Seeing is believing?: The effects of real-time, image-based feedback on emergency management decision-making

John M. McGuirl
Institute for Ergonomics
The Ohio State University
mcguirl.1@osu.edu

Nadine B. Sarter
Center for Ergonomics
University of Michigan
sarter@umich.edu

David D. Woods
Institute for Ergonomics
The Ohio State University
woods.2@osu.edu

ABSTRACT

Emergency management personnel often face feedback delays and a lack of reliable information. To address this problem, new information technologies have been developed that can provide real-time, image-based feedback. While potentially useful, this trend represents a fundamental shift in both the timing and format of the information used by incident commanders (ICs). Eight ICs took part in a simulation exercise to determine the potential impact of real-time imaging on their decision-making. Nearly all of the ICs failed to detect important changes in the situation that were not captured in the imaging but that were available via other, more traditional data sources. It appears that the ICs placed an inappropriately high level of trust in the imaging data, resulting in reduced data search activities and hypothesis generation. This research helps practitioners anticipate and guard against undesirable effects of introducing similar technologies on training and operational procedures.

Keywords

Decision-making, situation assessment, attention narrowing, trust miscalibration, emergency management

INTRODUCTION

The ability to gather relevant data in a timely fashion is paramount to properly assessing and responding to a crisis event. Unfortunately, emergency responders are often faced with both a scarcity and an uncertainty of information, as well as delays in feedback, particularly during the initial stages of an incident. Efforts to address these challenges often revolve around technologies that increase the amount and timeliness of data that is available to the decision-maker. However, previous research has shown that altering the amount, form, and timing of data that practitioners use can have unwanted side effects and introduce new vulnerabilities to the system they are meant to improve.

Some of these vulnerabilities include increased data management demands, the generation of new forms of errors, reduced adaptability, and a narrowing of data search activities and hypothesis generation (Billings, 1997; Norman, 1990; Sarter and Woods, 1992, 1995; Smith et al., 1997). The introduction of real-time image-based feedback into emergency management produces a point of change that offers a unique opportunity for analyzing how the cognitive work in that domain is conducted (Cook and Woods, 1996; Patterson et al, 2002). Our objective is to contribute to a better understanding of changes in the cognitive processes involved in this domain and to help avoid the potential breakdowns this new information source may create.

Challenges in incident command

Traditionally, emergency responders have relied heavily on either radio or face-to-face interactions for gathering information. These forms of verbal communication are supplemented with text information that is placed on various status displays and maps. There are several inherent limitations to verbal communication. Verbal communication is sequential, which means that it requires turn taking between the participants in order to establish the context, or "common-ground", that is needed to reduce ambiguities (Clark and Brennan, 1991). This context building requires time and continuous effort, particularly for non co-located team members, as the context changes in a highly dynamic environment. Also, verbal exchanges are transient in nature and do not afford a permanent record of the information. In addition, there are technical constraints, such as the limited availability of radio channels for communication. These limitations can foster the perceived need to obtain knowledge of the situation first-hand and to more effectively convey intent to subordinates (Danielsson and Ohisson, 1999).

In addition to its scarcity, feedback is often subject to delays, which can be the result of a) a lag in the execution of a given command, b) a lag in the response of the world to the action initiated by the command, c) a lag in the feedback mechanism, or d) some combination of the above**.** These delays pose an additional challenge to the exercise of control. In their study of fire-ground commanders, Brehmer and Allard (1991) demonstrated how even short feedback delays could greatly compromise an operator's ability to control a system.

Another consequence of delayed feedback is that it can lead to the creation of a "buggy model" i.e., incomplete and/or erroneous concepts of the situation, resulting in an inability to accurately predict how the situation may evolve (Johansson et al, 2002; Woods et al, 1994). This presents a challenge to the decision-maker whose expertise "is tuned to the future" (Woods, 2002). It is this need to anticipate events, rather than simply react to them, that often drives the IC's desire for more, and more immediate, feedback.

Supporting incident command

The previous discussion highlights the fact that in the course of an incident, the information requirements of the IC often outstrip the abilities of the available data channels to provide the desired feedback. As a result, ICs must develop their situation assessment with limited and uncertain information, and monitor its current state and plan for the future with information about the past. In response to this, several new communication and information technologies have been proposed to support his/her activities with the goals of providing the necessary information in a timely manner.

One technology in particular that has been proposed as a solution to these problems is the unmanned aerial vehicle (UAV), which can provide ICs with real-time, image-based feedback. By eliminating delays and providing a context-rich form of data, it could aid in assessing and comprehending the overall situation. The unique, overhead perspective afforded by the UAV would allow planners to consider challenges faced by local operators (e.g., restricted access, potential dangers) when they are developing their response strategy. As the mission evolves, it could serve as a tool for comparative analysis and for assessing the effectiveness of the current strategy.

Research objectives

The goal of the present research is to examine the impact of real-time image-based feedback on the various cognitive activities and strategies involved in diagnosis and situation assessments in an uncertain and highly dynamic environment. One concern in particular is that the detailed and real-time nature of this new data source will make it very compelling and thus adversely affects the use, interpretation, and integration of more traditional data. This concern is based on a growing body of anecdotal evidence from the use of current systems in the military that suggests that decision-makers over-rely on imagery feeds as a source of information. In other words, do they experience a narrowing and fixation of attention on the real-time imaging feedback?

METHODOLOGY

A simulation exercise was developed based on a 1994 HAZMAT accident at a chemical processing facility in England The use of an actual event as the basis for the simulation helped to ensure its face validity, an important factor in eliciting authentic performance from experienced practitioners (Woods and Sarter, 1993; Johansson et al, 2002). This particular event was chosen because of its complexity and the availability of post-event analysis and reference materials (Davis, 2002; HSE, 1996). It was also expected that the participants would not be familiar with the details of this event and thus not possess a performance advantage.

Scenario

The ICs are assuming command of a HAZMAT event at a large facility that produces motor fuel anti-knock compounds. There were approximately 300 employees on-site and a residential development 250 meters from the southwest boundary of the facility. A large gas cloud, approximately 10-15 meters high was spotted enveloping an ethyl chloride (EC) processing plant. This cloud was clearly visible in the UAV imaging.

The second hazard was a large pool of liquid EC collecting near the container tank. As it evaporates, liquid EC forms a flammable vapor and mist of hydrochloric acid (HCL) vapor. Contact with the water being used to disperse the gas would increase the rate of evaporation, thereby increasing the amount of vapor and increasing the flammability hazard.

Because the gas cloud obscured the collecting liquid EC, the imaging would not help in detecting this additional, and potentially more dangerous problem. Therefore, if the diagnosis of the problem were framed by what the imaging depicted, the conclusion would be that the only problem was the release of EC gas. However, several other sources of information, including update reports and the MSDS, provided evidence of this second hazard. A breach of the facility boundary by the gas was programmed to occur at the 70-minute mark and, if not addressed, the flammable vapors of EC gas would ignite and flash back to the EC tank at the 85-minute mark.

Performance standard

Although there was no prescribed sequence or set of activities for managing this scenario, a canonical performance standard was developed based on what could be expected from a proficient IC who has been provided with the tools and information provided in the exercise. All of the information necessary to detect and correctly diagnose the liquid EC hazard, as well as the threat that it presented, was made available to the ICs. However, because this information was distributed across different data sources and arrived at different times, it required the IC to sample multiple data sources over time.

Ethyl chloride has a low boiling point (12.2º C) and may exhibit the characteristics of either a liquefied flammable gas or a highly flammable liquid, depending on the ambient temperature. The tactics required to address gas and liquid EC are dispersion and suppression respectively. This is complicated by the fact that during the courses of the simulation, the ambient temperature rose from 9º C to the boiling temperature of 12.2º C. Therefore, the tactics used by the IC during the exercise, as well as the content of their situation assessments, provided converging evidence concerning their diagnosis of the situation and the data sources they used to develop that diagnosis.

EXPERIMENTAL DESIGN

Participants

Eight volunteers were recruited from the fire service and emergency management agencies of three different counties in Ohio, USA. All were experienced ICs, with between 2 and 16 years (avg. = 6.1) of experience in incident command.

Apparatus

The simulation exercise was conducted at the Butler County EMA Emergency Operations Center, which provided the communications tools, equipment, and information sources, such as maps, status boards, and material safety data sheets (MSDS), that are used during an emergency response. In addition, a 21” monitor was placed on the left side of the workstation to present the imaging feedback from the UAV (see Figure 1). The UAV imaging was simulated by digitally enhancing a high-resolution photo of the facility using Fire Studio 3.0 (Digital Combustion) to produce a video-like presentation (see Figure 2). Due to limitations in the simulation software, the imaging data was limited to a fixed view. The participants were informed of this limitation only if they asked to modify the vantage point.

Figure 1: Workstation setup

Figure 2: Imaging provided to the participant

Procedure

Prior to beginning the exercise, the participants were given time to ensure they were familiar with the equipment and resources available for conducting the operation. They were informed that they could request any additional information they needed, including details on the facility, the chemical processes, and surrounding areas. A total of 16 data cards (9 facility and 7 chemical hazard) were prepared for the most likely requests, however queries not included in this set were also answered. They were also informed that weather updates would be available every 10 minutes. The weather reports included ambient air temperature, wind speed and direction, cloud and precipitation conditions.

At the start of the simulation, a confederate gave the IC a 5-minute verbal report of events up to that point (a hard copy of this timeline was also provided). The confederate was also used to mediate communications and requests between the IC and the simulated on-site responders, facility employees, and outside agencies.

The simulation itself was approximately 90 minutes in duration. It should be noted that the simulation ran in real-time, with no “time-outs” or artificial advances in time. The exercise was followed by a 30-minute debriefing that was used to capture both the subjects’ attitude towards the new imaging data as well as to provide an opportunity to clarify any observations from the simulation.

Tasks

Diagnostic task

The main task was to identify all of the hazards at the facility and to develop a response to contain and minimize the threat to the surrounding population.

Verbal briefings

The participants were required to give three verbal situation assessments in the form of staff briefings at 30, 60, and 90 minutes into the exercise. These assessment reports included a detailed review of the current state and nature of the hazard, the planned response including resource management recommendations, and how they expected the situation to evolve until their next briefing in 30 minutes.

Management of operations

The main goals of diagnosing and assessing the situation were pursued while also responding, in real-time, to externally-driven events, such as new developments in the situation and fulfilling requests from on-site responders for information, decision-input, or for additional resources. These events were communicated via “injects” which were read to the IC and then presented as a text message by a confederate. There were 22 scheduled update reports and 22 requests from the field requiring a response. As much as possible, the scenario was adapted to account for the participant’s decisions by discarding a pre-planned message if it became irrelevant as a result of an earlier decision. The ICs had the option of verbalizing their responses, recording them on the message form, or both.

Data collection

In addition to manually recording which data cards the participants requested, audio and video recordings of the sessions were made. The process traces (Woods, 1995) developed from these recordings were used to analyze the participant’s model of the situation in situ. In addition to evaluating the accuracy of their assessments, these traces were also used to analyze the comprehensiveness of the participants’ use of available/relevant data and their skepticism about the fidelity of the data and its source, i.e., if it originated from fire service personnel or civilians.

RESULTS

The main task was to correctly identify the nature of a hazard and to apply the appropriate response as quickly as possible. All of the ICs initially mounted a dispersion attack to limit the spread of the EC gas beyond the facility boundaries. Only IC#4 detected and correctly diagnosed the flammability and explosive threat from the liquid ethyl chloride and ordered a foam suppression attack in concert with the water curtain dispersion. One other IC (IC#1) inquired about the availability of suppressant foam; however, during the debriefing, he indicated that this was a precautionary measure only and not tied to a specific threat. The remaining seven ICs maintained the initial assumption that the only hazard present was a toxic vapor. They failed to revise their situation assessment despite receiving information that indicated the possible presence of liquid EC collecting and the limited success of their initial tactic. As a result, the flammable EC vapors ignited and flashed back to the reactor area.

Patterns in data search strategies

As shown in Tables 1 and 2, there were no significant differences between IC#4 and the other ICs in terms of which data cards were requested or in the overall number of requests made regarding the facility and the chemical processes present. The only requests that were unique to IC#4 were for two hazard data cards: H5 and H7. These cards provided information about the boiling point and the potential hazards of liquid ethyl chloride, which suggests that he was exploring potential complications outside of the toxic gas/vapor scenario.

Facility data	Subject #							
	4	1	5	2	7	3	6	8
F1: Chemicals on site	♦	♦	♦	♦	♦	♦	♦	♦
F2: Description of surrounding area/population	♦	♦	♦	♦	♦	♦	♦	♦
F3: Fire protection systems (cladding, suppression)	♦	♦	♦	♦	♦	♦	♦	♦
F4: Types of chemical processes present	♦	♦	♦	♦	♦	♦	♦	♦
F5: Chlorine storage area description & location	♦	♦	♦					
F6: Number of employees						♦	♦	♦
F7: Hydrology of area								

Table 1: Facility data card requests

Hazard data	Subject #							
	4	1	2	5	3	8	6	7
H1: EC reactor contents	♦	♦	♦	♦	♦	♦	♦	♦
H2: Vapor hazards	♦	♦	♦	♦	♦	♦	♦	♦
H3: EC exposure standard		♦	♦	♦			♦	
H4: EC handling	♦	♦	♦	♦				♦
H5: Boiling point of EC	♦							
H6: HCL hazard		♦						
H7: Liquid EC hazards	♦							
H8: BLVE conditions								
H9: EC environmental effects								

Table 2: Chemical hazard data card requests

The process trace reveals that one key difference in IC#4's data search strategy was not as much what data was requested, but in verifying data as well as the underlying assumptions associated with it, particularly if it did not originate from fire service personnel. Although there was no error in the briefing that may have led to an alternate hypothesis, it indicates a certain level of skepticism regarding data that the other ICs took at face value, which possibly left him open to revising his assessment. For example, he requested confirmation of the information contained in the initial briefing (as provided by the facility employees), including the source of the leak and the nature and contents of the processor plant. He also confirmed the hazard and response strategies provided by the facility personnel with the MSDS in which he noted the low boiling point of liquid EC and that water can accelerate its evaporation, thus releasing more vapor.

While all of the ICs referenced the imaging data, IC#4's sampling behavior differed in several ways. While he initially used it like the other ICs, to orient himself and to observe the extent of the vapor, over the course of the scenario, his sampling frequency was lower and the duration of each sample was longer. The timing of his sampling also differed from the majority of the ICs. They tended to occur during lower tempo periods, i.e., times in which no updates or external demands had recently been received.

DISCUSSION

Although the presence of multiple hazards and the inherent time pressure alone were sufficient to create a technically challenging scenario, assessing the situation was particularly difficult in this case because it presented a "garden path" problem (Johnson et al., 1988). In such scenarios early evidence points to an incorrect, but plausible assessment, while later, less salient evidence, points to the correct assessment. Garden path problems often lead to a "this and nothing else" fixation in which the operator is unable to realize the need to revise his assessment and deviate from the original plan (De Keyser and Woods, 1990). In this mindset, the operator exhibits a confirmation bias, in which only data that supports the currently held belief is applied while disconfirming information is either rationalized away or discounted completely (Nisbett and Ross, 1980; Tversky and Kahneman, 1974).

The imaging supported the early assessment not only by highlighting the vapor hazard, but also by masking other cues, possibly resulting in a form of the "surface/deep" oversimplification in which the focus is on the interpretation afforded by the surface cues, when in reality, the important cues require a deeper search to be uncovered (Feltovich et al, 1997). However, while this fixation may explain why the ICs failed to use the full array of information sources that were available, it does not fully explain the apparent dominance of the real-time imaging in framing the ICs' situation assessments. To this end, we will examine the properties of this type of feedback in more detail and how they may influence the situation assessment process.

Compared to verbal messages, which are transient in nature, and textual data, which can be obscured or misplaced at a workstation, the real-time imaging was continuously available to the IC and easily observable. Thus, it could be argued that it was highly salient relative to the other information sources. This is relevant because the attention narrowing effect often associated with fixations can result in a shift from expectation-driven information search to event-driven attention capture (Folk, et al, 1992; Moray, 1986).

The perceived fidelity of the imaging data may also have contributed to its apparent dominance. Due to past field experience, many practitioners in emergency management and the fire services work under the principle of "seeing is believing" (Page, 2005). In other words, they place a high level of trust in visual evidence and particularly in their ability to evaluate a situation based direct observation. Therefore, the literal nature of the imaging was more familiar, and therefore garnered greater trust, than the pre-processed, filtered, and abstracted data sources normally provided. Note, however, that, while the imaging data was accurate, it was not complete and did not provide all of the relevant cues needed for an accurate diagnosis.

Another possible reason for the perceived value of the imaging data was its timeliness. Several of the ICs indicated that, given the high tempo of events in which circumstances can change quickly, the task relevance of information is highly correlated with its currency, or as one IC summarized, "newest equals best". Therefore, the reasonable assumption was that the real-time feedback was accurately reflecting the current state of the situation.

These initial results indicate that this form of real-time feedback can have a detrimental, narrowing effect in terms of what data sources are incorporated into a situation assessment with a corresponding reduction in exploration of the solution space. This narrowing effect was due, in part, to a perception that feedback that is unfiltered, unprocessed, and un-delayed is complete and high fidelity information. In turn, this "completeness fallacy" resulted in an inappropriate level of trust in and reliance on this feedback.

CONCLUSIONS

Real-time, image-based feedback introduces two fundamental changes in the nature of the information that decision-makers can access. It not only eliminates (or greatly reduces) the time delays they experience; it also introduces a shift from having to integrate preprocessed verbal/textual descriptions to perceiving a richer, un-processed visual representation. However, there are also several limitations and potential unintended consequences of this new feedback. Although all levels of command can now be supplied with the same image-based feedback in real-time, its value to practitioners at those different levels, and the methods in which it should be utilized, are different. While direct observation of an area of interest can reveal some spatial relationships, it still must be abstracted into a form that can be used for planning purposes.

These results suggest that the observed breakdowns could be addressed by applying a cognitive systems approach to the design and implementation of this new type of feedback. For example, design elements that highlight important changes in the visual field and track the evolution of various processes would enable operators to detect relevant details while still maintaining a global perspective. Integrating an interactive critiquing system into the design would encourage the exploration of alternative hypotheses and enhance anomaly detection.

This research was intended to help Emergency Management Agencies and Incident commanders identify and better understand the consequences of introducing this new form of feedback into the command and control loop. potential impact of introducing similar technologies on training and operational procedures. One extension of this work that is currently being developed is to explore how distributed teams might employ this technology to support collaborative work. At a more general level, it is hoped that it will contribute to the knowledge base of technology-mediated decision-making in event-driven environments.

ACKNOWLEDGMENTS

This research was prepared, in part, through collaborative participation in the Advanced Decision Architectures Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0009, and through support from the Eddowes Memorial Endowment, The Ohio State University. We wish to thank all of the study participants for their time, efforts and valuable feedback. Special thanks also go the Butler County Emergency Management Agency for its support of this project.

REFERENCES

1. Billings, C.E. (1997). *Aviation automation: The search for a human-centered approach*. Mahwah, NJ. Lawrence Erlbaum Associates.
2. Brehmer, B. and Allard, R. (1991). Dynamic decision-making: The effects of task complexity and feedback delays. In J.Rassmussen, B. Brehmer, and J. Leplat (Eds.) *Distributed decision making: Cognitive models for cooperative work,* (pp.319-334). Wiley.
3. Clark, H.H. and Brennan, S. (1991). Grounding in communication. In L. Resnick, J. Levine, and S. Teasley (Eds.), *Socially shared cognition.* Washington, DC: American Psychological Association.
4. Cook, R.I. and Woods, D.D. (1996). Adapting to new technology in the operating room. *Human Factors, 38,* 593-613.
5. Danielsson, M., and Ohisson, K. (1999). Decision making in emergency management: A survey study. *International Journal of Cognitive Ergonomics, 3,* 91-99.
6. Davis, D. (2002). Fire commander. In R. Flin and K. Arbuthnot (Eds*.) Incident Command: Tales from the hot seat,* (pp. 88-104) Burlington, VT, Ashgate Publishing Ltd.
7. De Keyser, V. and Woods, D.D. (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In A.G. Colombo and A. Saiz de Bustamante, (Eds.), *Systems reliability assessment*. Dordrechts, The Netherlands: Kluwer Academic.
8. Digital Combustion, Inc. (2004). Fire Studio 3.0 (computer software).
9. Feltovich, P.J., Spiro, R.J., and Coulson, R.L. (1997). Issues of expert flexibility in contexts characterized by complexity and change. In P.J. Feltovich, K.M. Ford, and R.R. Hoffman (Eds.), *Expertise in context: Human and machine*. Menlo Park: CA. AAAI/MIT Press.

10. Folk, C.L., Remington, R.W., Wright, J.H. (1992). The Structure of attentional control: Contingent attentional capture by apparent motion, abrupt Onset, and color. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 317-329. Washington DC: American Psychology Association.
11. HSE (1996). The chemical release and fire at the Associated Octel Company Ltd., Sudbury: HSE Books.
12. Johannson, B., Hollnagel, E. and Granlund, A. (2002). The control of unpredictable systems. In Johnsson C. (Ed.) *Proceedings of the 21st European annual conference on Human Decision Making and Control*, GIST Technical Report G2002-1, Department of Computing Science, University of Glasgow, Scotland.
13. Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (pp. Ch.40, 1-51). New York: Wiley.
14. Nisbett and Ross (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewoods Cliffs, NJ: Prentice-Hall.
15. Norman, D.A. (1990). The ‘problem’ of automation: Inappropriate feedback and interaction, not ‘over-automation’. *Philosophical Transactions of the Royal Society of London,* B327.
16. Page, Douglas (2005). Seeing is believing, *Fire Chief*, September, 2005. On-line at http://firechief.com/tactics/command_camera_092305/
17. Patterson, E. S., Cook, R.I., and Render, M.L. (2002) Improving patient safety by identifying side effects from introducing bar coding in medication administration. *Journal of American medical information association*, 540-553.
18. Sarter, N.B. and Woods, D.D. (1992). Pilot interaction with cockpit automation I: Operational experiences with the flight management system. *International Journal of Aviation Psychology, 4*, 1-28.
19. Sarter, N.B. and Woods, D.D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37*, 5-19.
20. Smith, P.J., McCoy, C.E., & Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: The effects on user performance. *IEEE Transactions on Systems, Man, and Cybernetics, 27*, 360-371.
21. Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1130.
22. Woods, D.D. (1993) Process-tracing methods for the study of cognition outside the experimental psychology laboratory. In: G.A. Klein, J. Orasanu, R. Calderwood, C.E. Zsambok (Eds*.) Decision making in action: models and methods,*(pp.228-251*).* Norwood, NJ*:* Ablex Publishing Corporation.
23. Woods, D.D. (2002). Steering the reverberations of technology change: On laws that govern cognitive work. *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*.
24. Woods, D.D., Johannesen, L., Cook, R.I., and Sarter, N. (1994) *Behind human error: Cognitive systems, computers, and hindsight*. Dayton, Ohio: Crew Systems Ergonomic Information and Analysis Center, WPAFB.
25. Woods, D.D. and Sarter, N.B. (1993). Evaluating the impact of new technology on human-machine cooperation., In J. Wise, V.D. Hopkin, and P. Stager, (Eds.), *Verification and validation of complex systems: Human factors issues,* (pp. 133-158). Berlin: Springer-Verlag.