# Using Simulation to investigate a Non-Anthropomorphic Framework for Communications within a Human-Agent War-Fighting Team

**Michael Proctor**
Department of Industrial Engineering
University of Central Florida
Orlando, FL, USA
*mproctor@mail.ucf.edu*

**Gilbert Cardona**
Major, Field Artillery, United States Army
Department of Industrial Engineering
University of Central Florida
Orlando FL, USA

**Christopher Hollander**
Institute for Simulation and Training
University of Central Florida
Orlando FL, USA

As complex, non-human agents become increasingly ubiquitous members of the US military war-fighting team, an effective and natural system of communications must be explored and developed to achieve human-agent collaboration across the entire team. A non-anthropomorphic communications framework does not exist that will support human-agent collaboration beyond current electronic control. This research surfaces a non-anthropomorphic framework of communications between human-human, human-agent, and agent-agent teams based on related literature review. The framework provides perspective on the potential breath of communication modalities that exist as well as the many challenges faced. Within this framework, this research uses simulation to explore the contribution to collaborative, covert military operations that non-verbal forms of communications between humans and agent team members might entail. Visual and audio cues considered include pose, motion, color, and non-speech sounds. In addition this article presents findings on contribution of these modalities to the military operation being considered as well as identifies issues for implementation, applications, implications, and areas for future research.

**Keywords:** agent, communication, haptic, human-agent team, multi-modal

## 1. Objective, Scope and Context

Information dominance is a critical dimension of US military superiority. Part of that dominance arises from distribution of electronic, optical, written, and/or oral messages and orders within combat formations down to the individual soldier [1]. As the dynamics of war-fighting and the organization of the war-fighting team evolve, US military dominance may be maintained and strengthened by the inclusion of new members to the war-fighting team with new forms of communications. These new team members may be embodied agents that may or may not possess extra-sensory input and output technologies.

Currently physically and/or electronically tethered robots routinely perform roles on the modern battlefield in the areas of explosive ordinance disposal, reconnaissance, surveillance, target acquisition, and target engagement [3]. Planning for warfare in the future envisions greater use of these systems as well as semi-autonomous and autonomous robots taking on characteristics of agents as defined by Rocha [4] [5]. We refer to the latter group of these systems as simply embodied agents.

The Future Combat System (FCS) is the planned next generation of US Army battlefield technology for the war-fighting team. FCS is made up of 14 + 1 + 1 systems with the last two +1 systems being the Soldier and the integrated communications network called the War-fighter Information Network – Tactical (WIN-T). Included in the Future Combat System is a network which enables multi-directional communication of electronic, optical, written, and aural information both for relay between [7] and command, and control of the FCS family of systems [6] including classes of unmanned aerial vehicles (UAV) and unmanned ground vehicles (UGV) that are on WIN-T.

Current technological mechanisms for human, robot, and embodied agent team member communications encompass only limited modalities. Limited modalities risk overburdening a particular mechanism to the point of inefficiency and risk communication failure should counter measures successfully block or make unreliable the current mechanism. A broader and more robust framework for communications using multiple mechanisms including the spectrum of human modalities may promote not only robustness but efficiency, effectiveness, interpretation without confusion, rapid individual learning, and consistency through team training with less demand upon time. To test these hypotheses is far beyond the scope of the research that we were sponsored to conduct, though we believe that our findings below add to the body of knowledge pertaining to these topics as contained in the open literature.

The scope of our funding awarded by the U.S. Army Research Laboratory enabled us to conduct literature research on communications between humans and agents as well as conduct a limited experiment on the use of signals within a simulated combat setting.

Given our military charter, our research is applied to a current military context.

Soldiers already conduct wartime missions, such as Presence Patrol or Movement to Contact operations in Iraq, in the context of an environment rich in multimodal communications. Unique sounds, odors, sights, touch, and even tastes communicate meaning. While technology dominates communications at the organizational level, soldier-to-soldier communications at the team level are dominated by non-technological communications methods. The most basic form of military communications are simple signals. From the Marine Corps "leg bump" in modern-day, urban house clearing operation [23] to the legendary hand and arm signals of Roger's Rangers, signals retain their importance. Although the specific movements and their desired meanings have changed, the basic purpose of hand and arm signals is to promote collaboration through covert communications between team members during combat operations [2].

As non-human, embodied agents go from robotic to autonomous they will take on greater human characteristics. While non-verbal signals by automated devices are as common in our society as the traffic light, as we ventured further into unchartered area of embodied agents taking on human characteristics, we felt it appropriate to better understand human information processing theory so as to guide our path. We assumed as self-evident that communications between non-human and human team members will naturally expand over time beyond the current confines of the tethered controller medium to mediums that enabled non-human, embodied agents to assume more human forms of communication to include non-verbal signals.

Using the concepts of human information processing for our non-human, embodied agent communication research efficiently capitalizes upon prior human multimodal capacities research to send and receive signals and avoids duplicating that enormous associated cost. Leveraging that prior research also enables a large step forward in applying technology enabled human-like mediums on the battlefield and making all elements of the war-fighting team more lethal and efficient. The question becomes; what input modalities or combinations of modalities should our human-like technology mediums provide for when communicating signals in an already input-output intensive environment? Modeling and simulation provides an inexpensive approach to gaining insight into different technology medium concepts applicable to this question and related questions. As way of clarification, by using the term, human-like, we do not intend to constrain our concept explorations to technology mediums of human form or capability. Rather, we recognize that technology can create embodied agents that have form different from and capabilities often beyond those inherently human.

Hence a theoretical communications framework provides communications researchers perspective to modeling and simulation of concepts for potential future communication technologies on the battlefield. A requirement of that framework is that it recognize the vital importance of covert communications in a combat setting between human-human and potentially human-agent team members. An objective of such modeling and simulation research would seek to add to the understanding of effective communications while not overburdening the users' input or output modalities within a combat environment. The objective of better understanding would better guide exploration of embodied agent sensory and extra-sensory technology concepts for future human-robotic agent communications.

## 2. Communications Background and a Theoretical Meta-language Framework

Humans instinctively process the world through multiple modalities. For instance when a car horn is blasted, people tend to look in the direction of the sound to encompass the scene holistically through multiple inputs. This is true for simple everyday events such as receiving the news, where people cite newspaper and television individually and together for their source of important news and current events [11]. It's interesting to note that 48% of the sampled group from Graber's research [11] read the newspaper and watch televised news (not simultaneously) to gather their information.

There may be at least two explainable reasons for human preference to multimodal information processing: perceptual integration and redundant-signal effect [12] [13] [14]. Perceptual integration is the concept of combining different input modalities into a singular multimodal representation of an object or rather an amplified form of that object. Redundant-signal effect is the concept of increased reaction time due to redundant bi-modal information processing from separate uni-modal sources. In essence redundant-signal processing enhances human perception and ultimately affects response selection and execution which in turn improves reaction times.

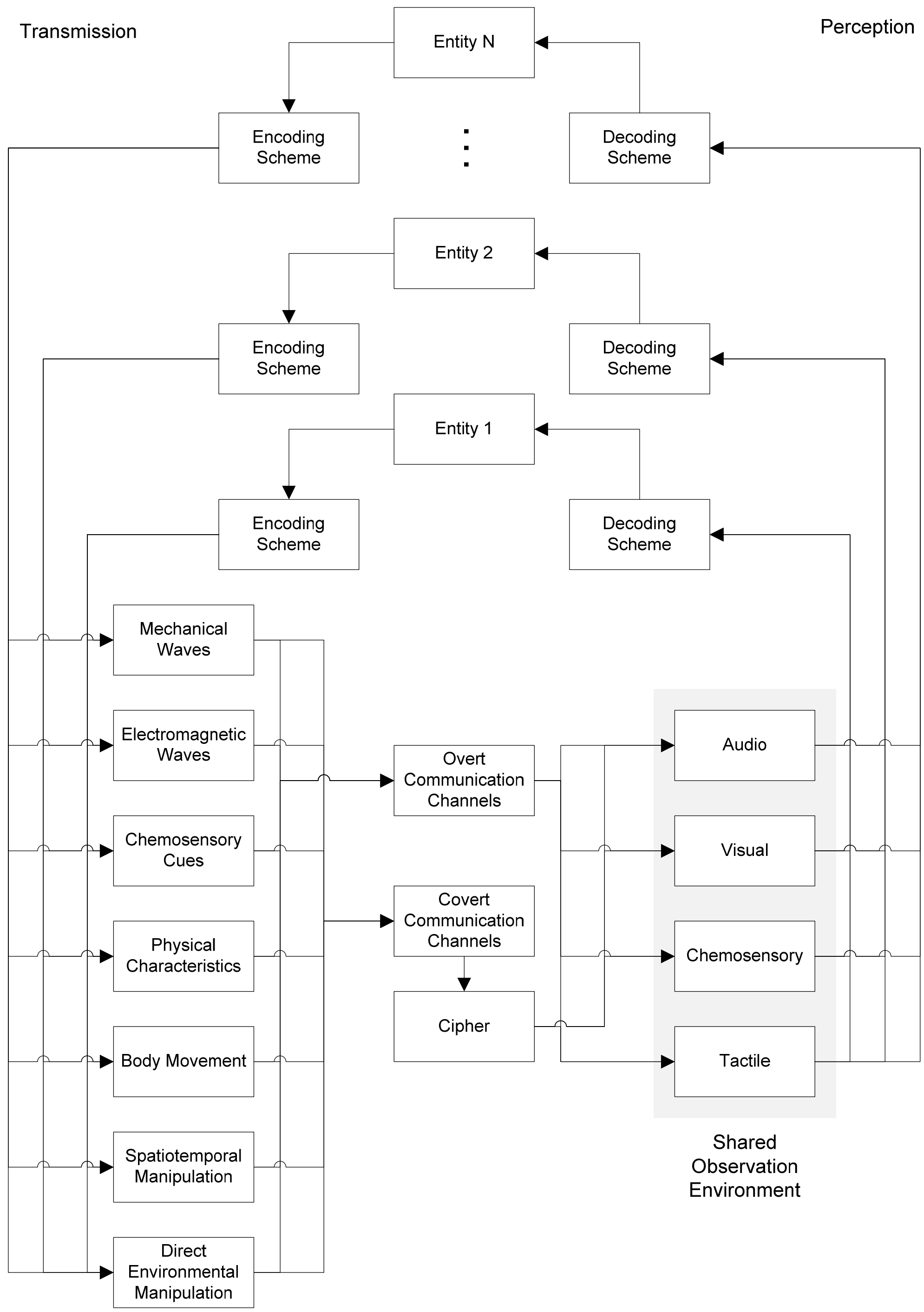



**Figure 1**. A multi-modal communication framework for trained and impromptu human-agent teams

In combat, team members implicitly often attempt to communicate with each other using multiple modalities - visually and through physical touch, not just audibly. Examples include audibly shouting, “INCOMING” while visibly gesturing for cover. Or physically touching or even shaking an individual while shouting their name to get their attention. Clearly embodied agent systems intended for autonomous, human-like behavior and missions must be capable of multi-modal communications far beyond traditional robotic behavior.

Figure 1 illustrates one framework proposed by Hollander and Proctor [24] for considering multi-modal communications between humans and agents. This framework accounts for both overt and covert communication mechanisms, making it ideal for military applications where security is of critical importance.

When meaning is associated with communications acts we have information. We have language when a set conventions to include vocabulary, syntax and grammar are associated with the information. Support for the development of a non-anthropomorphic meta-language framework is found in studies and research done on human information processing, multi-modal information processing and communications, and prior research into human-agent interaction [24][25][26][27].

### *2.1 Human Information Processing Research*

The human brain's ability to process information from different input modalities and the resultant effect upon overall time to react has significant bearing upon the development of non-anthropomorphic meta-language framework.

*The Psychology of Human-Computer Interaction* [8] began to look at the psychology of the human in its interaction with computers particularly but also measured the time for interaction using the different modalities and processing elements of the human brain. Card et al laid the foundation for the GOMS model (goals, operators, methods, and strategies) and the Human Information Processor.

There appears to be a gap on the exact understanding concerning processing times and decay time for haptic input modalities, but there is extensive research concerning haptic image processing. Newell et al determined the characteristics of haptic image encoding; with regards to a visual scene. They are encoded into the brain rather holistically, although recollection of specific details is incomplete, whereas haptic inputs are encoded individually and an entire image is rebuilt over time. This does imply that haptic input modalities are processed serially and not in parallel. It does not give insight into brain processing time for haptic input modalities, but [10] indicates that although the different input modalities are physically separate they share similar internal structures and properties (e.g. capacity and decay).

### *2.3 Human-Agent Interaction Research*

Work has been completed that created agents for the purpose of communicating with humans by mimicking human voice patterns, body language, and emotions [15]. The drawback to this particular system was that it was completed through a computer simulation that made the human user conceptually move into the virtual environment for the interaction to take place.

There are robotic agent systems available that communicate with human users, but given the highly specialized method by which these robotic agents are created and tested they can only communicate with their developers and programmers. This is a significant obstacle for the layperson or a Soldier to immediately communicate with an agent in a multimodal, highly stressful environment such as a combat engagement.

Another realm of prior research is developing agents to respond to human speech and body language, specifically in the case of Kismet at MIT [16]. Kismet is another example of a system where the robotic agent and developer are essentially training one another in communications. Kismet is unable to effectively communicate with a random adult human pulled off the street from somewhere in Cambridge, Massachusetts.

To date, most prior research has focused on an agent's ability to receive and understand input given by a human in the form of speech, gestures, and facial recognition. Work emphasizing an agent's ability to transmit information in a form more natural to humans (through non-verbal mechanisms) has been, until recently, largely unaddressed.

After reviewing the supporting material, several expectations arise for models and simulations of a multi-modal framework for human-agent communications. Based on the Card et al model of human information processing and work done on multimodal information processing, communications decoded through a haptic feedback device should excite the body's rapidly adapting mechanoreceptors and lead to quicker reaction time compared to audio or visual decoding. Combinations of input modalities may also lead to more effective means of communications.

## 3. Experimental Approach: Assessing Three Modes of Communications for the Human-Agent War-Fighting Team

### *3.1 Overview*

With the above charter, financial limitations, and communications framework and expectations, the objective of our research experiment was to establish a baseline of archival data and analysis from the open literature that assesses modes of agent-to-human communications consistent with the above framework yet in a simulated combat environment. We are not aware of any such data available in the open literature.

We took steps to insure both internal and external validation of our experiment. In terms of external validation, we based the synthetic environment, scenario, tactics, communications methods, and communications media on existing Army doctrine and sources as outlined below. Test subjects were trained through Army institutional training organizations in the military tactics that we employed. Subject Matter Experts validated the scenario once it was built. One additional external validation question that we also considered is: Do humans respond to stimuli in a simulation as the research literature indicates that they might be expected to respond in real life? We answer that question below based on our analysis of the data. In terms of internal validation measures, we designed the experiment as discussed below so as to avoid confounding of data while having sufficient power to yield statistically significant results where such differences exist. Since this experiment involved human test subjects, the entire experiment and data collection protocols were reviewed and approved by the University of Central Florida Institutional Review Board.

To limit complexity and to create a baseline of data, simple signal communication was chosen for analysis. Signal messages were sent using three different but widespread communication modalities – visual, auditory, and haptic. Effectiveness of the modalities was determined within the context of the interactive simulation used and of the experiment protocols.

The scenario modeled, using Advanced Interactive Systems, Desktop Soldier Visualization System (SVS), was the Movement to Contact mission conducted by a fire team in a contemporary middle-east urban setting. The movement to contact task as outlined in United States Army Field Manual 7-8 is an offensive technique to find and fix active enemy forces for a variety of purposes that include but are not limited to: destruction or bypassing of enemy forces, and/or engaging the enemy with the smallest element possible [17]. The war-fighting team in the scenario was composed of humanoid agents and a human avatar to be controlled by an actual human test subject. SVS is customizable PC simulation that was employed in this experiment using desktop monitors with attached speakers. SVS is used by the military and other agencies to train and experiment in various natural and urban environments and scenarios to meet the demands of military or law enforcement organizations.

Forty-five U.S. Army Reserve Officer Training Corps Cadets served as test subjects. Each cadet subject was exposed to the three input modalities in random order over the course of the simulation experiment.

With respect to the scenario, in a movement to contact as opposed to a presence patrol, urban civilian activities were not modeled as civilians were assumed to be in shelter. The scenario utilized a humanoid agent as *point man* (lead team member in the movement formation) in a position to directly communicate the message enemy in sight to human team members and elicit a response from the human team members. The experiment was solely focused on the movement to contact mission and the human response to the enemy in sight message between the humanoid agent *point man* and the human fire team member.

The three signals, visual, auditory, and haptic, conveyed the enemy in sight message. The humanoid agent point man, here after simply referred to as point agent, emitted the signal for the purpose of communicating the enemy in sight message to the human fire team member. The visual communication signal replicated current U.S. Army practices of hand and arm signals. Specifically when the point agent detected the enemy it moved from an erect walking posture to a prone position pointing its weapon in the direction of the enemy. The auditory communication signal emitted by the point agent and recommended for and used as the enemy in sight message was a low frequency sound typical of pack hunting hyena in the process of stalking prey. Similar audio signals have been used by military forces to covertly send messages throughout history. This includes American Indians famous for sending covert messages to each other by mimicking animal and bird calls. World War II Normandy paratroopers are famous for having used a device to send a cricket sound as a covert friendly force message to fellow soldiers. Finally the haptic communication signal emitted by the point agent was an electronic pulse that created a haptic signal on the human team member through a neoprene vest encapsulated with cell phone vibrators and worn by the human team member.

*3.2 Equipment*

The simulation was executed through two linked and synchronized Dell D820 laptop computers enabled with the same database and scenarios. In addition to the laptop computers, a neoprene vest impregnated with encapsulated cell phone vibrators and a Bluetooth® reception module enabled the haptic signal to be transmitted from the point agent to the human team member (figures 2 and 3). This device delivered a short vibration to the test subject in synchronization with activities occurring in the simulation to appropriately time point agent communications with vest vibration.

Figure 2: Haptic Communication Device (View 1).

Figure 3: Haptic Communications Device (View 2).

*3.3 Experiment Design*

Each test subject conducted the experiment individually under the supervision of the same test controller. After introductory remarks and prior to conducting the scenario, each test subject experienced three communications signals that represented the enemy in sight message used in the experiment. So that test subjects would be familiar with each signal and recognize it as the message enemy in sight, each signal was presented to each subject in random order in order to mitigate the potential effect of all subjects consistently hearing the signal in the same order. Specifically, we were mitigating short term memory bias by avoiding all subjects experiencing one signal last.

In each instance the point agent proceeded along a specific path within the synthetic middle-east urban environment during daylight hours with team members in trail conducting a movement to contact mission. Upon enemy contact the point agent sent the signal to the human team member and then proceeded to assault the enemy.

Each test subject was scored by two criteria: did they correctly interpret the point agent signal as an enemy in sight message and how long did it take to react to the message. The interpretation of agent team member communication was a go or no-go score while reaction times from test subjects averaged within signal category and compared to the other two communication signal test groups using an analysis of variance.

## 4. Results and Statistical Analysis

Data on participants and their performance were recorded after the execution of the experiment for each individual and completed before the execution of the experiment for the next individual. This allowed for the managing of samples to ensure that a minimum of thirty samples were recorded for each signal type and to allow for assumption of sample normality and the use of the z statistic. Empirical studies have been conducted and indicate that moderate departures from the assumption of normality do not seriously affect the confidence coefficients [18]. A confidence interval of 95.0% and $\alpha = 0.05$ were used throughout the calculations of statistical information.

The breakdown and demographics for the forty-five test subjects are presented in Table 1. The demographic survey was used mainly to provide insight into each test subject's computer familiarity for later use in regression analysis to determine whether computer familiarity influenced reaction times within the experiment. The immersive tendencies questionnaire was used to measure the test subject's proclivity to presence [19]. The immersive tendencies questionnaire was used to purely draw conclusions concerning a test subject's performance and their proclivity to immersion in simulation.

The test subjects were 86.6% male. This is not uncommon for Army ROTC Departments nor the United States Army whose active duty force is 85.7% male [20]. The mean military science level for the test group was 2.38 indicating the test subjects were mostly between their Military Science (MS) level 2 and MS level 3 years, while the mode for military science level was MS 3. The overwhelming majority of test subjects had no prior military background. There were 18 cadets with some military experience with the mode being up to one year. The mean age for the test subjects is 21.27 years and this closely corresponds with the mean military science level of the test subjects. A cadet who attends college immediately after graduating high school, joins the ROTC program, and proceeds at a full academic load would be a MS level 3 at the approximate age of 21 years. The last two demographic parameters are computer familiarity and proclivity to presence. The average computer familiarity score of three indicates that the majority of the test subjects are familiar with multiple software packages but are unable to program in a computer language.

| | *M* | *Median* | *Mode* | *SD* |
|---|---|---|---|---|
| Gender (Male = 1) | | | | |
| Female<br>Male | 0.87 | 1 | 1 | 0.34 |
| Military Science Level | | | | |
| 1 to 4 | 2.38 | 2 | 3 | 1.39 |
| Military Experience | | | | |

| | | | | |
|---|---|---|---|---|
| Range: 0 to 8 years | 1.42 | 0 | 0 | 2.31 |
| Age | | | | |
| Range: 18 to 28 | 21.27 | 21 | 19 | 2.6 |
| Computer Familiarity | | | | |
| Level 1 to 5 | 2.87 | 3 | 3 | 0.59 |
| Proclivity to Presence | | | | |
| Range 43 - 106 | 77.20 | 77 | 77 | 13.99 |

**Table 1**: Test Subject Demographic Breakdown

The simulation time clock was used to determine the time period from the moment the point agent sent a signal until there was a reaction made by the test subject. Test subjects failing to respond to any or all signals were given a maximum score of 10.00 seconds. Table 2 depicts the results for the 45 test subjects along with the mean and standard deviation for each communication signal and the percentage of successful reactions indicating the effectiveness of that communication. The percent of successful reactions is the total number of reactions observed for each communication signal compared to the total number of test subjects.

| | **VISUAL (1)** | **AUDITORY (2)** | **HAPTIC (3)** |
|---|---|---|---|
| **MEAN** | **4.20** | **5.39** | **2.78** |
| **STD DEV** | **3.49** | **3.36** | **2.40** |
| **% EFF** | **75.56%** | **66.67%** | **91.11%** |

**Table 2**: Test Subject Performance Data

Tables 3, 4, and 5 illustrate the mean, standard deviation, and 95.0% confidence interval for each sample. All calculations were performed in Minitab® 15 Statistical Software.

| **ONE SAMPLE Z TEST AND CONFIDENCE INTERVAL** | | | |
|---|---|---|---|
| The standard Deviation = 1.17 | | | |
| N | Mean | SE Mean | 95.0% CI |
| 34 | 2.327 | 0.201 | (1.934, 2.720) |

**Table 3**: Z Test and Confidence Interval for Visual Communications Signal

| **ONE SAMPLE Z TEST AND CONFIDENCE INTERVAL** | | | |
|---|---|---|---|
| The standard Deviation = 0.796 | | | |
| N | Mean | SE Mean | 95.0% CI |
| 30 | 3.082 | 0.145 | (2.797, 3.367) |

**Table 4**: Z Test and Confidence Interval for Auditory Communications Signal

| **ONE SAMPLE Z TEST AND CONFIDENCE INTERVAL** |
|---|
| The standard Deviation = 0.778 |

| N | Mean | SE Mean | 95.0% CI |
|---|---|---|---|
| 41 | 2.076 | 0.122 | (1.838, 2.314) |

**Table 5**: Z Test and Confidence Interval for Haptic Communications Signal

Comparing the values in Tables 3, 4, and 5, 95% of the population would react between 1.9 seconds and 2.7 seconds to a visual communications signal, between 2.8 seconds and 3.4 seconds to a auditory communication signal, and finally between 1.8 seconds and 2.3 seconds to a haptic communication signal.

*4.1 ANOVA Analysis of Performance Scores*

The null hypothesis for the analysis of variances for the three sample group is that $\mu 1 = \mu 2 = \mu 3$, while the alternative hypothesis is that $\mu 1 \neq \mu 2 \neq \mu 3$. Table 7 describes the values determined by performing one-way ANOVA and comparing each sample mean to each other sample mean. Figure 3 graphically displays the differences in means and the confidence intervals for each sample.

| **ONE-WAY ANOVA: VISUAL, AUDITORY, HAPTIC** | | | | | |
|---|---|---|---|---|---|
| Source | DF | SS | MS | F | P |
| Factor | 2 | 15.495 | 7.747 | 8.18 | 0.001 |
| Error | 87 | 82.376 | 0.947 | | |
| Total | 89 | 97.871 | | | |
| | | | | | |
| S = 0.9731 | | R-Sq = 15.83% | | R-Sq(adj) = 13.90% | |

**Table 6**: One-way ANOVA for Visual, Auditory, and Haptic

```
Individual 95% Confidence Intervals For Mean Based on Pooled
                    Standard Deviations

Level  N   Mean    StDev     ----+---------+---------+---------+-----
VIS    34  2.3268  1.1705        (------*-----)
AUD    30  3.0823  0.7963                       (------*-----)
HAP    41  2.0763  0.7777     (-----*----)
                             ----+---------+---------+---------+-----
                                2.00      2.50      3.00      3.50

Pooled StDev = 0.9277
```

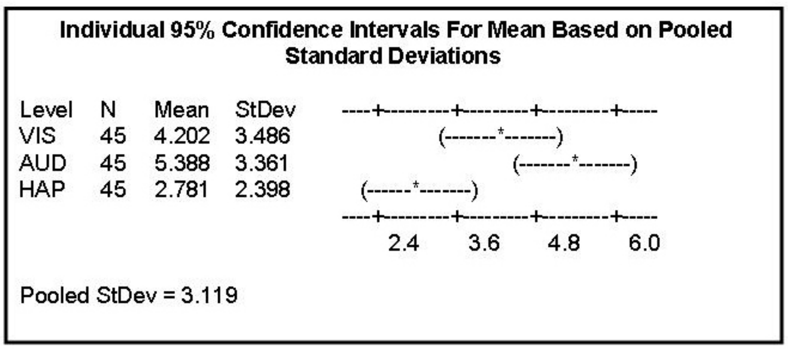
**Figure 4**: Individual 95% CI's for Mean Based on Pooled Standard Deviation.

The one-way ANOVA test of the three samples (table 7, figure 4) is statistically significant due to the calculated p-value of 0.001. With a confidence interval of 95.0% and $\alpha = 0.05$, the resulting p-value $< \alpha$, and therefore the null hypothesis can be rejected whereas the alternative hypothesis, $\mu 1 \neq \mu 2 \neq \mu 3$, cannot be rejected. The input for this ANOVA test was the values from Table 3 and do not include max scores.

A second ANOVA test of the three samples (table 7, figure 5) was performed maintaining max scores within the samples. This second ANOVA is also statistically significant due to the calculated p-value of 0.001. With a confidence interval of 95.0% and $\alpha = 0.05$, the resulting p-value $< \alpha$, and therefore the null hypothesis can be rejected whereas the alternative hypothesis, $\mu 1 \neq \mu 2 \neq \mu 3$, cannot be rejected.

| **ONE-WAY ANOVA: VISUAL, AUDITORY, HAPTIC w/ MAX SCORES** | | | | | |
|---|---|---|---|---|---|
| Source | DF | SS | MS | F | P |
| Factor | 2 | 153.4 | 76.7 | 7.88 | 0.001 |
| Error | 132 | 1284.49 | 9.73 | | |
| Total | 134 | 1437.9 | | | |
| | | | | | |
| S = 3.119 | | R-Sq = 10.67% | | R-Sq(adj) = 9.32% | |

**Table 7**: One-way ANOVA for Visual, Auditory, and Haptic Incorporating Max Scores.



```
Individual 95% Confidence Intervals For Mean Based on Pooled
Standard Deviations

Level  N   Mean   StDev   ----+---------+---------+---------+-----
VIS    45  4.202  3.486            (-------*-------)
AUD    45  5.388  3.361                     (-------*-------)
HAP    45  2.781  2.398     (------*-------)
                          ----+---------+---------+---------+-----
                             2.4       3.6       4.8       6.0

Pooled StDev = 3.119
```

**Figure 5**: Individual 95% CI's for Mean Based on Pooled Standard Deviation Incorporating Max Scores.

Given the large F-statistic for each of the ANOVA tests, a second analysis of the data was conducted to using Tukey Post Hoc analysis. The results of each communication signal (both with and without 10.00 second max scores) were pair-wise analyzed against one another and are depicted in tables 8 and 9.

| **Scenario** | **Calculated Value** | **Tukey Critical Value** | **Significance** |
|---|---|---|---|
| Audio vs. Haptic | 5.662 | 3.49 | Yes |
| Visual vs. Haptic | 1.410 | 3.49 | No |

Table 8: Tukey Post Hoc Analysis (No Max Scores)

| **Scenario** | **Calculated Value** | **Tukey Critical Value** | **Significance** |
|---|---|---|---|
| Audio vs. Haptic | 5.606 | 3.42 | Yes |
| Visual vs. Haptic | 3.055 | 3.42 | No |

**Table 9**: Tukey Post Hoc Analysis (w/ Max Scores)

Given the final results from the Tukey Post Hoc Analysis there does not appear to be a significant improvement of haptic communication signals compared to the visual communications with using max scores and no max scores. Yet in both instances there appears to be a significant improvement of using haptic communication signals when compared to audio communication signals.

*4.2 Regression Analysis of Mitigating Factors in Overall Performance*

The main purpose of the demographic survey was to gain insight into the computer familiarity of each test subject. Since the main vehicle for the experiment was a PC simulation, regression was used to investigate and mitigate computer use as a factor in overall performance. The immersive tendencies questionnaire was another factor used to mitigate overall performance and was used as a independent variable for regression. There were wide variations within the responses to each communication cue but the regression data shows that computer familiarity and the test subject's proclivity to presence had no effect upon their overall score.

## 5.0 Discussion

This research was done in order to establish a body of literature and gain insight into effectiveness of signals from a conceptual embodied agent to a human team member using three different communication modalities within a simulated combat environment.

The final results of the ANOVA for the experiment do not support the null hypothesis that the mean of each sample are equal and therefore the null hypothesis of equality can be rejected. From an external validation perspective, these findings are consistent with pre-existing human-centered non-combat, communications literature. That is haptic messages often invoke a more rapid response than other messaging modalities. Although the resultant p-value provides a statistically significant result, the statistical results do not infer that the result is of practical significance or indicates a large effect in the overall US Army population [21].

However, though these time savings may not appear to be exceptionally significant, Subject Matter Experts returning from recent combat experiences indicate that in the life and death situations of combat any advantage over the enemy should be used and exploited to maximize friendly forces operational effectiveness.

Additionally and not a point of insignificance, there was a large portion of the sample that did not even receive the signal for the visual and auditory communications. Subject Matter Experts indicate reliability of communications in combat is a life or death concern. Further, in this experiment, the visual signal was communicated in an environment free of visual obscurants. Under these ideal conditions and determining signal reception as percentage of effectiveness, visual communications were 75.56% effective, auditory communications were 66.67% effective, while haptic was 91.11% percent effective.

Given these external factors, the reliability of the haptic signal combined with the faster reaction time indicates that the haptic signal is the best of the three signals in our experiment.

## 6. Conclusion, Research Limitations, and Recommended Future Research

Visual and auditory are but two channels for communications within the human-agent teams be considered for future U.S. Army combat systems. Haptic communications have the potential to be a versatile form of communications. An excellent example is the cell phone vibrator where separate pulses communicate to the user different signals such as incoming call, incoming text message or voicemail, and even low battery. Given short time periods, haptic communications potentially have the same or greater depth to messages that are available with visual or auditory communications.

A benefit of haptic communications versus visual and auditory is how it may fit in a war-fighting environment that will someday contain humanoid agents serving as soldiers in various combat capacities. Through the use of simulation, we consider communications within a human-agent war fighting team wherein the very dangerous position of point man is replaced by a humanoid soldier agent, which we simply refer to as a point agent. Examining three modes of communications, we found a haptic signal to be the most reliable signal between the point agent and a human team member. We also found that the haptic signal resulted in the fastest response by the human to the signal. Advantages also noted include soldiers within the team do not have to be in a position to see an autonomous agent nor do they have to be within range to hear a signal from the agent. Further, covertness of communications is maintained regardless of environment, time of day, or type of operation. This provides a complication for those attempting to intercept communications between team members and units in an effort to disrupt friendly unit operations. Additionally, through the available literature research it is evident that humans cannot only handle multimodal interactions but excel and prefer those types of interactions. Multiple input modalities provide the human brain a holistic view of any given scene/event and enhance situational awareness [22].

### *6.1 Limitations of Research*

Although the Soldier Visualization System enables the exercise controller to develop a scenario to train a specific task it does also have its limitations. At the time of this research the simulation did not have the library of models or tools or programming interface to enable a user to create entities to mimic the full range of hand and arm signals in use by the United States Military. Hand and arm signals continue to be used successfully in war fighting and we would expect a humanoid soldier agent will need to be able to perform these signals across a number of military combat positions. Visualization limitations with the version of SVS used in this experiment limited the capabilities and realism of visual communications to be used by the point agent. Alternative hand and arm signal for enemy in sight is a soldier in the standing, kneeling, or prone position pointing in the direction of the enemy. As a result, a full range of visual communications (lights, poses, individual movements, etc.) was limited to those mentioned above.

As for the experiment itself, the length of the scenarios did not enable us to model or simulate use of haptic messaging and wear of the haptic vest over an extended period of time. Since this was a modeling and simulation experiment whose objective was to examine a concept and not either extended operational or fit and form of the haptic vest, these limitations were not seen as significant to our findings. To obtain meaningful insight into actual long term operational or human factors would require a completely different experiment over more realistic scenarios over multiple effectiveness measures, so that one can have a better insight about the performance of selected factor levels in alternative contexts and potential effectiveness indicators of interest.

Further, we did not address the flip side of this experiment; that is, human to agent communications. This is a legitimate area of research that of course would also involve another experiment whose focus could involve agent sensory hardware, communication recognition software, memory, cognition, decision making, and/or motor behaviors rather than human perception and response which this experiment focused on.

*6.2 Areas for Future Research*

Building upon the results of the experiment and taking it a step further there appear to be several areas of future research:

- Randomize the agent models (i.e. humanoid vs. robotic in appearance) and re-evaluate the experiment and track performance.
- Incorporate a team setting - with 3 or 4 people going through the scenario at once and re-evaluate the experiment in and track performance in an extended period of time.
- Human-Agent forms of communications must be considered in various environment elements to include dust, dust storms, night, fog, and mixed lighting situations common to urban environments.
- Human-Agent forms of communications must be considered in military scenarios such as the presence patrol that involves civilian populations.

Other potential research topics include:

- Confounding the haptic input and re-evaluate the experiment.
- Evaluation of location for haptic communication reception.
- Development and testing of a system of communications that is multimodal (e.g. uses a combination of visual, auditory, and haptic signals).
- Development of a haptic communications language, one that encompasses the current library of hand and arm signals and the scenarios for which they are used.
- Evaluation of haptic communications against visual communications using active duty combat arms Soldiers.
- Incorporation of haptic messaging device into Land Warrior.
- Development of a human interface that will enable human team members to communicate with one another with minimal keystrokes while maintaining a minimum of one hand on a weapon.
- Communications security for wireless communications between team members.

## 7. References

[1] Zieniewicz, M, & Goodman, E (2003). Battlefield Information Concepts for the Highly Mobile Warrior. *Army AL&T. Mar-Apr*, 23-25.

[2] Headquarters, Department of the Army. (1987). *Visual Signals* (FM 21-60). Washington D.C: U.S. Government Printing Office.

[3] Lerner, P (2005). Robots go to War. *Popular Science*, Retrieved October 1, 2007, from http://www.popsci.com/popsci/technology/generaltechnology/f59ca60b2e948010vgnvcm1000004eecbccdrcrd.html

[4] Hill, R, Champagne, L, & Price, J (2004). Using Agent-based Simulation and Game Theory to Examine the WWII Bay of Biscay U-boat Campaign. *Journal of Defense Modeling and Simulation. 1, Issue 2*, 99-109.

[5] Rocha, L. M. (1999). Artificial Life to Semiotic Agent Models: Review and Research Directions. Los Alamos National Laboratory Report, LA-UR-99-5475. Los Alamos, New Mexico.

[6] U.S. Army. "FCS Network." FCS, On the Network the Strong Become Stronger. Available from http://www.army.mil/fcs/network.html. Internet; accessed 1 October 2007.

[7] U.S. Army. "Class 1 Unmanned Aerial Vehicle (UAV)." FCS, On the Network the Strong Become Stronger. Available from http://www.army.mil/fcs/uav1.html. Internet; accessed 1 October 2007.

[8] Card, S., Moran, T., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Publishers.

[9] Newell, F.N., et al. (2006). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research*, *161, No. 2*, 233-242.

[10] Stanney, K., et al. (2004). Functional Optical Brain Imaging Using Near-Infrared During Cognitive Tasks. *International Journal of Human-Computer Interaction*. *17, No. 2*, 229-257.

[11] Graber, D. (1988). *Processing the News: How People Tame the Information Tide*. White Plains, NY: Longman, Inc..

[12] O'Hare, J. (1991). Perceptual Integration. Journal of the Washington Academy of Sciences, 81, 44-59.

[13] Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of cognitive neuroscience, 11*(5), 473-490. Retrieved March 26, 2007, from Biological Sciences database.

[14] Miller, J. (1982). Divided Attention: Evidence for Co-Activation with Redundant Signals. Cognitive Psychology, 14, 247-279.

[15] Bruce, A., Nourbakhsh, I., Simmons, R., The Role of Expressiveness and Attention in Human-Robot Interaction. (2002) Retrieved Jan, 2007 from http://www.cs.cmu.edu/~illah/PAPERS/bruce_icra2002.pdf

[16] MIT Humanoid Robotics Group, "Kismet". (2006) Retrieved Feb, 2007 from http://www.ai.mit.edu/projects/sociable/overview.html

[17] Headquarters, Department of the Army (1992). *Field Manual No. 7-8 The Infantry Rifle Platoon and Squad*. Washington D.C.: Department of the Army.

[18] Mendenhall, W, & Sincich, T (1995). *Statistics for Engineering and the Sciences*. Upper Saddle River, NJ: Prentice-Hall, Inc..

[19] Singer, M, Witmer, B., Goldberg, S. (1996). Presence Measures for Virtual Environments: Background & Development.

[20] US Army G-1, (2005). Army G-1. Retrieved March 6, 2007, from Army G-1 Human Resources Policy Web site: http://www.armyg1.army.mil/hr/demographics.asp

[21] Gelman, A., Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. The American Statistician, November 2006, Vol. 60, No. 4.

[22] Proctor, M, & Gerber, W (2004). Acoustic Considerations for Land Combat, Entity-Based Simulation. *Journal of Defense Modeling and Simulation*. *Volume 1, Issue 4*, 225-233.

[23] Proctor, M. & Woodman, M. (2007) Training "Shoot House" Tactics using a Game. *Journal of Defense Modeling and Simulation*. *Volume 4, Issue 1*, 55-63.

[24] Hollander, C. & Proctor, M. (2007) A Multi-Modal Communication Framework for Trained and Impromptu Teams of Agents and Humans.
[25] N. Matsumoto, H. Fujii, M. Okada, "Minimal design for human–agent communication," Artificial Life and Robotics vol. 10 pp. 49-54, 2006.
[26] C. Bethel and R. R. Murphy, "Non-facial/Non-verbal Affective Expressions for Appearance Constrained Robots," IEEE Transactions on Systems, Man and Cybernetics Part C, accepted pending revisions.
[27] T. Fong, I. Nourbakhsh, K. Dautenhahn, "A survey of socially interactive robots: Concepts, Design, and Applications," Technical Report CMU-RI-TR-02-29, 2002.

**Author Biographies**

**Michael D. Proctor** is an Associate Professor at the University of Central Florida. He is a retired U.S. Army Lieutenant Colonel and holds a Ph.D. in Industrial Engineering from North Carolina State University (1991).

**MAJ Gil Cardona** is currently assigned to Fort Bragg as Simulation and
Exercise Officer, G5, Plans, 82nd Airborne Division, U.S. Army. He is a graduate of the US Army Command and Staff College and holds a Masters in Industrial Engineering (2007) from the University of Central Florida.

**Christopher Hollander** is a PhD student at the University of Central Florida, Department of Modeling and Simulation, Orlando, Florida, USA. He is also a research assistant in the ACTIVE lab at the Institute of Simulation and Training, Orlando, Florida, USA.