Improving Reliability Awareness to Support Appropriate Trust and Reliance on Individual Combat Identification Systems

Lu Wang
University of Virginia
Charlottesville, VA

Greg A. Jamieson
University of Toronto
Toronto, ON

Justin G. Hollands
Defence R & D Canada
Toronto, ON

Individual combat identification (CID) systems have been developed to help soldiers distinguish friends from foes in combat situations. However, these automated systems are not perfectly reliable. Previous studies have found that participants often do not rely properly on such systems and consequently their identification performance was not improved by them. We present an experiment that tested the effectiveness of providing aid reliability information to support participants' appropriate trust in and reliance on a CID aid. The results indicated that participants had difficulty in estimating the aid reliability. Participants who were not informed of the aid reliability trusted in and relied on the aid feedback less than those who were aware of the aid reliability. Providing the aid reliability information led to more appropriate reliance on the aid. This research has implications for the design of interfaces for individual CID systems and the training of infantry soldiers.

INTRODUCTION

Individual Combat Identification Systems

Individual combat identification (CID) systems perform an interrogation/response process to confirm whether an unknown target is friendly or not (Sherman, 2000). To interrogate a target, a soldier first manually activates the process. Then, if a correct electronic reply is received from the target, his CID system will give a 'friend' feedback. Otherwise, the target will be deemed 'unknown'. Because electronic communication can occasionally fail in the chaotic battlefield, these systems cannot always recognize friendly soldiers.

Several experimental studies have been conducted to examine whether humans could use these imperfect CID systems to improve their identification performance (Dzindolet, Pierce, Beck, Dawe, & Anderson, 2001a; Dzindolet, Pierce, Pomranky, Peterson, & Beck, 2001b; Karsh, Walrath, Swoboda, & Pillalamarri, 1995; Kogler, 2003). In these studies, participants sometimes under-relied on (i.e., disuse) a reliable CID system (Karsh et al., 1995) and sometimes over-relied on (i.e., misuse) the feedback from an unreliable system (Dzindolet et al., 2001a, 2001b). As a result, their identification accuracy was not improved by the automation unless the automation was perfectly reliable or their manual performance was severely deficient (Kogler, 2003). Researchers suggested that the suboptimal use of the CID systems was likely caused by the participants' inappropriate trust in these automated systems (Dzindolet et al., 2001a). However, since trust was not recorded in these previous studies, it is impossible to determine whether the improper reliance was attributable to inappropriate trust or not.

Trust in Automation

Lee & See (2004) define trust as, "the attitude that an agent will help achieve an individual's goal in a situation characterized by uncertainty and vulnerability" (p. 54). Trust in automation mediates the relationship between users' belief about the automation capability and their reliance on the automation (Lee & See, 2004). On the one hand, trust determines, in part, users' strategies to use the automation. On the other hand, trust itself largely depends on users' perceptions of the capability of the automation (Sheridan & Parasuraman, 2006). Hence, trust in and reliance on an automated system is more likely to be appropriate when information about the automation's capability is available.

The two objectives of this experiment are: first, to examine the effectiveness of using aid reliability

10.1518/107118108X350366

Crown Copyright

information to support appropriate trust in and reliance on the aid; second, to scrutinize the relationships among the participants' belief about the aid reliability, their trust in the aid, and their reliance on the aid.

METHODS

Participants

Twenty-six students with normal visual acuity from the University of Toronto were recruited. Complete data were collected from twenty-four participants and used for analysis. Each participant was paid $30 CAD for their participation, and a bonus $10 CAD was given to the performer with the highest CID accuracy.

Apparatus

The experimental simulation was modified from a commercial shooter game. Figure 1 shows a screenshot of the participants' view. This simulation was installed on two Dell OptiPlex GX270 desktop computers with 20-in UltraSharp 2000FP flat panel monitors. For each participant, only one computer and one monitor were used.

Figure 1. Participant's view in simulation

Experimental Design

A 3 (aid reliability: no aid, 67% reliability, and 80% reliability) x 2 (reliability awareness: uninformed, informed) mixed design was employed. The aid reliability was manipulated within participants. In the no aid condition, participants conducted the experimental task manually. In the two aided conditions, participants were assisted by a CID aid that could provide two types of feedback – 'friend' and 'unknown'. The 'friend' feedback was always correct while the 'unknown' feedback was fallible. The reliability of x percent means that when the aid sends out an 'unknown' feedback, x percent of the time it correctly identifies a hostile target. The reliability awareness was manipulated between participants. Only the informed group was informed of the reliability of the 'unknown' feedback before they started an aided condition. In contrast, the uninformed group only knew the reliability was less than 100%.

Procedures

This experiment took approximately 2.5 hours to complete. Each participant first went through a vision test, signed a consent form and filled out a demographic survey. The participant was then asked to read the instructions and observe the pictures of friendly and hostile targets. After that, the participant went through a training session (60 trials, 10 seconds per trial). Only one target would appear in each trial. The participant was asked to identify the target in the scene and shoot it if it was hostile and hold fire if it was friendly. After a target was killed or a trial ended, a screen would pop up to ask the participant to indicate their level of confidence in their engagement decision. Then an experimenter would inform the participants the correct identity of the target.

After training, participants completed three mission blocks of varying aid reliability. The order of the reliability conditions was counterbalanced across the participants in each reliability awareness group. The participants' task was the same as in the training session except that the experimenter would not inform them of the target identity after each trial.

All of the participants were informed that in the 120 trials of each mission block half of the targets were friendly and the other half were hostile. The order of target identity was randomized.

Participants were advised that they would have a CID aid to assist them in two of the three blocks. The aid would respond when the weapon was pointed at a target. When the aid identified a friendly solider, it would give 'friend' feedback – a blue light. Otherwise it would give 'unknown' feedback – a red light. All of the participants were made aware that while the 'friend' feedback would always be correct, the

Crown Copyright

'unknown' feedback was fallible. But only the participants in the informed group were told the failure rate of 'unknown' feedback in the aided mission blocks. At the end of the two aided mission blocks, participants filled out a trust questionnaire that assessed their trust in the whole aid and each type of aid feedback. For the participants who were not informed of the aid reliability, one additional question asking them to estimate the failure rate of the 'unknown' feedback.

Measures

Belief. The uninformed group's belief about the aid reliability was measured using their estimate of the 'unknown' feedback failure rate. The informed group was assumed to have a correct belief about the aid reliability.

Trust. Trust in the whole aid, the 'unknown' feedback, and the 'friend' feedback was measured using 7-point scales. Most of questions were extracted from an empirically validated trust questionnaire (Jian, Bisantz, & Drury, 2000).

Reliance. Participants' reliance on the fallible 'unknown' feedback was measured using the shift of their response bias from the no aid condition to the 'unknown' feedback trials in an aided condition (i.e., the trials in which participants received 'unknown' feedback). The assumption is that the more the participants rely on the 'unknown' feedback, the more liberal they would become when they receive 'unknown' feedback. Similar measures of automation reliance were used in several previous studies (e.g., Maltz & Shinar, 2003). The appropriateness of the response bias was assessed by comparing the participants' response bias shift with the optimal shift. For example, the optimal shift of response bias from the no aid condition to the 'unknown' feedback trials in the 67% reliability condition could be defined as the difference of optimal response bias between these two settings (see Equation 1).

$$
\begin{aligned}
&\textit{Optimal } \ln\beta \textit{ Shift} \\
&= \ln\beta_{optimal\_noaid} - \ln\beta_{optimal\_67\%reliability} \\
&= \ln\frac{P(Friendly)}{P(Hostile)} - \ln\frac{P(Friendly\,|\,'Unknown')}{P(Hostile\,|\,'Unknown')} \\
&= \ln\frac{50\%}{50\%} - \ln\frac{33\%}{67\%} \\
&= .693 \qquad \text{.................. } \textbf{Equation 1}
\end{aligned}
$$

Performance. Performance was measured using the rate of false alarm errors (i.e., shoot when targets are friendly), the rate of miss errors (i.e., hold fire when targets are hostile) and average response time (RT) in each mission block. RT was recorded only for those trials that participants shot at a target, and no RT was recorded when participants held fire in a trial.

RESULTS

To increase the normality and stabilize the variances of the probability data, an arcsine transformation was applied to all the probability data (Winer, 1991). Effect size, *r*, was calculated for the significant effects.

Belief

A one sample *t*-test (2-tailed) revealed no significant difference between the uninformed group's estimate (M=38.17%) and the real failure rate (i.e., 33%) in the 67% reliability condition, $t(11)=1.52$, $p=.16$. However, participants' estimate in the 80% reliability condition (M=31.25%) was significantly larger than the real value (i.e., 20%), $t(11)=2.72$, $p=.02$, $r=.63$. In addition, a paired samples *t*-test (2-tailed) indicated that there was no significant difference between the estimates in the two aided conditions, $t(11)=1.63$, $p=.13$.

Trust

Whole aid. A 2 (aid reliability: 67%, 80%) x 2 (reliability awareness: uninformed, informed) ANOVA on the participants' trust in the whole aid revealed a significant main effect of aid reliability, $F(1, 22)=7.18$, $p=.01$, $r=.50$. Participants trusted the whole system more in the 80% reliability condition (M=4.30) than the 67% reliability condition (M=3.74). The other effects were not significant.

'Friend' feedback. Most participants indicated near-complete trust in the 'friend' feedback on the 7-point scale in both aided conditions (M=6.67). This result likely indicates that they understood that the 'friend' feedback was always correct.

'Unknown' feedback. The Wilcoxon Signed-Rank test and Mann-Whitney test were used to analyze participants' trust in the 'unknown' feedback due to the data's non-normality. Both the uninformed and informed groups trusted the 'unknown' feedback significantly more in the 80% reliability condition

Crown Copyright

than the 67% reliability condition: for the uninformed group, *z*=-2.23, *p*=.03, *r*=-.64, *M*=3.42 vs. 2.25; for the informed group, *z*=-2.00, *p*=.05, *r*=-.58, *M*=4.72 vs. 3.92. In addition, regardless of aid reliability, the trust in the 'unknown' feedback was consistently higher in the informed group than the uninformed group: for the 67% reliability condition, *U*=19.50, *p*<.01, *r*=-.63, *M*= 3.92 vs. 2.25; for the 80% reliability condition, *U*=33.00, *p*=.02, *r*=-.48, *M*=4.75 vs. 3.42.

Reliance

As shown in Figure 2, we examined the shift in the response bias from the no aid condition to the 67% reliability condition, from the 67% reliability condition to the 80% reliability condition, as well as the shift from the no aid condition to the 80% reliability condition. The 3 (shift condition) x 2 (reliability awareness) ANOVA on lnß shift revealed that the informed group (*M*=0.43) changed their response bias more dramatically than the uninformed group (*M*=0.10), *F*(1, 22)=5.11, *p*=.03, *r*=.43. The other effects were not significant.

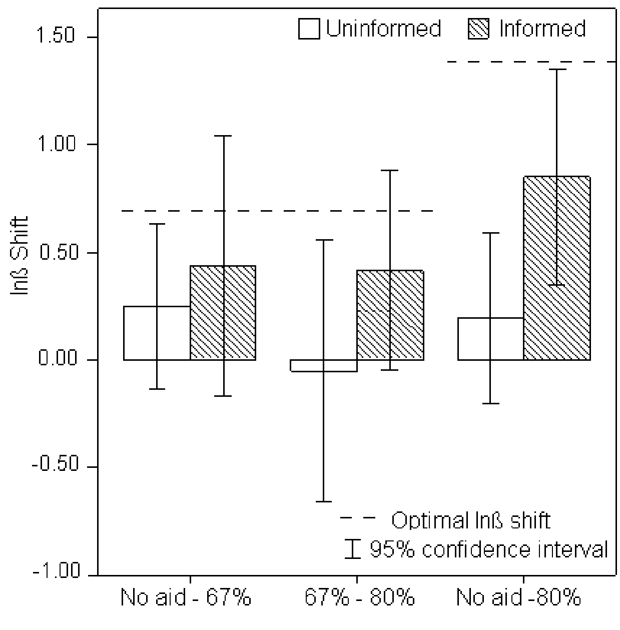

Figure 2. Response bias shift between conditions

The informed group's adjustment of response bias was not significantly different from the optimal value (.693) from the no aid condition to the 67% reliability condition or from the 67% reliability condition to the 80% reliability condition. However, their adjustment was significantly less than the optimal value (1.386) from the no aid condition to the 80% reliability condition, *t*(11)=-2.36, *p*=.04, *r*=.58. The uninformed group was worse than the informed group. Their shift in response bias was significantly less than optimal in all three shift conditions, *t*(11)=-2.57, -2.89 and -6.60, *p*=.03, .02, .00, *r*=.61, .65,.89, for no aid -67%, 67% - 80%, and no aid – 80%, respectively.

Relationship among Belief, Trust and Reliance

The relationship among participants' belief, trust and reliance were examined using the one-tailed Kendall's τ correlation analysis. Pooled over the two aided conditions for the uninformed group, the analysis revealed that participants' trust in the whole aid and trust in the 'unknown' feedback were both negatively correlated with their estimate of the 'unknown' feedback failure rate, *τ*(24)=-.27 and -.50, *p*=.39 and .00, respectively. However, participants' estimate of 'unknown' feedback failure rate was not significantly correlated with the participants' shift of response bias, *τ*(24)=-.11, *p*=.23. Pooled over the two aided conditions for both groups, the analysis indicated that participants' shift of response bias had a significant positive relationship with their trust in the 'unknown' feedback, *τ*(48)=.21, *p*=.03, but not the trust in the whole aid, *τ*(48)=.11, *p*=.14.

Performance

The 3 (aid reliability: no aid, 67%, 80%) x 2 (reliability awareness: uninformed, informed) ANOVA revealed a significant effect of aid reliability on the transformed false alarm rate, *F*(2, 44)=10.75, *p*<.01. Contrasts revealed a significant difference between the no aid condition (*M*=.24) and two aided conditions (*M*=.14), *F*(1,22)= 9.86, *p*<.01, *r*=.56 and a significant difference between the 67% reliability condition (*M*=.17) and the 80% reliability condition (*M*=.11), *F*(1, 22)=13.95, *p*<.01, *r*=.62. The other effects on false alarm rate were not significant. In addition, no significant difference was detected in the transformed miss rate or response time.

DISCUSSION

Although the estimate of the 'unknown' feedback failure rate and the reliance on the 'unknown' feedback was not correlated, these two factors were both significantly correlated with the trust in the 'unknown' feedback. These results support the

Crown Copyright

explanation that trust in an aid acts as an attitude mediating the relationship between the users' belief about the aid capabilities and their reliance on the aid (Lee & See, 2004). The uninformed group tended to underestimate the aid reliability and they could not differentiate the two reliability levels. This misbelief in the aid's reliability might have contributed to their different trust and reliance as compared to the informed group. The informed group generally relied on the 'unknown' feedback more appropriately than the uninformed group. However, even the informed group did not change their response bias enough from the no aid condition to the 80% reliability condition. This might be attributed to the 'sluggish beta' phenomenon which suggests that human response to the change of target probability is likely to be less than the ideal magnitude (Chi & Drury, 1998). This result may also suggest that factors other than belief could influence reliance on automation. Therefore, supporting correct belief is helpful but does not necessarily lead to the optimal reliance on automation (Beck, Dzindolet, & Pierce, 2007).

The current findings suggest that informing participants of aid reliability will be beneficial in generating appropriate reliance on CID systems. One area where these findings may provide a useful guide is in the design of information displays for CID systems to convey reliability information. Vendors could conduct field tests of the CID systems and determine the reliability levels in various situations. Another important area in which this research may be applied is in training soldiers to assess the system reliability by themselves. Soldiers may be able to better trust and rely on CID systems if they are alert to contextual factors that affect the system reliability.

REFERENCES

Beck, H. P., Dzindolet, M. T., & Pierce, L. G. (2007). Automation usage decisions: Controlling intent and appraisal errors in target detection task. *Human Factors, 46*(1), 50-80.

Chi, C., & Drury, C. G. (1998). Do people choose an optimal response criterion in an inspection task? *IIE Transactions, 30*(3), 257-266.

Dzindolet, M. T., Pierce, L. G., Beck, H. P., Dawe, L. A., & Anderson, B. W. (2001a). Predicting misuse and disuse of combat identification systems. *Military Psychology, 13*(3), 147-164.

Dzindolet, M. T., Pierce, L., Pomranky, R., Peterson, S., & Beck, H. (2001b). Automation reliance on a combat identification system. In *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 532-536). Minneapolis/St. Paul, MN: Human Factors and Ergonomics Society.

Jian, J., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53-71.

Karsh, R., Walrath, J. D., Swoboda, J. C., & Pillalamarri, K. (1995). *Effect of battlefield combat identification system information on target identification time and errors in a simulated tank engagement task.* (Technical report ARL-TR-854). Aberdeen Proving Ground, MD: Army Research Lab.

Kogler, T. M. (2003). *The effects of degraded vision and automatic combat identification reliability on infantry friendly fire engagements.* Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50-80.

Maltz, M., & Shinar, D. (2003). New alternative methods of analyzing human behavior in cued target acquisition. *Human Factors, 45*(2), 281-296.

Sheridan, T., & Parasuraman, R. (2006). Human-automation interaction. *Reviews of Human Factors and Ergonomics, 1,* 89-129.

Sherman, K. (2000). Combat identification system for the dismounted soldier. In *Proceedings of SPIE 2000: Digitization of the Battlespace V and Battlefield Biomedical Technologies II* (pp.135-146). Orlando, FL: the International Society for Optical Engineering.

Winer, B. J. (1991). In Brown D. R., Michels K. M. (Eds.), *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

Crown Copyright