



ARL-TR-8411 • JULY 2018



Wearable Smart Interaction Device (SID) for Advanced Human–Robot Interaction Using Gestures and/or Speech

by Rodger Pettitt, Linda R Elliott, and Glenn Taylor

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.



Wearable Smart Interaction Device (SID) for Advanced Human–Robot Interaction Using Gestures and/or Speech

by Rodger Pettitt and Linda R Elliott
Human Research and Engineering Directorate, ARL

Glenn Taylor
Soar Technology, Inc., Ann Arbor, MI

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) July 2018		2. REPORT TYPE Technical Report		3. DATES COVERED (From - To) 2 October 2017–21 May 2018	
4. TITLE AND SUBTITLE Wearable Smart Interaction Device (SID) for Advanced Human–Robot Interaction Using Gestures and/or Speech				5a. CONTRACT NUMBER W911NF-13-C-0071	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Rodger Pettitt, Linda R Elliott, and Glenn Taylor				5d. PROJECT NUMBER SBIR A13-068	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory ATTN: RDRL-HRB-AB Aberdeen Proving Ground, MD 21005-5425				8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-8411	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Robotic assets used by dismounted Soldiers have usually been controlled through continuous and effortful teleoperation; however, more autonomous capabilities have been developed that reduce the need for continuous control of movements. While greater autonomy can make robotic systems more useful, users still need to interact with them, and operator control units for deployed robots still primarily rely on manual controllers such as joysticks. This report describes an evaluation of a multimodal interface that leverages speech and gesture through a wrist-worn device to enable an operator to direct a robotic vehicle using ground guide-inspired or infantry-inspired commands through voice or gesture. A smart watch is the primary interaction device, allowing for spoken input (through the microphone) and gesture input using single-arm gestures. Both modalities were shown to be useful, with each approach having particular advantages, depending on operational context.					
15. SUBJECT TERMS Army robotic systems, ground robots, autonomous systems, gesture commands, speech commands					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 56	19a. NAME OF RESPONSIBLE PERSON Rodger Pettitt
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) 706-545-9142

Contents

List of Figures	v
List of Tables	v
Acknowledgments	vi
1. Introduction	1
1.1 Statement of the Problem	1
1.2 Overview of Assessment	2
1.3 Objectives	2
2. Method	2
2.1 Participants	2
2.1.1 Pretest Orientation and Volunteer Agreement	3
2.1.2 Medical Review and Screening	3
2.2 Instruments and Apparatus	3
2.2.1 Smart Interact Device (SID) for Advanced Human–Robot Interaction	3
2.2.2 National Aeronautics and Space Administration (NASA)-Task Load Index (TLX)	5
2.2.3 Questionnaires	5
2.3 Procedures	5
2.4 Data Analysis	7
3. Results	7
3.1 Demographics	7
3.2 Training Evaluation	8
3.3 Observer-Based Performance Measures	9
3.4 Post-Session Questionnaires	11
3.5 End of Experiment Comparative Evaluation Questionnaire	14
3.6 Advanced Speech Commands	14

3.7 NASA-TLX	16
4. Conclusions and Recommendations	17
5. References	20
Appendix A. National Aeronautics and Space Administration (NASA)-Task Load Index (TLX) Mental Workload Rating Scale	21
Appendix B. Questionnaires	23
Appendix C. Soldier Comments	35
List of Symbols, Abbreviations, and Acronyms	47
Distribution List	48

List of Figures

Fig. 1	Examples of SID gestures	4
Fig. 2	SID for advanced human–robot interaction and Lockheed Martin’s SMSS	5

List of Tables

Table 1	Commands currently implemented in SID for controlling a robotic mule.....	6
Table 2	Soldier-based training ratings	8
Table 3	Feedback on command reliability	9
Table 4	Mean number incorrect response errors (N = 24).....	10
Table 5	Mean number of no response errors (N = 24).....	10
Table 6	Mean number of communication errors (N = 24).....	11
Table 7	Ratings of difficulty of commands, by speech and gesture	12
Table 8	Ratings of reliability of commands by speech and gesture.....	12
Table 9	Opinion survey ratings.....	13
Table 10	Summary count of system preference regarding command reliability	14
Table 11	Advanced speech command errors	14
Table 12	Ratings for each advanced speech command.....	15
Table 13	Overall experience ratings	15
Table 14	Mean NASA-TLX ratings for speech, gesture, and advanced speech.....	16
Table 15	Paired-comparison t–tests comparing speech and gesture for NASA-TLX.....	16

Acknowledgments

We would like to thank and acknowledge the contributions of the project managers supporting the effort described in this report: Arthur W Evans and Susan G Hill (US Army Research Laboratory/Human Research and Engineering Directorate), who are leading multiyear investigations of human–robot interaction. We would also like to thank SoarTech and Lockheed Martin staff who provided support during the evaluation. We also thank all the support staff that collected performance data alongside each Soldier and those who diligently organized and reduced the data. We are particularly indebted to the Soldiers who participated readily and steadily throughout this event. Their expert perspective provided great insight with regard to critical issues pertaining to operational effectiveness in battlefield contexts.

1. Introduction

1.1 Statement of the Problem

Robotic assets used by dismounted Soldiers have usually been controlled through continuous and effortful teleoperation; however, more autonomous capabilities have been developed that reduce the need for continuous control of movements. While greater autonomy can make robotic systems more useful, users still need to interact with the more advanced systems, and operator control units for deployed robots still primarily rely on manual controllers such as joysticks. While not yet fielded, some robotic control devices have begun to incorporate more naturalistic controls such as speech and gesture that can ease the burden of interacting with the robots by leveraging how people communicate with each other.

The potential of speech and gesture to enhance Soldier-robotic system interaction is supported at various levels. At the user level, interviews with active duty junior and senior enlisted Soldiers regarding advanced concepts for robotic systems have shown high support ratings for more naturalistic controls and autonomous systems and controls (Swiecicki et al., 2015; Elliott et al., 2016; Pettitt et al., 2017). A review of gesture-based systems showed significant progress in use of gestures for robot control (Elliott et al., 2014; Elliott et al., 2016) and concluded that wearable gesture-based systems were more suited to military operations when compared to camera-based systems. Wearable glove-based devices for robot control have been demonstrated successfully for Soldier use (Baraniecki et al. 2017). From the literature review, results suggested strongly that integration of gesture-based and speech-based controls would maximize advantages of each approach; gestures (e.g., pointing) can intuitively communicate direction, while speech can provide action detail (e.g., go around that car; open the door over there). In addition, gestures based on wearable technology can be effective in high-noise, low-visibility situations.

One such implementation of a speech and gesture-based robotic control system is SoarTech's Smart Interaction Device (SID), funded in part through US Army TARDEC (SBIR A13-068, contract number W56HZV-15-C-0050). SID is a multimodal human-machine interface designed specifically for the control of unmanned vehicles and has been used with a variety of unmanned air and ground platforms (Taylor 2017). SID's supported interactions have focused on ground guide-inspired commands, such as turning in place or moving forward/backward, or infantry-inspired commands, including following a leader or setting rally points. Unlike some other approaches that use sensors on the robot to capture user input, SID employs a worn smart watch as the primary interaction device, allowing for

spoken input (through the microphone) and gesture input (using the inertial sensors on the watch), while keeping the user's hands free for other tasks. Since it is important to maintain the user's situation awareness, SID also provides multi-modal feedback to the user through the watch, including visual (on the watch face), speech, and vibration, to let the user know the status of commands or of the vehicle.

1.2 Overview of Assessment

The assessment of the SID was based on the evaluation of a common set of SID operational commands for each command modality (e.g., speech, gesture). Soldiers were introduced to system capabilities and provided with sufficient training and practice to complete operational tasks using SID commands. Assessment tasks were constructed to evaluate command effectiveness within an operational context and included commands such as Forward, Backward, Turn (left or right), Stop, and Follow me. Each Soldier used both SID speech and gesture commands for robot control. Also, they used advanced SID speech commands that explored more autonomous robotic functions such as setting up routes, specifying degrees of turn, and establishing rally points. After performance of the scenarios/assessment tasks, Soldiers participated in after-action reviews and provided feedback through structured ratings regarding ease of use, workload, and operational relevance.

1.3 Objectives

The objectives of this effort were to evaluate options for control of a robotic mule using SID. In this preliminary evaluation, data were collected to answer the following questions, to better refine systems for operational use.

- How well does SID recognize inputs from Soldiers?
- Can the Soldiers successfully use SID to control the robotic mule in a variety of maneuver tasks?
- Do the Soldiers find SID useful for controlling the robotic mule?
- What improvements to the system would the Soldiers find useful?

2. Method

2.1 Participants

Twenty-four Soldiers assigned to operational and training units located at Fort Benning, Georgia, participated in the 6-day assessment. Participants did not receive

any compensation for participating in this investigation. Participant demographic data are summarized in Section 3.

2.1.1 Pretest Orientation and Volunteer Agreement

The Soldiers were given an orientation on the purpose of the assessment and their participation. They were briefed on the objectives and procedures for the assessment, as well as the equipment they would be required to use throughout the assessment. They were also told how the results will be used and the benefits the military could expect from the assessment. It was made clear that Soldier participation in the assessment was voluntary.

2.1.2 Medical Review and Screening

After arriving at the assessment site, the investigators asked the Soldiers if any of them had a medical profile or history that would jeopardize them if they participated in the study.

2.2 Instruments and Apparatus

2.2.1 Smart Interact Device (SID) for Advanced Human–Robot Interaction

SoarTech designed this particular version of SID to be used with robotic mules that move along with small dismounted units to offload some of the weight carried by Soldiers (Taylor et al. 2017). SID's supported interactions have focused on ground guide-inspired commands, such as turning in place or moving forward/backward, or infantry-inspired commands, including following a leader or setting rally points. Figure 1 shows examples of SID gestures. Unlike some other approaches that use sensors on the robot to capture user input, SID employs a smart watch as the primary interaction device, allowing for spoken input (through the microphone) and gesture input (using the inertial measurement unit on the watch), while keeping the user's hands free for other tasks. Since it is important to maintain the user's situation awareness, SID also provides multisensory feedback to the user including visual (on the watch face), speech, and vibration, to let the user know the status of commands or of the vehicle.

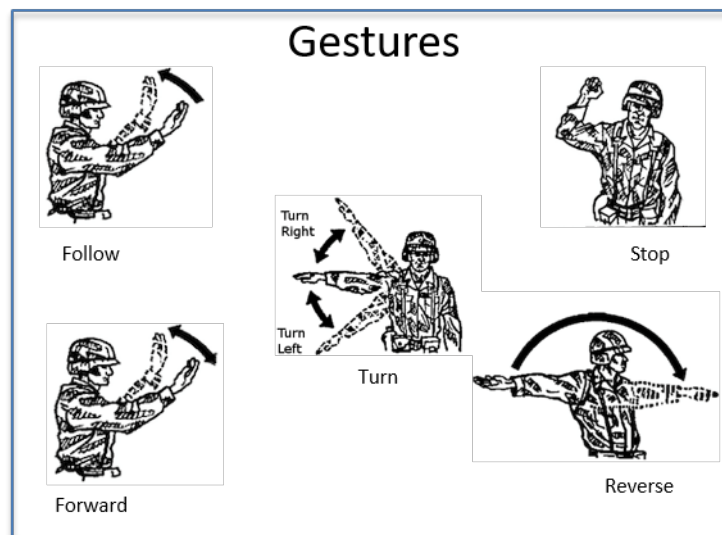


Fig. 1 Examples of SID gestures

Components:

- **Smart Watch:** The primary user interface is an Asus Zen Watch 2 smart watch worn by the operator (Fig. 2). The watch captures user gestures and speech input, and transfers inputs for further processing to the Computing/Relay platform using Bluetooth.
- **Computing/relay platform:** An Android-based smartphone or tablet in the operator's pocket or ruck. For the evaluation, this was a Nexus 7 tablet connected via Bluetooth to the smart watch and over Wi-Fi to the robotic platform and other computing resources. The smartphone/tablet also provides audio feedback to the user (e.g., through an earpiece or speaker).
- **Robotic platform:** Lockheed Martin's Squad Mission Support System (SMSS; Fig. 2) served as the robotic platform during the evaluation. SMSS uses a number of sensors, including a Light Detection and Ranging (LIDAR) for person tracking and obstacle avoidance. It also hosts a number of computing devices for sensor and other processing.



Fig. 2 SID for advanced human–robot interaction and Lockheed Martin’s SMSS

2.2.2 National Aeronautics and Space Administration (NASA)-Task Load Index (TLX)

The NASA-TLX is a subjective workload scale that measures subjective workload of operator(s) working with human–machine systems (Hart and Staveland 1988). It uses a multidimensional rating procedure based on six subscales. These subscales include degree of mental demand, physical demand, temporal demand, effort, frustration, and self-assessment of performance. It can be used to assess workload in various human–machine environments such as aircraft cockpits; command, control, and communication workstations; supervisory and process control environments; simulations; and laboratory tests. The NASA-TLX assessment scale can be found in Appendix A.

2.2.3 Questionnaires

Questionnaires were designed to elicit Soldiers’ opinions about SID and their ability to maneuver the robot using speech and gesture commands. The Soldiers rated system capabilities using 7-point semantic differential scales ranging from “extremely good/easy” to “extremely bad/difficult.” Questionnaires were administered to each Soldier at the completion of training and after each event. Questionnaires were also used to gather information concerning the Soldiers’ demographic data. Questionnaires can be found in Appendix B.

2.3 Procedures

Soldiers performed maneuver tasks in three different trials. The tasks were equivalent in gesture and speech trials to allow comparison of task performance and subjective workload demands. Individual Soldier training was conducted prior to each trial. Training included “hands-on” activity and practice controlling the system. Soldiers performed one trial with the SID speech commands and a trial

with gesture commands. The order of speech and gesture trials were counterbalanced across Soldiers. After each Soldier completed both speech and gesture trials, they completed a third trial designed to evaluate advanced speech-based commands. Table 1 shows the commands used in each trial. This report focuses on the evaluation of commands used by both speech and gesture.

Table 1 Commands currently implemented in SID for controlling a robotic mule

Command	Gesture	Speech	Advanced speech
Forward	X	X	...
Backward	X	X	...
Turn left	X	X	...
Turn right	X	X	...
Stop	X	X	...
Follow me	X	X	...
Start recording route <name>	...	...	X
Stop recording route	...	...	X
Follow route <name>	...	...	X
Set rally point <name>	...	...	X
Go to rally point <name>	...	...	X
Move forward/backward <x> meters	...	...	X
Turn left/right <x> degrees	...	...	X

Measures of performance for each trial were recorded by an observer. Measures of performance were based on observations regarding outcomes to each command in terms of SMSS errors in response to commands by SID. Soldiers completed three trials, one for each condition. Each trial consisted of three sets of the commands. During each set, Soldiers were instructed to give the command up to three times in order to get the correct response from the robot. Based on the robots position and in order to avoid obstacles some commands were attempted more often than others. All correct and incorrect responses were recorded.

At the end of each trial, Soldiers were asked to provide questionnaire-based feedback with regard to relevance of command, suggestions for improvement, and any operational issues that need to be considered. Feedback from the Soldiers pertaining to ease of use, comfort, usefulness, workload (NASA-TLX), and operational relevance was collected using 7-point semantic differential rating scales and verbal/written comments.

2.4 Data Analysis

Data collected from observation and questionnaire-based techniques were summarized using descriptive statistics. Tests of significance were based on SPSS (Statistical Package for the Social Sciences) paired-sample t-tests for the speech versus gesture comparisons (using standard $\alpha p = 0.05$). Holm-Bonferroni correction for familywise error rates were applied for multiple comparisons (Holm 1979). When considering several hypotheses, the problem of multiplicity arises: the more hypotheses we check, the higher the probability of a Type I error (false positive). The Holm-Bonferroni method is one of many approaches that control the familywise error rate (the probability that one or more Type I errors will occur) by adjusting the rejection criteria of each of the individual hypotheses or comparisons. Correlation-based analyses were also conducted for between-subjects variables (e.g., Soldier-based vs. trainer-based ratings). Corresponding descriptive statistics for user-based feedback are provided.

3. Results

3.1 Demographics

Twenty-four enlisted Soldiers agreed to participate in this evaluation. All Soldiers were male and awaiting special training (e.g., Airborne). The Soldiers were from the following Combat Specialties: Infantry (11B; $n = 18$), Indirect Fire Infantryman (11C; $n = 2$), Special Forces Candidate (18X; $n = 2$), and M1 Armor Crewman (19K; $n = 2$). Their ages ranged from 18 to 26 years (mean = 19.8). Six reported at least some college experience. All had served in the Army from 4 to 6 months. None reported any injury of any type; all reported they could pass the Army physical fitness test.

All but three Soldiers reported wearing their watch on their left hand (one reported the right hand, two did not answer). All reported English as their first language. When asked which region of the country described their speech, 4 reported Southeast, 5 reported Southwest, 10 reported Midwest, 4 reported Northeast, and 1 reported North West. Only two reported previous experience with robots, which were a) manipulator arm on firetruck, and b) commercial quadcopter. Eighteen reported experience with remote control vehicles. Eight reported experience with speech control capabilities, and reported fair-to-positive experience regarding effectiveness of use.

3.2 Training Evaluation

Soldiers provided feedback after each training session and speech, gestures, and advanced speech command iteration. Table 2 shows the mean ratings, based on a 7-point scale ranging from 1 = extremely ineffective/negative to 7 = extremely effective/positive, and standard deviation (SD).

Table 2 Soldier-based training ratings

Aspect	N	Speech mean (SD)	Gesture mean (SD)	Advanced mean (SD)
Overall effectiveness of training for commands	24	6.2 (0.7)	5.9 (1.1)	6.5 (0.6)
Hands-on training	24	6.4 (0.6)	6.2 (0.9)	6.6 (0.6)
How prepared did you feel for using the device	24	6.3 (0.8)	6.0 (1.0)	6.6 (0.6)
How easy was it to learn and remember commands	24	6.8 (0.5)	6.1 (1.1)	6.4 (0.9)
How well do you think you performed end of training?	24	5.8 (1.0)	5.4 (1.1)	6.3 (0.9)
How well do you expect to perform robot control tasks	24	5.8 (1.0)	5.4 (1.3)	6.3 (1.1)
Trainer rating of trainee	20	4.8 (0.4)	4.4 (0.7)	4.8* (0.4)

*N=15

Ratings of how well each Soldier thought they were going to do in speech or gesture were not significantly correlated with how well they performed; however, their expectations of performance for speech and for gesture were significantly related ($r = 0.771$, $p < 0.001$), suggesting a measure of generalizable confidence and/or humility (e.g., if a Soldier thought they were going to perform very well in speech, they were very likely to think they were going to perform very well in gesture, and vice versa. Soldier comments regarding training can be found in Appendix C.

Trainer-based ratings were also collected to compare to Soldier self-ratings, with regard to how well each participant performed in training. This was collected in order to 1) establish construct validity of trainee performance that can 2) be related to measures of performance. Trainers rated each trainee on a 5-point scale. Trainees were not aware of these trainer ratings. Trainer ratings of Soldier training performance were not correlated to self-ratings of performance (for speech, $r = 0.024$, $p = 0.91$; for gesture, $r = 0.034$, $p = 0.87$). However, this is a function of low variance; the ratings by the trainer only ranged from 4 to 5.

Soldiers also provided 7-point ratings on the reliability of each command, as experienced in training. Ratings ranged from 1 = extremely unreliable to 7 = extremely reliable, as shown in Table 3. Comments regarding reliability are provided in Appendix C.

Table 3 Feedback on command reliability

Command	N	Speech mean (SD)	Gesture mean (SD)
Move forward	24	6.5 (0.7)	6.0 (1.0)
Move backward	24	6.5 (0.7)	4.0 (2.1)
Turn left	24	6.0 (1.4)	5.3 (1.3)
Turn right	24	6.3 (0.9)	4.8 (1.7)
Follow me	24	5.7 (1.4)	5.7 (1.2)
Stop	24	5.5 (1.6)	6.2 (1.0)
Commands in general	23	6.1 (0.9)	5.3 (1.3)

3.3 Observer-Based Performance Measures

Measures of performance were based on observations regarding outcomes to each command in terms of responses to commands when the commands were correctly executed by the Soldier. Error in response included the following:

- **Incorrect:** An incorrect response from the robot when a command was given and the response was wrong (i.e., command the robot to go forward and it turns left instead).
- **No response:** No response from the robot was when a command was given and acknowledged by the robot, but the robot did not perform the task.
- **Communication error:** The robot did not acknowledge receiving the command given by the Soldier.

Errors due to incorrect response. Table 4 provides the percent of errors for each command. The sum of all attempted commands is also provided. The number of attempts varied across Soldiers and commands, as Soldiers were requested to try at least three times to get the correct response. The nature of the evaluation made it so the “stop” command was given more often, as it typically followed other commands to make the robot discontinue its actions. Speech commands were interpreted extremely well (e.g., 0.00 indicates no incorrect responses). Paired t-tests show the speech modality was associated with significantly less error than the gesture modality for the majority of commands. However, gesture commands were also highly accurate, particularly for “forward” and “follow”.

Table 4 Mean number incorrect response errors (N = 24)

Command	Sum speech ^a	% Speech incorrect	Sum gesture ^a	% Gesture incorrect	Paired t-test (p)
Forward	96	3.1	93	8.6	0.170
Backward	79	0.0	160	46.3	0.000 ^b
Turn left	102	0.0	99	17.2	0.007 ^b
Turn right	96	0.0	129	34.9	0.001 ^b
Stop	385	0.0	460	7.4	0.021 ^b
Follow	83	0.0	76	13.2	0.30 ^b

^a Total number of command attempts^b Differences significant at $p \leq 0.05$

No response to commands. Table 5 provides the percent of non-responses for each command. Differences were generally not significant, with the exception of “backward”. It is clear that the incidence of no response was very low for both speech and gesture.

Table 5 Mean number of no response errors (N = 24)

Command	Sum speech ^a	% Speech no response	Sum gesture ^a	% Gesture no response	Paired t-test (p)
Forward	96	13.5	93	4.3	0.224
Backward	79	1..3	160	17.5	0.013 ^b
Turn left	102	30.4	99	9.1	0.059
Turn right	96	21.8	129	11.6	0.485
Stop	385	24.6	460	12.0	0.259
Follow	83	38.8	76	25.0	0.110

^a Total number of command attempts^b Differences significant at $p \leq 0.05$

Communication errors. The percent of communication errors for each command (Table 6) describes error resulting from the robot not acknowledging having received the command. This reflects an error of system communication. Percentages were consistently low, with the exception of no audio responses to speech commands of “stop” and “follow”.

Table 6 Mean number of communication errors (N = 24)

Command	Sum speech^a	% Speech comm. error	Sum gesture^a	% Gesture comm. error	Paired t-test (p)
Forward	96	3.1	93	0.0	0.185
Backward	79	3.8	160	2.5	0.747
Turn left	102	2.8	99	1.0	0.328
Turn right	96	4.1	129	2.3	0.747
Stop	385	8.6	460	0.7	0.025 ^b
Follow	83	7.2	76	0.0	0.031 ^b

^a Total number of command attempts

^b Differences significant at $p \leq 0.05$

3.4 Post-Session Questionnaires

The following summarizes responses to questionnaires administered after performance with the SID controller. Each Soldier used SID, using commands with speech and gesture modalities.

Fit/comfort of SID watch controller. Each Soldier rated SID with regard to weight and comfort after each session.

Weight was rated using a 7-point semantic differential scale ranging from 1 = extremely heavy to 7 = extremely lightweight. The overall mean rating was 6.6 (SD = 0.65), indicating that subjects generally found the SID to be extremely lightweight.

Comfort was rated using a 7-point semantic differential scale ranging from 1 = extremely uncomfortable to 7 = extremely comfortable. The overall mean rating was 6.1 (SD = 0.99), indicating that subjects found the SID to very comfortable.

Ease of use was rated using a 7-point semantic differential scale ranging from 1 = extremely difficult to 7 = extremely easy. Results are summarized in Table 7. Ratings were generally higher for speech commands. For gesture commands, ratings were lowest for “move backward” (mean = 3.9) and highest for “stop” (mean = 6.2), indicating them to be the most and least difficult to perform, respectively. Ratings for “commands in general” were significantly higher for speech (paired-sample t-test = 4.44, df = 23, $p < 0.001$). Soldier comments regarding use of the SID can be found in Appendix C.

Table 7 Ratings of difficulty of commands, by speech and gesture

Difficulty of command (N = 24)	Speech mean (SD)	Gesture mean (SD)
Move forward	6.6 (0.6)	5.9 (1.1)
Move backward	6.7 (0.6)	3.9 (1.9)
Left turn	6.1 (1.6)	5.2 (1.6)
Right turn	6.4 (0.9)	4.7 (1.7)
Follow me	6.0 (1.3)	5.8 (1.3)
Stop	5.7 (1.7)	6.2 (1.1)
Commands in general	6.3 (0.9)	5.3 (1.1)

Reliability was rated using a 7-point semantic differential scale ranging from 1 = extremely unreliable to 7 = extremely reliable. Results are summarized in Table 8. Ratings were generally higher for the speech modality. For gesture commands, ratings were lowest for “move backward” (mean = 4.0) and highest for “stop” (mean = 6.2), indicating the least and most reliable, respectively. Ratings for “commands in general” were significantly higher for speech (paired-sample t-test = 3.042, df = 23, p = 0.006). Soldier comments regarding the reliability of the commands can be found in Appendix C.

Table 8 Ratings of reliability of commands by speech and gesture

Reliability of command (N = 24)	Speech mean (SD)	Gestures mean (SD)
Move forward	6.5 (0.7)	6.0 (1.0)
Move backward	6.5 (0.7)	4.0 (2.1)
Left turn	6.0 (1.4)	5.3 (1.3)
Right turn	6.3 (0.9)	4.8 (1.7)
Follow me	5.7 (1.4)	5.7 (1.2)
Stop	5.5 (1.6)	6.2 (1.0)
Gestures in general	6.1 (0.9)	5.3 (1.2)

Opinion survey. Soldiers used a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree) to respond to statements summarized in Table 9. For positive statements, agreement was generally high for both modalities, with the mean for speech commands (6.1) higher than for gestures (5.7). This difference, using paired comparison t-test, was statistically significant ($t = 3.087$, $df = 23$, $p = 0.005$). For negative statements, the mean for speech commands were lower than for gestures; however, the difference was not significant ($t = 0.995$, $df = 23$, $p = 0.330$). Overall,

responses indicated a general preference for speech over gesture modality. Soldiers agreed that the smart watch was a good concept for operations. Opinion survey comments can be found in Appendix C.

Table 9 Opinion survey ratings

Statement	Speech mean (SD)	Gesture mean (SD)
Positive Statements		
“It was easy to learn the commands for robot control”	6.9 (0.3)	6.5 (0.8)
“It was easy to execute the commands”	6.7 (0.6)	5.8 (1.4)
“The robot responded effectively to commands”	5.9 (0.8)	5.0 (1.4)
“Using this modality is a good concept for robot control”	6.3 (1.1)	5.6 (1.6)
“Commands helped keep my attention on my surroundings”	6.0 (1.0)	5.3 (1.5)
“Commands are useful for ‘heads-up’ ‘eyes-free’ performance”	5.5 (1.1)	5.5 (1.9)
“I trusted the robot to obey my commands”	5.5 (1.1)	5.0 (1.5)
“Using this modality is a good concept for Army dismount operations”	5.9 (1.0)	5.6 (1.5)
“The robot was very responsive to commands”	5.6 (1.0)	4.9 (1.6)
“The smart watch was comfortable to wear”	6.5 (0.9)	6.3 (1.2)
“The smart watch is a good concept for robot control”	6.2 (1.4)	6.5 (0.7)
Mean	6.1 (0.6)	5.7 (0.8)
Reverse statements		
“It was difficult to maneuver the robot using this modality ”	2.5 (1.7)	3.3 (1.9)
“Commands would be difficult to remember in operations”	2.3 (1.7)	2.9 (1.8)
“The robot was slow to obey”	4.5 (1.9)	4.8 (1.8)
Mean	3.1 (1.2)	3.5 (1.3)

3.5 End of Experiment Comparative Evaluation Questionnaire

Soldiers were asked to indicate which type of input they preferred for each command; results are summarized in the following Table 10. Command preference comments can be found in Appendix C.

Table 10 Summary count of system preference regarding command reliability

Reliability	Speech	Gesture	Same
Move forward	11	3	10
Move backward	16	3	5
Left turn	14	3	7
Right turn	15	3	6
Follow me	12	6	6
Stop	9	9	6
All commands in general	14	7	3

Soldiers were also asked, if the system worked perfectly for every command, how useful would the system be in operations, using a 7-point scale, ranging from 1 = extremely useless to 7 = extremely useful. The mean for speech was 6.33, and the mean for gesture was 5.62. Soldier comments regarding operational relevance can be found in Appendix C.

3.6 Advanced Speech Commands

Advanced speech command errors. Advanced commands were only given using the speech modality. Thus, there are no comparative analyses with the gesture modality. Table 11 provides a breakdown of errors.

Table 11 Advanced speech command errors

Advanced Speech Command	Sum correct response	Sum no response	Sum incorrect response	Sum communication error
Start recording route <name>	24	0	1	0
Stop recording route	27	0	1	0
Follow route <name>	25	0	0	0
Set rally point <name>	24	0	0	0
Go to rally point <name>	26	2	1	0
Move forward <x> meters	81	1	7	0
Turn left <x> meters	45	5	2	0
Turn right <x> meters	37	0	0	0

Results indicate a high degree of accuracy for these advanced commands. Errors were typically related to the command being misunderstood. Examples include, when the soldier said “turn left” and the acknowledgement was “turn right”, or the soldier said “10 meters” and the acknowledgement was “2 meters”.

Soldiers provided feedback on advanced speech commands. Ratings for each command are shown in Table 12. Ratings were based on a 7-point scale ranging from 1 = extremely negative to 7 = extremely positive. Ratings were generally positive, ranging from 5.67 (move faster/slower) to 6.63 (go to rally point). Soldier comments regarding advanced speech commands can be found in Appendix C.

Table 12 Ratings for each advanced speech command

Command N = 24	Rating mean (SD)
Move faster/slower	5.7 (1.3)
Turn left/right x degrees	6.1 (0.9)
Go forward/backward x meters	6.3 (1.0)
Set route	6.3 (0.9)
Follow route	5.9 (1.4)
Set rally point	6.5 (0.7)
Go to Rally Point	6.6 (0.6)

Overall experience (Table 13) ratings were based on a 7-point scale ranging from 1 = extremely negative to 7 = extremely positive. Soldier comments regarding their overall experience can be found in Appendix C.

Table 13 Overall experience ratings

User experience N = 24	Rating mean (SD)
Reliability – degree of overall reliability	5.8 (0.8)
Trustworthiness – degree you trust the system overall	5.5 (1.0)
Ease of use – degree of difficulty	6.4 (0.6)
Comfort/fit – degree system is comfortable over time	6.4 (0.7)
Operational relevance – degree system would be useful in operations	6.0 (1.2)
Maintaining situation awareness – degree system supports awareness of surroundings	5.8 (1.1)

3.7 NASA-TLX

Soldiers rated perceived workload using the NASA-TLX rating scales after each performance session. Table 14 shows the mean unweighted ratings for each NASA-TLX construct. Five have to do with aspects of workload, while performance reflects perception of self-efficacy; therefore, the constructs were not rolled into one overall measure.

Table 14 Mean NASA-TLX ratings for speech, gesture, and advanced speech

Aspect N = 20	Speech mean (SD)	Gesture mean (SD)	Advanced mean (SD)
Mental demand	2.3 (1.4)	3.8 (2.3)	3.2 (2.0)
Physical	1.3 (0.6)	2.5 (2.1)	1.6 (0.7)
Temporal	2.8 (1.6)	3.1 (1.9)	2.0 (0.8)
Effort	2.3 (1.8)	3.1 (2.1)	2.7 (1.9)
Frustration	2.3 (1.9)	4.0 (2.7)	1.9 (2.0)
Performance	8.1 (1.4)	7.1 (1.4)	8.5 (1.4)

Table 15 shows paired comparison t-tests for each workload construct and the self-efficacy construct, along with the criteria for Holm-Bonferroni correction for familywise error rates between speech and gesture control. The Holm-Bonferroni criteria are listed for each comparison. It can be seen that using this more conservative method, the NASA-TLX ratings for physical demand do not meet significance. The results indicate that workload was rated significantly higher for gestures, with regard to mental demand and frustration.

Table 15 Paired-comparison t-tests comparing speech and gesture for NASA-TLX

Gesture vs. speech	t	df	Sig. (2-tailed)	Holm- Bonferroni	Sig.
Frustration	-2.981	18	0.008	0.008	Yes
Mental	-2.933	18	0.009	0.01	Yes
Physical	-2.549	18	0.020	0.0125	No
Performance	2.500	18	0.022	0.0167	No
Effort	-1.073	18	0.297	0.025	No
Temporal	-1.035	18	0.315	0.05	No

NASA-TLX ratings of mental demand and frustration were analyzed for correlation to performance measures when using speech and gesture. Correlations were not significant; the only significant correlation was between ratings of frustration for speech and ratings of frustration for gesture, suggesting some individuals experience generalizable levels of frustration that did not relate to performance.

4. Conclusions and Recommendations

The objective of this study was to evaluate effective control of a robotic mule using SID. Results indicate Soldiers were able to successfully perform all control task using gesture, speech and advanced speech commands using SID. Soldier ratings indicate the commands were easy to learn, remember, and execute.

Data were collected and analyzed to answer the following questions:

How well does SID recognize inputs from Soldiers? Results showed that while both command methods were fairly accurate, the speech commands were particularly effective. There were errors resulting from no response and incorrect response from the robot when a command was given correctly for both types of commands. This suggests further improvements are needed in terms of pure recognition accuracy of the system. Smart watches do not seem to have high quality microphones, and seem to have trouble in noisy environments. This had a noticeable impact on speech recognition. The SMSS's engine revving during turns was problematic for the speech recognizer. The off-the-shelf speech recognizer used in this particular implementation of SID (the Android version of Think-a-Move's SPEAR recognizer) also did not work as well in noisy conditions with the SMSS and a smart watch as it had in earlier trials with an RGator robotic system and a smartphone microphone (Taylor et al. 2017). Using a more typical headset with a directional, noise-cancelling microphone would likely yield better results. Gesture recognition could also be improved. The gesture recognizer used the native inertial measurement unit on the smart watch, which sometimes had sensitivity to ambient arm motion, and generally had too fine a window in which gestures would be successfully recognized.

Can the Soldiers successfully use SID to control the robotic mule in a variety of maneuver tasks? Based on Soldier feedback and observations, SID was easy to learn and use for both gesture and speech commands. With few exceptions, all Soldiers successfully maneuvered the robot using all the gesture and speech commands.

Do the Soldiers find SID useful for controlling the robotic mule? Overall, Soldiers found SID to be useful for controlling the robotic mule. When Soldiers

were asked to indicate which type of input they preferred for each command, more preferred speech than gesture. Also, workload was rated significantly higher for gestures, with regard to mental demand and frustration. For ease of use, ratings were generally higher for speech commands.

What improvements to the system would the Soldiers find useful? Many Soldiers stated that the gesture command was too sensitive and had to be performed at a very precise speed and motion in order for the robot to recognize the command. Also, several commented that the “follow me” command was too similar to the forward command. When asked if the system is usable in its current form, Soldiers noted that the robot not receiving speech commands due to background noise and imperfect gesture recognition leading to either wrong commands or no commands received would be problematic in combat environments.

To summarize, fundamental findings are as follows:

The SID smart watch form factor is very promising. Soldiers rated this form factor very highly in terms of comfort, fit, and ease of use. Given Soldiers normally wear a watch, this form factor allows the Soldier to command semiautonomous systems without having to carry a separate controller device. This allows more hands-free operation and enables the Soldier to focus visual attention to his or her surroundings.

The SID speech commands were extremely reliable and effective. The results show great promise; however, further evaluations should be conducted in more high-noise environments, particularly when the source of noise is very nearby. In this evaluation, the vehicle was at least 30 m away for safety reasons, and no personnel were speaking other than the operator. The SID gesture commands were also very reliable and effective, in general. One command, Go Backward, was identified as needing refinement, as the gesture can be intuitively confused with Go Forward, as the hand and arm gesture is a forward movement, and would only be construed as “go backward” if the vehicle is directly in front of the operator. Soldiers indicated that the use of gestures could be highly relevant to combat operations, particularly in situations of high noise.

While time constraints prohibited further investigations of the use of speech and gesture together, the advantage of having both as command options is fairly self-evident. Speech and gesture commands are complementary in combat operations. Speech is the normal means of Soldier communications and is relatively hands- and eyes-free. However, effectiveness declines in conditions of high noise. Gestures can overcome this vulnerability through its independent channel of communications. However, it is imperative that the gestures be intuitive, easily accomplished, and easily distinguished. There should be an easy means of

switching modalities from speech to gesture. Further development should focus on refining the set of gestures. This may be accomplished most effectively through further refinements to smart watch capabilities, which in this proof-of-concept evaluation, was limited to those of a commercial off-the-shelf smart watch.

5. References

- Baraniecki L, Elliott L, Hartnett G, Riddle K, Pettitt R. An intuitive wearable concept for robotic control. Proceedings of the International HCI Human Computer Interface Conference; 2017 July; Vancouver, Canada.
- Elliott L, Hill S, Barnes M. Gesture-based controls for robots: overview and implications for use by Soldiers. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2016 July. Report No.: ARL-TR-7715.
- Elliott L, Skinner A, Vice J, Walker A. Utilizing glove-based gestures and a tactile vest display for covert communications and robotic control. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2014 June. Report No.: ARL-TR-6971.
- Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Human mental workload. Amsterdam (Holland): North-Holland Press, Elsevier Science Publishers; 1988. p. 139–183.
- Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 1979;6(2):65–70
- Pettitt R, Elliott L, Swiecicki C. Soldier-based concepts for squad level autonomous/intelligent robotic assets. Proceedings of the International HCI Human Computer Interface Conference; 2017 July; Vancouver, Canada.
- Swiecicki C, Elliott L, Wooldridge R. Squad-level Soldier-robot dynamics: exploring future concepts involving intelligent autonomous robots. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2015 Feb. Report No.: ARL-TR-7215.
- Taylor G. Natural, Multi-modal interfaces for unmanned Systems. Proceedings of the International HCI Human Computer Interface Conference; 2017 July; Vancouver, Canada.
- Taylor G, Quist M, Lanting M, Dunham C, Muench P. Multi-modal interaction for robotic mules. Paper presented at the SPIE Defense and Security: Unmanned Systems Technology XIX; 2017; Anaheim, CA.

Appendix A. National Aeronautics and Space Administration (NASA)-Task Load Index (TLX) Mental Workload Rating Scale

This appendix appears in its original form, without editorial change.

Approved for public release; distribution is unlimited.

Please place an “X” along each scale at the point that best indicates your experience with the robot controller you just used.

Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc)? Was the mission easy or demanding, simple or complex, exacting or forgiving?

Low | High

Physical Demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the mission easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low | High

Temporal Demand: How much time pressure did you feel due to the rate or pace at which the mission occurred? Was the pace slow and leisurely or rapid and frantic?

Low | High

Performance: How successful do you think you were in accomplishing the goals of the mission? How satisfied were you with your performance in accomplishing these goals?

Low | High

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low | High

Frustration: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during your mission?

Low | High

Appendix B. Questionnaires

This appendix appears in its original form, without editorial change.

Approved for public release; distribution is unlimited.

DEMOGRAPHICS

Roster # _____

Date _____

<u>MOS</u>	<u>RANK</u>	<u>DUTY POSITION</u>

Age _____ Gender (M) (F)

On which hand do you wear your watch? Left Right

1. Which area of the US best defines your speech (accent)? (check)

Northeast US	Atlantic Seaboard US	Southeast US	Midwest US	Southwest	Northwest
-----------------	----------------------------	-----------------	---------------	-----------	-----------

If not US, please specify country _____

2. Is English your first language? _____

3. What is your highest level of education/major?

4. How long have you served in the military? _____

5. Have you been deployed in a combat area? ☐ No ☐ Yes

If Yes, Describe duty position and theater of operation

4. Have you ever driven a robotic vehicle? ☐ No ☐ Yes

If Yes, Describe

5. Have you ever driven a remote control car? ☐ No ☐ Yes

If Yes, Describe

6. Have you ever used speech/voice activated control to operate a device?

☐ No ☐ Yes

If Yes, Describe

What is your usual experience with speech-based systems? (e.g., smartphone, telephone automation, etc.)

1	2	3	4	5
Very Inefficient				Very Efficient
Consistent Problems				No problems

Regarding above, what kind of system were you referring to?

6. Using the scale below, rate your level of experience with the following computer software and computer-related activities.

1	2	3	4	5	6	7
No experience	Below average	Slightly below average	Average	Slightly above average	Above average	Expert

	RESPONSE
a. Microsoft Windows 98, 2000, XP, etc.	
b. Computer-based games	
c. Army digital systems (e.g., FBCB2)	
d. I would self-rate my computer skills as:	

7. Using the scale below, please self-rate the following Knowledge, Skills, and Abilities (KSAs) related to military duties.

1	2	3	4	5	6	7
No experience	Below average	Slightly below average	Average	Slightly above average	Above average	Expert

	RESPONSE
a. Knowledge of tactics, techniques, and procedures (TTP)	
b. Knowledge of map reading and orientation in field setting	
c. Knowledge of reconnaissance, surveillance, and target acquisition procedures	
d. Robotic operations	
d. Knowledge relating to communications equipment and communications procedures	
e. Communication skills (ability to use communications equipment and face-to-face communications to enhance mission accomplishment)	

Training Evaluation of Smart Watch Speech and Gesture Controller

SPEECH

GESTURE

Advanced

Roster number _____ Date _____ Time _____

- 1. Training of commands.** Please indicate how well you were prepared for the tasks, using the following ratings.

1	2	3	4	5	6	7
Extremely Ineffective/ Negative	Very Ineffective/ Negative	Ineffective/ Negative	Neutral	Effective/ Positive	Very Effective/ Positive	Extremely Effective/ Positive

Overall effectiveness of training for commands	1	2	3	4	5	6	7
Hands-on training	1	2	3	4	5	6	7
How prepared did you feel for using the device?	1	2	3	4	5	6	7
How easy was it to learn and remember the commands?	1	2	3	4	5	6	7
How well do you think you performed in training at the end?	1	2	3	4	5	6	7
How well do you expect to perform robot control tasks?	1	2	3	4	5	6	7

Comments

- 2. Reliability of commands.** Please indicate degree of reliability for each command. Reliability indicates the degree to which the robot performed the action as commanded, upon (first) request.

1		2		3		4		5		6		7	
Extremely Unreliable		Very Unreliable		Unreliable		Neutral		Reliable		Very Reliable		Extremely Reliable	
Reliability of Command						1	2	3	4	5	6	7	
Move Forward						1	2	3	4	5	6	7	

Move Backward	1	2	3	4	5	6	7
Left Turn	1	2	3	4	5	6	7
Right Turn	1	2	3	4	5	6	7
Follow me	1	2	3	4	5	6	7
Stop	1	2	3	4	5	6	7
Commands in general	1	2	3	4	5	6	7

Comments

Evaluation of Smart Interaction Device (SID)

Gesture-based control

Roster number _____ Date _____ Time _____

Using the following scales, please rate the following aspects of using gestures for robot control

1. Fit / Comfort of SID watch controller (please circle number)

1	2	3	4	5	6	7
Extremely Heavy	Very Heavy	Heavy	Neutral	Lightweight	Very Lightweight	Extremely Lightweight

1	2	3	4	5	6	7
Extremely Uncomfortable	Very Uncomfortable	Uncomfortable	Neutral	Comfortable	Very Comfortable	Extremely Comfortable

2. Ease of use/Difficulty. Please indicate degree of difficulty/effort to execute each gesture correctly.

1	2	3	4	5	6	7			
Extremely Difficult	Very Difficult	Difficult	Neutral	Easy	Very Easy	Extremely Easy			
Difficulty of Gesture Command			1	2	3	4	5	6	7
Move Forward			1	2	3	4	5	6	7
Move Backward			1	2	3	4	5	6	7
Left Turn			1	2	3	4	5	6	7
Right Turn			1	2	3	4	5	6	7
Follow me			1	2	3	4	5	6	7
Stop			1	2	3	4	5	6	7
Gestures in general			1	2	3	4	5	6	7

Comments. (Ease of Use)

Reliability of gesture commands. Please indicate degree of reliability for each command. Reliability indicates the degree to which the robot performed the action as commanded, upon (first) request. (Please circle number in table)

1	2	3	4	5	6	7
Extremely Unreliable	Very Unreliable	Unreliable	Neutral	Reliable	Very Reliable	Extremely Reliable

Reliability of Gesture Command	1	2	3	4	5	6	7
Move Forward	1	2	3	4	5	6	7
Move Backward	1	2	3	4	5	6	7
Left Turn	1	2	3	4	5	6	7
Right Turn	1	2	3	4	5	6	7
Follow me	1	2	3	4	5	6	7
Stop	1	2	3	4	5	6	7
Gestures in general	1	2	3	4	5	6	7

If the system was unreliable (4 or below), please describe issues/problem

4. Opinion Survey. Using the scale below, please indicate whether you agree or disagree with the following statements. (Please circle number in table)

1	2	3	4	5	6	7
Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree

Statement	Rating						
"It was easy to learn the gesture commands for robot control"	1	2	3	4	5	6	7
"It was easy to execute the gesture commands"	1	2	3	4	5	6	7
"The robot responded effectively to commands"	1	2	3	4	5	6	7
"Using gestures is a good concept for robot control"	1	2	3	4	5	6	7
"Using gestures helped keep my attention on my surroundings"	1	2	3	4	5	6	7

"It was difficult to maneuver the robot using gestures "	1	2	3	4	5	6	7
"Gestures are useful for 'heads-up' 'eyes-free' performance"	1	2	3	4	5	6	7
"I trusted the robot to obey my gestures"	1	2	3	4	5	6	7
"Using gestures is a good concept for Army dismount operations"	1	2	3	4	5	6	7
"Gestures would be difficult to remember in operations"	1	2	3	4	5	6	7
"The robot was very responsive to commands"	1	2	3	4	5	6	7
"The robot was slow to obey"	1	2	3	4	5	6	7
"The smartwatch was comfortable to wear"	1	2	3	4	5	6	7
"The smartwatch is a good concept for robot control"	1	2	3	4	5	6	7

Comments

Evaluation of Smart Interaction Device (SID)

Speech-based control

Roster number _____ Date _____ Time _____

Using the following scales, please rate the following aspects of using Speech for robot control

2. Fit / Comfort of SID watch controller (please circle number)

1	2	3	4	5	6	7
Extremely Heavy	Very Heavy	Heavy	Neutral	Lightweight	Very Lightweight	Extremely Lightweight

1	2	3	4	5	6	7
Extremely Uncomfortable	Very Uncomfortable	Uncomfortable	Neutral	Comfortable	Very Comfortable	Extremely Comfortable

3. Ease of use/Difficulty. Please indicate degree of difficulty/effort to execute each Speech correctly.

1	2	3	4	5	6	7			
Extremely Difficult	Very Difficult	Difficult	Neutral	Easy	Very Easy	Extremely Easy			
Difficulty of Speech Command			1	2	3	4	5	6	7
Move Forward			1	2	3	4	5	6	7
Move Backward			1	2	3	4	5	6	7
Left Turn			1	2	3	4	5	6	7
Right Turn			1	2	3	4	5	6	7
Follow me			1	2	3	4	5	6	7
Stop			1	2	3	4	5	6	7
Speech in general			1	2	3	4	5	6	7

Comments. (Ease of Use)

Reliability of Speech commands. Please indicate degree of reliability for each command. Reliability indicates the degree to which the robot performed the action as commanded, upon (first) request. (Please circle number in table)

1	2	3	4	5	6	7
Extremely Unreliable	Very Unreliable	Unreliable	Neutral	Reliable	Very Reliable	Extremely Reliable

Reliability of Speech Command	1	2	3	4	5	6	7
Move Forward	1	2	3	4	5	6	7
Move Backward	1	2	3	4	5	6	7
Left Turn	1	2	3	4	5	6	7
Right Turn	1	2	3	4	5	6	7
Follow me	1	2	3	4	5	6	7
Stop	1	2	3	4	5	6	7
speech in general	1	2	3	4	5	6	7

If the system was unreliable (4 or below), please describe issues/problem

4. Opinion Survey. Using the scale below, please indicate whether you agree or disagree with the following statements. (please circle number in table)

1	2	3	4	5	6	7
Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree

Statement	Rating						
"It was easy to learn the Speech commands for robot control"	1	2	3	4	5	6	7
"It was easy to execute the Speech commands"	1	2	3	4	5	6	7
"The robot responded effectively to commands"	1	2	3	4	5	6	7
"Using speech is a good concept for robot control"	1	2	3	4	5	6	7
"Using speech helped keep my attention on my surroundings"	1	2	3	4	5	6	7
"It was difficult to maneuver the robot using speech "	1	2	3	4	5	6	7

“speech are useful for ‘heads-up’ ‘eyes-free’ performance”	1	2	3	4	5	6	7
“I trusted the robot to obey my speech”	1	2	3	4	5	6	7
“Using speech is a good concept for Army dismount operations”	1	2	3	4	5	6	7
“speech would be difficult to remember in operations”	1	2	3	4	5	6	7
“The robot was very responsive to commands”	1	2	3	4	5	6	7
“The robot was slow to obey”	1	2	3	4	5	6	7
“The smartwatch was comfortable to wear”	1	2	3	4	5	6	7
“The smartwatch is a good concept for robot control”	1	2	3	4	5	6	7

Comments

Appendix C. Soldier Comments

This appendix appears in its original form, without editorial change.

Approved for public release; distribution is unlimited.

Comments regarding training

SPEECH
I felt the commands given was received and performed in a timely manner from the moment I gave a command to the watch until the robot performed the command
The robot/watch was easy to use and effective.
Responded to commands quickly
Couldn't get robot to use follow me command effectively
It was nice. It responded to every command except left and difficulty with my accent as we don't pronounce the "t".
Some commands need to be said louder than others
Different ways of saying the same command would be useful.
GESTURE
The training was very simple and easy to use.
Much easier than the verbal commands
Trade reverse gesture command with follow me
When watch is loose it has a hard time picking up commands.
Commands were delayed.
There's a delay that is a problem in my opinion the delay affects the performance.
Forward and follow should be a little more separate from the similarity they have in the gestures used.
A few gestures were confused with others.
Turns need to be more defined, maybe change from the starting position between left/right.
ADVANCED SPEECH
The advanced speech makes the robot easier and far more usable.
Was effective

Comments regarding command reliability

SPEECH
I feel confident the commands given will be performed effectively majority of the time.
Sometimes it would not pick up my voice, however it would after repeating it.
Worked overall well but there's still a delay with the echo before the robot responds.
GESTURE
Move backwards did not work.
Move backwards was a motion it would not pick up.
Felt like forward and follow me was getting mistaken during some commands
Sometimes it would mistake "stop" for "turn right", other than that, effective.
Need's specification on speed of movements and such.
"Move backwards" command is hit or miss at the moment.
First page comments.
Reverse, follow me, and left were mixed up at times.

Commands were too sensitive to slight movements not allowing the command to be read.
It mixed up commands. Commands weren't received.
ADVANCED SPEECH
Very easy to use.
Very effective and simple to use.
Responds to most very well, the turns take two tries most of the time for me.
Advanced commands works well.
Robots memorization of exact recorded route was a little off, other than that, no issues with advanced speech.
All commands were heard and carried out. Background noise affects the commands too much.
Reliable but slow.

Ease of use comments

SPEECH
Positive
Extremely easy for command using voice commands.
The use of the SID was very easy, and required little to no effort mentally or physically to perform task.
It worked well and having it repeat your voice is a good touch.
Easy.
Suggestions/Comments
Make it able to turn left by saying either left turn or turn left, or right the same way.
Might respond better during contact with a throat mic.
The only problem seemed to be the sensitivity of the watch. I.E., most background noise.
All the commands make sense and are fairly easy to remember. I had issues with getting response from "follow me".
I do not know if it was because of the speed of my command or my accent but follow me occasionally took a second try.
Your voice has to be relaxed and constant to initiate commands
Sometimes the watch would not detect the voice commands. But it does work.
Would not respond to stop.
The "stop" command was the most difficult to use. It seemed the watch wanted a very specific enunciation of the word or it just wouldn't register.
Very easy to use unless you have an accent.
The watch could use multiple ways of taking the one command
Hard to stop in specific direction with delay.
GESTURE
Positive
The device is easy to pick up and learn in a matter of moments,
Commands were easy to pick up and remember.
As long as commands were crisp it worked great.

Suggestions/Comments
Had some issues with right turn gesture, but all other gesture commands respond well.
Reverse seemed difficult
The turning motions are slightly awkward.
The full extension of my arm seemed weird. Would make it difficult to give commands while seeking cover or concealment.
Move backwards wouldn't work
It got almost all of my signals, it only messed up on turns and it only did that twice; Only downside may be the amount of space required for signals.
Some gestures such as stop, left, and right were not reading or reading as incorrect commands. Felt like angle of hand has a lot of impact on reading of gestures by robot.
"Follow me" was mistaken for "move forward" quite often.
Commands seemed delayed.
There were certain speeds at which you have to do the gesture for it to work and all of the time when you rest your hands it calls out random gestures. I also noticed every time after a follow me command the very next command is usually wacked out and non-compliant.
"Move backwards" is somewhat hit or miss.
Follow me and forward are similar enough to where the computer/device takes the wrong command at times. Move backward rarely worked and was confused by the device as follow me too many times.
However, it will take much practice to master.
Left and right turns need to be more distinct from each other.
The commands were too sensitive to the precise motion. In a combat situation you do not have time to get exact motion.

Reliability of command comments

SPEECH
The "stop" command had some minor issues processing the command depending on distance, and noise.
It only responded to follow me once out of many times. Stop was difficult but could've been due to rain. Left turn had a slight challenge.
Only unreliable when background noise is present.
Stop-had trouble picking up or receiving the command; In general-There seemed to be a delay
Follow me was difficult to get it to register it took multiple times each time to get it to confirm, everything else was great.
Would not respond to stop when turning.
The command stop was difficult at times.
Follow me and move forward are unreliable only sometimes
It's just semi-slow

GESTURE
Right turn only seem to work when it wants to, and move backwards will work even from different gesture.
"Right turn" gesture would have issues, and would sometimes perform other commands.
Gesture for reverse was often mistaken for forward. Though this could have been user error.
Move backwards wouldn't work.
It would tend to do a left turn.
Would not recognize move backwards.
Follow me wouldn't work half the time.
One turns it tended to not register command or would confuse commands.
Didn't want to turn right, mixed up follow me and reverse.
Gesture had to be within very specific parameters.
I had no problems, the turns got confused twice but that's all.
Move backward always came out as follow me.
"Follow me" was often mistake for "move forward".
Move backward got mixed up with follow me, same but reversed for follow me.
Moving backwards didn't really register, but it sometimes would when I did the "Stop" gesture.
Delay between commands, when your hand is at a resting state it starts calling things off and moving backwards is a pain.
Reverse would not get picked up most of the time. Follow me was picked up by half the other commands.
Unreliable commands weren't read. Commands were mistaken for others.

Opinion survey comments

SPEECH
I feel this robot would be of great use for military operations, and very easy to use.
Something more durable and reliable. Would be better form combat uses.
I feel that some kind of helmet mount to speak into would help make it easier in a battle field condition.
It responded effectively to every command except "follow me" and the only other problem I had was I found myself having to press the speech button multiple times because it would flicker instead of activate until I took my finger off.
Although it did not pick up my voice 100% of the time, the watch/robot combo is an effective tool.
Very responsive but slow to react.
Still a delay from command to execution.
The smartwatch needs to be more tactically made and it would be useful to use the robot for mine sweeping and enemy detection/reconnaissance as well.
Push to talk might need to be a button you can use without looking at it, the same way you just feel a fire selector switch on a rifle. Using the watch seems pointless if you need the other hand to operate the watch.

There should be away to key it into a radio so you don't need to look down at the watch.
There was a large delay from the command to the movement.
GESTURE
With the gestures you have to think for a second. And you have to make sure you execute gesture properly. I think voice worked better.
A glove might be better for gesture control. Also, maybe a throat mounted mic for voice controlled operations.
The response time is decently slow, and in high stress environments it may be difficult to remember commands. But it was very responsive and easy to use. I could tell before I got conformation whether or not the motion was captured correctly, it was a breeze to use.
The watch felt loose and needs to be secure to successfully use gestures for commands. Robot was very slow to use commands.
The watch was useful and effective to control the robot.
I think that using gestures is more effective on a base for logistical operation rather than combat. Voice command in my opinion works better.
I believe a smartwatch that goes into the field should be more shockproof, waterproof, etc.
Stop and forward are easier to remember, reverse only seemed to work half the time.
Faster response time from the robot. Gestures should more closely follow normal *unknown* hand signals.
Gestures could get confusing come gestures mean different things when controlling squads and the members could think your gestures is for them and not the robot.

Advanced speech command comments

Advanced Speech Commands
Positive
Defining "route" and "rally point" was very useful and worked great.
Everything worked as it was supposed to I had no problems at all.
Very useful, the advanced commands made it easy to use
I think it's genius and the capabilities in the logistical aspect are good along with the combat potential for IED sweeps having robot potentially ride in front scanning and also the ease of the watch which keeps you combat effective.
Issues/suggestions
Everything worked well for me except for the command "stop" while already in motion (too loud). I feel like this could be a problem in a combat situation where the surroundings are much louder.
Did not do the move faster/slower command while testing.
Had trouble remembering the route.
The robot could not repeat its route. The robot completed half of the route.
When saying move 10 meters forward it moved 2 repeatedly.
When turning X degrees, the robot doesn't follow exact degrees. Ex. 45 degrees it would go about 55 degree - 60 degree, 180 degree about 240 degree.

Overshot on turn commands
Didn't test faster or slower
It turned way further then the command
Comments regarding other useful interactions
I prefer the speech over the gesture method. The gestured would be much more difficult to perform correctly with a full kit of combat gear/weapon.
I was able to tell it to follow a route I had not yet finished defining and it tried to follow it.
Maybe it could mark individuals in a squad or team and then be able to be moved from one to the other by voice command.
"Follow me x meters" causing the robot to follow a certain distance behind you.
Break contact to go into cover. Keep distance from point X by X meters.
Setting follow distance, face me/face NSEW
Be able to make it turn at any degrees and move any distance. Add feet
If we took contact, it immediately seeks cover

Comments on overall experience

What did you have problems with?
The degrees of turn movements weren't exact and gave some trouble maneuvering the robot at times, however was still very easy to operate and easy to adjust
Some commands did not work or the robot registered one command and mistook it for another.
Following routes, reading correct distances
I didn't like the wrist watch idea, something that I don't have to look at to use would be better.
none
I had no problems it responded to everything it was commanded
I had problems with the arm signals and the robots ability to retrace a route.
Some commands were not followed or interpreted correctly
Nothing
Giving commands while it was turning
The gesture commands weren't always registered.
Just the bot following the remote
Nothing
Gestures
The watch had to be reset, not picking up some commands the first time
Not hearing commands over ambient noise; imperfect gestures led to either wrong commands or no commands received.
Acknowledging and performing commands in timely manner. Getting commands confused with others. Mixing up rally points
Is the system usable in its current form? What changes would be necessary to make it useful?
Though the system is functional, its connection strength is still unstable.

I believe this system could be usable in current form for on base procedures. However if used in combat, this would still need some minor adjustments, but this system could benefit the military greatly.
I think the system is usable in its current form. Minor tweaks to gesture would improve it immensely though.
Yes.
Yes, only the changes to the problems previously stated.
It is usable, but not fully practical. I would add some sort of collision detection. Instead of using a watch the audio input system should be headphones with some sort of toggle for audio.
Not yet. Needs to be able to follow routes and go the distance given.
Yes, I think it could be used behind the front lines, but when it gets noisy there would be problems.
A more hands free device so you could focus on your weapon i.e., a headset.
This part I believe could be used, the voice recognition by itself needs more data to handle more accents but other than that I believe it was usable.
It is usable in its current form but it needs consistency to be worth deploying repeatedly.
No some fine tuning with speech and gestures needs to be done. Had to be restarted multiple times.
Yes it is usable, I think.
Yes, more reliable mike/communications device.
Make the gesture commands a little easier to register and I think it's usable.
Less delay from command to execution. It's currently unusable now in my opinion due to the delay, but the design is adaptable and I foresee in the near future great things from the project.
Yes, for light duty work, still needs to be more accurate with demands.
The system would need to be more accurate when it comes to degrees. Needs a way to be commanded to keep itself/hidden/or from harm. Should be able to detect enemies and perform reconosaince. Should be able to mine sweep. Should also be able to have a weapon to be commanded to direct cover fire/suppressive fire.
Speech is gestures seem like they need more work. One-handed use for speech. I don't want to completely put down my rifle to work a watch. Why not put the watch or similar system on a rifle somewhere.
Yes it's usable, just make sure to have 100% reliable from the watch.
Part of it. The voice operation needs to be able to hear commands over loud noises.
Not usable. Needs improvement.

Command preference comments

SPEECH
Much clearer response, but sensitive to sound.
I experienced less error when using the speech commands to operate the system.
Speech worked way better. Response was quick

Easily recognizes voice, can give accurate commands, less noticeable to others around.
The speech was easier to make a right turn than with the gestures.
Had a hard time picking up the voice command for "stop"
Speech commands worked most of the time and did not have much reliability issues.
It came easier and could be done quickly.
I didn't have to wave my arms around like I was waving down a plane.
Felt easier to communicate with.
Speech was mostly good, I had some issues with some of the commands "follow me" was an issue, other than that it was easy.
Follow me is a command that could easily be communicated and would most likely be used while not taking fire.
Most speech commands were easier for robot to read whereas many gestures were misinterpreted.
The turning signals did not work as well as the voice commands for turning.
I had very few time where it didn't follow my commands
Is amazing and it works well except left turn due to my accent
Left and right turn were not always accurate with hand movements so I preferred speech for these commands.
More clear to the machine
Speech was more reliable. During the gesture test the robot un-paused too fast and stopping it with speech was already instinctive.
Speech worked a lot better than gestures simply because it responded nearly every time with 100% accuracy when it heard the command
All speech commands were registered
Gesture commands would not acknowledge or would get confused with other commands
GESTURE
Quiet and reliable, easily read as different commands
The gesture commands were easy to use, but weren't always effective and would sometimes perform the wrong gesture given.
If gesture commands were performed slightly wrong it gave a different command. But overall a great concept.
Sometimes it doesn't recognize gestures. Can misinterpret gestures.
Would not recognize "backward movement" gesture.
Not reliable, nor practical in a combat situation.
Robot had trouble understanding gestures.
Seems like too much movement for small tasks and you have more specific control with the voice control.
More difficult to get a response.
The gestures all registered flawlessly, also it was the least taxing and it's a silent method of directions which is always good
Stop is a command that needs an instant reaction and is a command that needs to take as little thought as possible.

Way more responsive and easy to use, over voice commands
I prefer all commands in gesture form, robot was more responsive to me in this category
Move backwards never really registered with the robot.
A disaster mostly and difficult to execute. Easy to remember signals though.
The "stop" command when spoken did not always seem to work, but was much easier with the gesture command versus the speech command.
The gesture commands can be misinterpreted by the robot a lot.
The stop command was the most reliable.
The gestures commands needs to be worked through more so it follow squad hand signals more and pick up all commands with 100% accuracy.
Had several issues with the gestures and it misreading me.
Ease of access in real life scenario.
What did you have problems with?
left and right, some on gestures
Unresponsive to commands without repeating
The gesture commands gave myself the most trouble. However was still easy and convenient to operate system with.
Mainly gesture commands
Gestures for move backwards and turning
The right turns had been slightly problematic with the gesture feature.
Previously stated
Gestures didn't work all the time and sometimes robot did not recognize speech.
Turns with gestures.
Speech: it was hard for the robot to respond when the surrounding environment is too noisy. Gesture: feels like I'm trying to flag down a plane.
Gestures
Speech "follow me" command had issues
The gestures would not read consistently
The robot had a hard time with me saying stop at the reverse gesture was hard to get the robot to read, it took multiple attempts for it to finally read the gesture.
Sometimes the mic could not hear me while turning.
The stop command
Just "moving backwards" with gesture commands
Turn left speech/Gestures-move backward, left/right turn and stop.
The "stop" command when spoken
Move backwards and follow me
Gestures getting mixed up; It tried to follow a tree at one point.
The gestures would pick up different commands
Gestures
Command performance in effective time needed to be executed faster
Is the system usable in its current form? What changes would be necessary to make it useful?

This system can be very useful on the battlefield
Though the system is functional in its current form the responsivity needs to be strengthened.
The sensitivity of the microphone on the watch and possibly the loudness of the engine which would interfere with some voice controls depending on distance.
Yes.
Maybe the gesture controls more broad and fluid and have more specific commands with the voice feature.
Yes, no other changes (body armor)
See other survey
No, for speech it needs to pick up the voice more clearly regardless of weather. Waterproof the watch.
Noise reduction, equipment, smaller movements for gesture, application
Yes.
More trial for wider variety of accents for speech control.
Yes, but only if it becomes a little more crisp.
No, needs lots of fine tuning to reading speech and hand gestures. It is useful, just needs to read and act commands clearly and quickly.
Quieter engine/better mic
Yes, better communications device.
Gesture commands just needs to be sharper, but yes I feel it's usable.
Get rid of gesture, in my opinion, and stop delay from command to execution.
It is usable for missions with low chance of combat.
No, it needs to be made more for tactical scenarios, and have more tactical use.
Make speech one handed/hands free
The speech system is the gestures needs to be worked on.
No, Less perfect gesture reading. Being able to head commands over loud noises
No, needs to be quieter
This system can be very useful on the battlefield

Operational relevance comments

SPEECH
This would be useful when at a base. As well as long marches.
Very useful
May be hard for robot to understand speech in high volume areas
This would be beneficial in a tactical environment.
GESTURE
Radio commands
Extremely useful.
Might be hard to perfectly enact gestures in certain situations or during combat if speech cannot be heard by watch.

Needs work, good concept.
Commands could be confused for troop commands.
If in the prone, commands will be hard to give.
Can you think of any other way of communicating to a robot that could be helpful?
Headset but still by speech.
Possibly inputting simple commands on a controller size of watch or wristband so gestures or speech cannot be misinterpreted just tap a button and have degrees and distance available.
Selecting commands off a touch screen.
Through a headset.
A keypad so that it's quieter.
Laser, like a cot. Or making a head trail by same means from the sky to have bot follow route - just random ideas.
No
Touchscreen/buttons on watch.
For a squad leader or officer, using the radio he's already carrying. Follow the watch or tablet instead of a person.
Through com in headgear.

List of Symbols, Abbreviations, and Acronyms

LIDAR	Light Detection and Ranging
NASA	National Aeronautics and Space Administration
SD	standard deviations
SID	Smart Interaction Device
SMSS	Squad Mission Support System
SPSS	Statistical Package for Social Sciences
TLX	Task Load Index
TTP	tactics, techniques, and procedures

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

2 DIR ARL
(PDF) IMAL HRA
RECORDS MGMT
RDRL DCL
TECH LIB

1 GOVT PRINTG OFC
(PDF) A MALHOTRA

1 ARL
(PDF) RDRL HRB B
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL
35898-7290

8 ARL
(PDF) SFC PAUL RAY SMITH CTR
RDRL HRO COL H BUHL
RDRL HRF J CHEN
RDRL HRA I MARTINEZ
RDRL HRR R SOTTILARE
RDRL HRA C A RODRIGUEZ
RDRL HRA B G GOODWIN
RDRL HRA A C METEVIER
RDRL HRA D B PETTIT
12423 RESEARCH PARKWAY
ORLANDO FL 32826

1 USA ARMY G1
(PDF) DAPE HSI J WARNER
300 ARMY PENTAGON
RM 2C489
WASHINGTON DC 20310-0300

1 USAF 711 HPW
(PDF) 711 HPW/RH K GEISS
2698 G ST BLDG 190
WRIGHT PATTERSON AFB OH
45433-7604

1 USN ONR
(PDF) ONR CODE 341 J TANGNEY
875 N RANDOLPH STREET
BLDG 87
ARLINGTON VA 22203-1986

1 USA NSRDEC
(PDF) RDNS D D TAMILIO
10 GENERAL GREENE AVE
NATICK MA 01760-2642

1 OSD OUSD ATL
(PDF) HPT&B B PETRO
4800 MARK CENTER DRIVE
SUITE 17E08
ALEXANDRIA VA 22350

ABERDEEN PROVING GROUND

12 ARL
(PDF) RDRL HR
J LOCKETT
P FRANASZCZUK
K MCDOWELL
K OIE
RDRL HRB
D HEADLEY
RDRL HRB AB
R PETTITT
RDRL HRB C
J GRYNOVICKI
RDRL HRB D
C PAULILLO
RDRL HRF A
A DECOSTANZA
RDRL HRF B
A EVANS
RDRL HRF C
J GASTON
RDRL HRF D
A MARATHE