant wandering gaze • Little/no eye contact • Distracted by other events	• Inconsistent eye contact • Occasionally looking over shoulder, down at notebook.	• Consistent eye contact • Focused, not distracted • Seldom looks at his notebook
Body Positioning	• Positions body at ≥ 90° • Uncomfortable distance	• Positions body ≤ 90° • Comfortable distance	• Positions body parallel • Relatively close proximity

At the end of the scenario, the external observer evaluated engagement using a three-item scale targeting visual engagement, active listening, and effort. The three items were evaluated using a five-point Likert scale: "The soldier maintained consistent eye contact with the citizen(s)."; "The soldier actively listened and was receptive to the citizen(s)."; and, "The soldier demonstrated effort to engage with the citizen(s)."

Activity. This is an individual-level construct captured by the Sociometric Badges that is an aggregated measure of the magnitude of body movement, posture, volume and pitch of voice. The measure captured changes in behavior and voice and was captured at a 0.5-second resolution. The possible scores ranged from 0 to 1 (i.e., the normalized volume as captured by the badge's microphone or the body movement energy in units of gravity (g) as captured by the accelerometer). If a soldier sat completely still and was completely silent, the activity score would essentially be zero.

Mirroring. This is a dyadic construct captured by the badges that compared body movement, posture, volume, and pitch of voice between the soldier and the ethnic role player. The possible scores ranged from 0 to 1. Mirroring was computed every 0.5 seconds using a 60-second moving window across which the scores were averaged. The more similar the body behaviors and non-verbal features between the participants, the higher the score.

Analysis

In this study we conducted two analyses: a cross sectional correlational analysis to evaluate post hoc survey data and a time series analysis to evaluate data obtained from SPOTLITE and the Sociometric Badges.

The correlational analysis compared data from the engagement and rapport constructs. Since both constructs were measured from different sources, several correlations were computed. For instance, rapport was measured by the external observer(s), the soldier, and the role player(s), engagement was measured by the external observer and listening was measured by the role player(s). The number of external observers and role players (i.e., either one or two) depended on the scenario.

The time series analysis compared two types of honest signals captured by the Sociometric Badges (activity and mirroring) and the observer-based engagement construct. Observer-based engagement ratings were collected continuously yet intermittently, meaning that new ratings were entered only when the changes in soldier engagement were observed. In order to calculate correlations with the Sociometric Badge data, we extrapolated the observer-based engagement scores to match the 0.5-second resolution rate of the badge data for each of the 13 scenarios. It is important to mention that even though the number of scenarios is 13 (see scenario labels in Figure 1), the number of data points used to compute the correlation ranges is much larger. The range of data points obtained from Sociometric Badges across all scenarios was 680 for the shortest scenario lasting approximately five minutes to 6800 for the longest scenario lasting approximately 50 minutes. In other words, when a correlation is observed between a badge feature and the engagement rating, the data contributing to that correlation originates from potentially thousands of data points over the duration of a scenario.

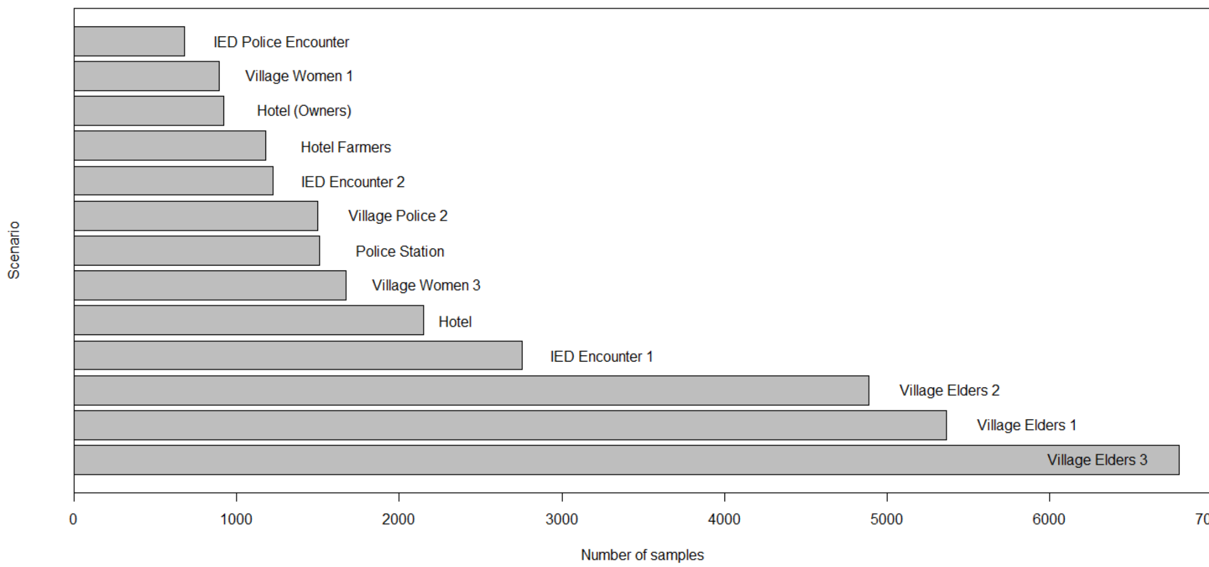

Figure 1. Time Series Sample Sizes by Scenario

RESULTS

The descriptive statistics for the correlational analysis are provided in Table 2.

Table 2. Descriptive Statistics of Engagement and Rapport Variables

Variable	Source	N	Mean	Standard Deviation
Engagement	External Observer(s)	15	2.87	1.64
Listening	Role Player(s)	11	3.50	1.22
Rapport	External Observer(s)	15	3.18	1.26
Rapport	Soldier	10	4.30	2.25
Rapport	Role Player(s)	10	3.75	0.59

For those scenarios when multiple raters were available, correlations were calculated to determine the interrater reliability. Table 3 shows the level of agreement between the two raters (External Observers or Role Players). There was very high agreement between the external observers' ratings and between the Role Players' ratings. It was not surprising to see a negative correlation between the Soldier and the Role Players, especially considering that the soldiers have not received any formal social interaction training as part of the AWALP curriculum, and the participants may have perceived the interaction quite differently.

Table 3. Interrater Reliabilities—Bivariate Correlations

Variables	Comparing External Observer Responses	Comparing Role Player Responses	Comparing Soldier and Aggregated Role Player Responses
Engagement	0.90 (n=5)	N/A	N/A
Listening	N/A	0.63 (n=9)	N/A
Rapport	0.90 (n=8)	0.08 (n=9)	-0.18 (n=8)

Note: N/A indicates that only one rater scored the variable.

The set of correlations between engagement and rapport varied widely as shown in Table 4. The largest correlation was observed when the two variables were both measured by the external observers, which is not surprising considering the potential for common source variance (Podsakoff et al., 2003). The remaining correlations ranged from 0.17 to 0.29, suggesting a weak, but positive relationship between engagement and rapport, partially supporting hypothesis H1.

Table 4. Correlations between Engagement and Rapport

	Engagement (source)	Listening (source)
Rapport (source)	External Observer(s)	Role Player(s)
Role Player(s)	0.17	0.26
Soldier	0.20	0.23
External Observer(s)	0.82	0.29

The time series analysis consisted of evaluating the correlations between the observer-based measure of engagement and data obtained from the Sociometric Badges, consisting of both activity (pitch, volume, posture, and body movement) and mirroring (pitch, volume, posture activity, and body movement). We calculated the set of four correlations for each scenario and then constructed box plots of those correlations in Figures 2 and 3. Figure 2 displays the box plots of the correlations with the individual-level data while Figure 3 displays the dyadic (i.e., mirroring) data. Figure 2 shows that the posture and body movement data were more negatively correlated with the observer-based measure of engagement. For the posture activity variable, we observed negative correlations in 11 of the 13 scenarios, nine of which were significant at $p<.05$. A similar result was observed for the body movement activity variable, with same number of negative correlations.

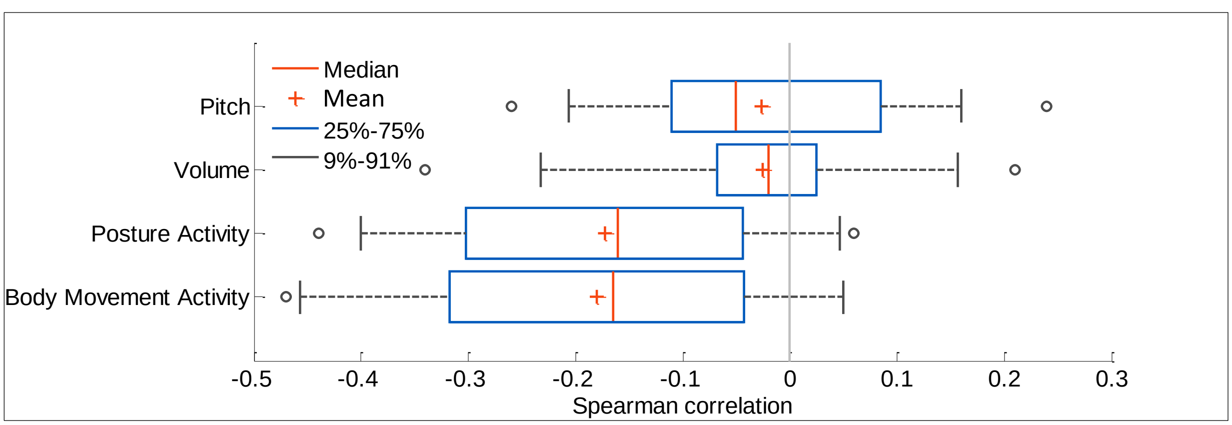

Figure 2. Box Plots of Time Series Analyses of Correlations

This suggests that the higher the soldier's body movement energy and the more changes in posture, the lower the engagement ratings. When looking at pitch and volume badge features, however, the results did not show a clear pattern. Pitch activity generated negative correlations in only eight of 13 scenarios (five of the negative correlations were significant at $p<.05$). Volume activity generated negative correlations in only seven of 13 correlations (three of the negative correlations significant at $p<.05$). Considering this, hypothesis H2a was partially supported.

The results for the mirroring badge features for hypothesis H2b were similarly mixed as shown in Figure 2. Note that the hypothesized direction of the correlation was positive, suggesting that mirroring is related to engagement. For pitch mirroring (11 scenarios) and volume mirroring (nine scenarios), the observed correlations with engagement were consistently positive and statistically significant ($p<.05$) across all scenarios, except the pitch feature in one scenario with a very small negative correlation of -.03. However, the body movement mirroring feature was negatively correlated with engagement in six of 11 scenarios (five of the negative correlations were significant at $p<.05$) and the posture mirroring feature was negatively correlated in six of 12 scenarios (five of the negative correlations were significant at $p<.05$). Although the volume and pitch mirroring features indicate a positive relationship with engagement, the body movement mirroring features were less consistent, and therefore, only partially supporting hypothesis H2b.

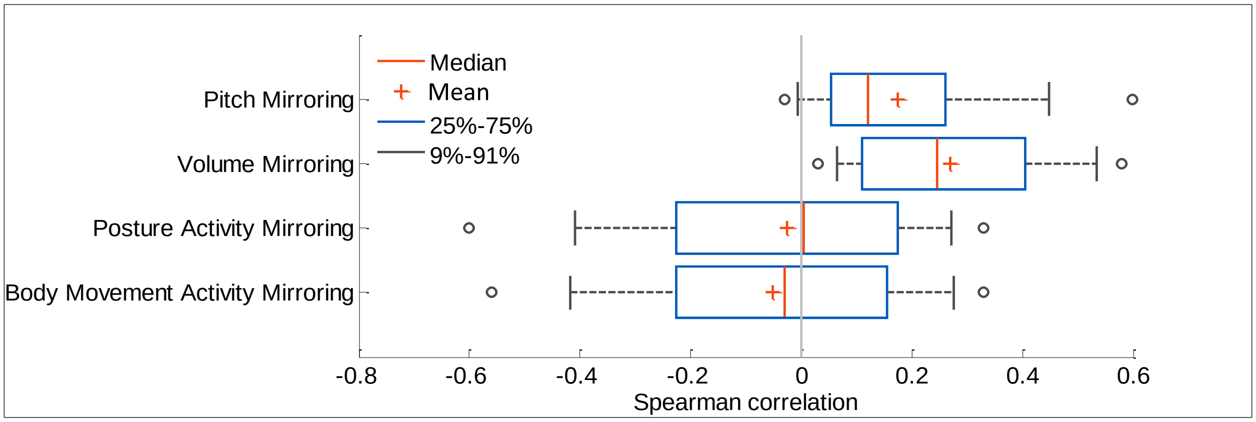

Figure 3. Box Plots of Time Series Analyses of Correlations

DISCUSSION

This purpose of this study was to begin investigating hidden behavioral signals within social interactions and to experiment with new measurement technologies to capture those subtleties. The results suggest that there might be measurable patterns of behavior that are difficult for humans to perceive, but can be captured by new measurement instruments. The data obtained from the study indicated that soldiers who are more engaged in social interactions are more likely to develop rapport with the role players (hypothesis 1), which seems logical. In addition, the data suggest that engagement can be made manifest in not only overt behaviors perceptible by observers, but also in more subtle behaviors displayed by the soldier (e.g., activity; hypothesis 2a) or between the soldier and role players (e.g., mirroring; hypothesis 2b). Specifically, there seemed to be two trends that emerged from the data. Of the two types of activity (body and vocal), only the body activity showed a consistent (negative) relationship with engagement. Of the two types of mirroring, only the vocal mirroring showed a consistent (positive) relationship with engagement. This might suggest that when a soldier is truly engaged, he minimizes body activity which involves more gross motor behaviors, but increases vocal mirroring behavior which involves more acute sensory behaviors. This result is an exciting step toward a better understanding of social interactions between soldiers and the strangers they encounter in their evolving role in an asymmetric warfare environment. The application of such knowledge in military operations will require that such behaviors are refined and translated to skills that are easily trainable throughout the military.

CONCLUSIONS

The most insightful outcome of this study was the promising ability to not only extend the envelope of measurement to capture previously imperceptible behaviors, but to link those behaviors to meaningful outcomes that could have tactical and strategic implications. The evolution of warfare to its current state, where lines between friend and foe are blurry and dynamic, requires that soldiers are equipped with more than kinetic instruments of coercive influence, and must be trained in the use of more tactful non-kinetic social influence techniques such as the behaviors investigated in this study. Although this study was limited based on the small sample sizes obtained from one training site with role players, the results suggested some exciting potential with respect to the role of subtle, honest signals. More important, perhaps, is that the methods introduce significant advances in the field of social science. The ability to capture and analyze unconscious social signals at a millisecond level using unobtrusive lightweight sensors will facilitate the design of real-time feedback tools and social prosthetics that may be used to enhance critical social competences of members of the armed forces in unfamiliar social environments. As this research stream matures, and subsequent studies gather more data, it is likely that additional patterns of behavior will emerge that could have similar impacts to training and operations. It is critical that soldiers are equipped with the skills and tools commensurate with the challenges they face as their roles evolve in an asymmetric warfare environment. This study is but one initial attempt to provide the understanding of social interaction mechanisms that will become the foundation of future soldier training.

ACKNOWLEDGEMENT

This work was supported by The Defense Advanced Research Projects Agency (Contract #W911NF-11-C-0266). The views, opinions, and/or findings contained in this article are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

REFERENCES

Beach, M., Roter, D., Wang, N., Duggan, P., Cooper, L. (2006). Are physicians' attitudes of respect accurately perceived by patients and associated with more positive communication behaviors? *Patient Education Counseling*. 62(3), 347–354.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher student interactions. *Journal of Nonverbal Behavior*, 12, 120-138.

Caligiuri, P., Noe, R., Nolan, R., Ryan, A., & Drasgow, F., (2011). *Training, Developing, and Assessing Cross-Cultural Competence in Military Personnel.* U.S. Army Research Institute: Arlington, VA.

Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Costa, P. T. Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Manual*. Odessa, FL: Psychological Assessment Resources.

Drolet, A. L., & Morris, M. W. (2000). Rapport in conflict resolution: Accounting for how nonverbal exchange foster cooperation on mutual beneficial settlements to mixed motive conflicts. *Journal of Experimental Social Psychology*, 36(1), 26–50.

Dunn, J. R., & Scheitzer, M. E. (2005). Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology*, 88(5), 735-748.

Exner, J. E. Jr. (1966). Variations in WISC performances as influenced by differences in pretest rapport. *Journal of General Psychology*, 74(2), 299-306.

Feldman, S. E., & Sullivan, D. S. (1971). Factors mediating the effects of enhanced rapport on children's performance. *Journal of Consulting and Clinical Psychology*, 36(2), 302.

Fisher, R. P., & Geiselman, R. E. (1992). Memory-*enhancing techniques for investigative interviewing: The cognitive interview.* Springfield, IL: Charles C. Thomas.

French, J. R. P., & Raven, B. (1959). The bases of social power. In D. Cartwright & A. Zander (Eds.). *Group Dynamics*. New York: Harper & Row.

Gibbs, R., & Clark, N., (2012). No need for instinct: Coordinated communication as emergent self-organized process. *Pragmatics and Cognition*, *20*(2), 241-262.

Grammer, K., Honda, M., Juette, A., & Schmitt, A. (1999). Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of Personality and Social Psychology*, 77(3), 487-508.

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/). Internet Web Site.

Lakoff, R. (1976). *Language and woman's place*. New York: Harper.

Levinson, S. (2006). On the "Human Interaction Engine." In N. Enfield & S. Levinson (Eds.), *Roots of Human Sociality: Culture, Cognition and Interaction.* New York: BERG.

Levitan, R., Gravano, A., & Hirschberg, J. (2011). Entrainment in speech preceding backchannels. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pp.113–117.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734.

Olguin Olguin, D (2007). *Sociometric Badges: Wearable Technology for Measuring Human Behavior*. MAS Thesis. Massachusetts Institute of Technology. Cambridge, MA.

Olguin Olguin, D. & Pentland, A. (2010). Sensor-based organizational design and engineering. *International Journal of Organizational Design and Engineering*. 1(1/2), 69-97.

Pentland, A. (2010). *Honest signals: how they shape our world*. Cambridge, MA: MIT Press.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. New York: Oxford University Press/Washington, DC: American Psychological Association.

Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*. 88(5), 879-903.

Rasmussen, L., & Sieck, W. (2012). Strategies for Developing and Practicing Cross-Cultural Competence. *Military Review*, *92(2)*, 71-80.

Ross, W., & Weiland, C. (1996). Effects of interpersonal trust and time pressure on managerial mediation strategy in a simulated organizational dispute. *Journal of Applied Psychology*. 81(3), 228–248.

Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology*, *73*, 1296-1312.

Selmeski, B. (2007). *Military Cross-Cultural Competence: core concepts and individual development.* US Air Force Culture and Language Center.

Sociometric Solutions (2013). http:// http://www.sociometricsolutions.com/index.html.

Sun, X., Nijholt, A., Truong, K. P., & Pantic, M. (2011). Automatic understanding of affective and social signals by multimodal mimicry recognition. S. D'Mello et al. (Eds.): ACII 2011, Part II, LNCS 6975, 289–296.

Tomasello, M. (2008). *The Origins of Human Communication*. Cambridge, MA: Massachusetts Institute of Technology Press.

Vallano, J. P., & Comp, N. S. (2011). A comfortable witness is a good witness: rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology*, 25, 960–970.

Watson, A. C., Angell, B., Vidalon, T., Davis, K. (2010). Measuring perceived procedural justice and coercion among persons with mental illness in police encounters: the police contact experience scale. *Journal of Community Psychology*, 38(2), 206-226.

Yu, C., Aoki, P. M., & Woodruff, A. (2004). Detecting user engagement in everyday conversations. *Proc. 8th Int'l Conf. on Spoken Language Processing* (ICSLP) (Vol. 2), Jeju Island, Republic of Korea, 1329-1332.