# An investigation of training strategies to improve alarm reactions

James P. Bliss*, Eric T. Chancey

*Psychology Department, Old Dominion University, 5115 Hampton Boulevard, Norfolk, VA, USA*

## ARTICLE INFO

## ABSTRACT

Researchers have suggested that operator training may improve operator reactions; however, researchers have not documented this for alarm reactions. The goal of this research was to train participants to react to alarms using sensor activity patterns. In Experiment 1, 80 undergraduates monitored a simulated security screen while completing a primary word search task. They received spatial, temporal, single sensor, or no training to respond to alarms of differing reliability levels. Analyses revealed more appropriate and quicker reactions when participants were trained and when the alarms were reliable. In Experiment 2, 56 participants practiced time estimation by simple repetition, performance feedback, or performance feedback and temporal subdivision. They then reacted to alarms based on elapsed time between sensor activity and alarm onset. Surprisingly, results indicated that participants did not benefit differentially from temporal interval training, focusing instead on advertised system reliability. Researchers should replicate these findings with realistic tasks and real-world complex task operators.

© 2013 Elsevier Ltd and The Ergonomics Society. All rights reserved.

## 1. Introduction

As technology has become more complex and capable, task operators have increasingly relied on sensor-driven signals to inform them about dangerous conditions and event. Researchers investigating the link between the reliability of automated signaling systems and operator reactions have drawn several conclusions. Chief among these is the idea that signaling systems producing false alarms inspire lack of trust (termed the "Cry-Wolf Effect") and that subsequent reactions to such systems may become slower or less frequent (Breznitz, 1984). Initial work by researchers such as Janis (1962) and Breznitz (1984) documented well the degradation of reaction behaviors. Subsequently, other researchers have focused on the characteristics of personal reactions to unreliable alarms, the variability of alarm reaction behavior associated with competing tasks and personal motivators, and the influence of information availability and performance consequences on behaviors. Such investigations have demonstrated that alarm reaction behavior is a complex phenomenon that is inextricably tied to the target task, signaling equipment, and operator cognition.

To broaden the understanding of human reactions to unreliable signals, researchers have relied on existing theories of learning and human cognition. These have included probability matching, where observed alarm responses approximate the perceived true alarm rate (Bliss et al., 1995), Signal Detection Theory, where signal detection and response is more rapid for historically reliable alarm systems (Getty et al., 1995), and automation trust, where exhibited trust corresponds to behavioral patterns of system use and misuse (Parasuraman and Riley, 1997). A clear conclusion is that, regardless of specific task domain, alarm reaction patters are undeniably influenced by perceived alarm system reliability. Specific investigations have shown application to domains as diverse as aviation (Pritchett, 2001), surgical theaters (Xiao and Seagull, 1999) and security monitoring (Marra and Playford, 2009). Because of the implications of low alarm reliability, designers and researchers have endeavored to remedy the problem. Designers of signaling systems have focused their efforts on sensor and display technologies such as likelihood alarm displays (Sorkin et al., 1988). This approach to signal design may be used to embed information about anticipated alarm validity within the signal itself. Bustamante (2008) showed the power of this approach. However, the benefits of likelihood alarm displays are evident only after signals have occurred; furthermore, individual operators frequently exhibit variability in responding, even when signal urgency and reliability are unequivocally high (Bliss, 2003). In addition, designed likelihood alarm displays may not be flexible enough to retain effectiveness across operational environments and task situations. The current work is complementary, targeting behavioral change strategies to insulate task operators from signal unreliability.

## 2. Experiment 1

One of the most common strategies advocated for task performers to manage signal unreliability is responding to every signal, regardless of its perceived validity. Such a strategy is common in high-consequence task environments such as medicine, where failure to acknowledge a valid signal may outweigh the impact of responding to false alarms (Xiao et al., 2004). One issue, however, is that task operators are forced to allocate their attention to stimuli that may not be consequential. This imparts inefficiency in response behavior and may actually lengthen response times in cases where multiple signals occur simultaneously. A second problem is that over time, operators may react more slowly to all signals, assuming that a certain proportion of them will be false (Getty et al., 1995). Implications such as these illustrate the need for a more discriminative training solution, especially in task environments where false alarms are frequent such as medical care or security monitoring (Xiao et al., 2004; Marra and Playford, 2009).

An alternative approach to training is to train operators to recognize and anticipate trends in the underlying data that drive signal annunciation. Such an approach capitalizes on the tendency of operators to rely on redundant sources of information to judge signals (Bliss, 2003). Our first experiment was intended to provide empirical support for the idea that data pattern training could improve operator reactions.

Our approach involved presenting participants with a simulated security monitoring task where periodic signals warned of motion in certain building rooms. Participants were advised about the signaling system's reliability before participating and received training to recognize spatial or temporal sensor signal patterns or focus on one sensor. We hypothesized that participants would respond most to high-reliability alarms, regardless of training (Bliss et al., 1995). We also anticipated that participants receiving single-sensor training would respond most often and most appropriately because of that method's simplicity. Researchers have not directly compared spatial and temporal training effectiveness; therefore, we made no hypotheses about the relative benefit of these training conditions.

### 2.1. Method: experiment 1

#### 2.1.1. Design

For Experiment 1 we employed a 2 × 4 split-plot design. The within subjects variable was the stated reliability of the alarm system and had two levels: 20% and 40% true alarms. These levels were chosen to reflect the low reliability rates associated with security alarm systems (Sampson, 2002). The between subjects variable was alarm reaction training type and had four levels: single sensor training, spatial pattern sensor training, temporal pattern sensor training, and no training (control).

Performance based dependent measures for the signaling task included reaction time to the sensors (time taken to click on the acknowledge button after a sensor activated), "respond" and "ignore" rates for the alarm signals (proportion alarms to which participants clicked on the "respond" or "ignore" buttons) and reaction time for alarm reaction behaviors ("respond" or "ignore"). A running alarm score reflected the number of times participants correctly responded to and ignored alarms. If a correct choice was made (responding to a true alarm or ignoring a false alarm), 1.5 points were added to the score; the same number was deducted from the score after each incorrect choice (ignoring a true alarm or responding to a false alarm). The score was presented visually to the participants during the task. Correct choices were accompanied auditorily with a female voice saying "Correct." Incorrect choices were accompanied by "Incorrect." These components facilitated participant decision-making during the task.

Subjective trust ratings for the alarm systems were also recorded, using Jian, Bisantz, and Drury's subjective trust questionnaire. Word search primary task performance was also recorded as the number of words identified during each experimental session. Word search was required as a loading task to more realistically simulate dual-task signal monitoring conditions. No statistically significant differences were noted for this variable, indicating similar primary task loading across groups and conditions.

#### 2.1.2. Participants

Eighty undergraduate students enrolled in psychology courses at Old Dominion University in Norfolk, Virginia, were tested. The mean age of the sample was 23.3 (SD = 2.25). The sample consisted of 17 males and 63 females. No participant indicated hearing loss or color deficiency. Participants were awarded course credit for participating in the 2-h experiment. A $20 performance incentive was also offered to the participant with the highest combined primary and secondary task score.

#### 2.1.3. Materials

The alarm task was modeled after a building security-monitoring scenario (see Fig. 1). Participants viewed a building schematic that contained motion sensors in selected rooms. Participants acknowledged every sensor signal (red light in a room and a 1000-Hz. tone) by clicking on the "ACK" icon. An alarm signal followed five prior sensor signal activations. The alarm signal was a red area containing the word "INTRUDER!" at the bottom of the screen and a fire bell from a Boeing 757. Participants could react to an alarm by selecting between two icons, "RESPOND" (true alarm) or "IGNORE" (false alarm). As noted above, correct reactions depended on prior sensor activations.

Referring to Fig. 1 below, single sensor training involved focusing on a single sensor (highlighted in red) to determine the validity of the subsequent alarm. Participants in this group were told that if the sensor activated at any time before the subsequent alarm signal, the alarm would be true. Spatial pattern training required detection of a clockwise sensor activation pattern to judge subsequent alarm validity. Participants were told to note the spatial pattern of sensors that activated prior to an alarm signal. If the pattern was clockwise, this indicated that the alarm signal would be true. Temporal pattern training required detection of a short sensor–alarm time interval to judge alarm validity. Participants in
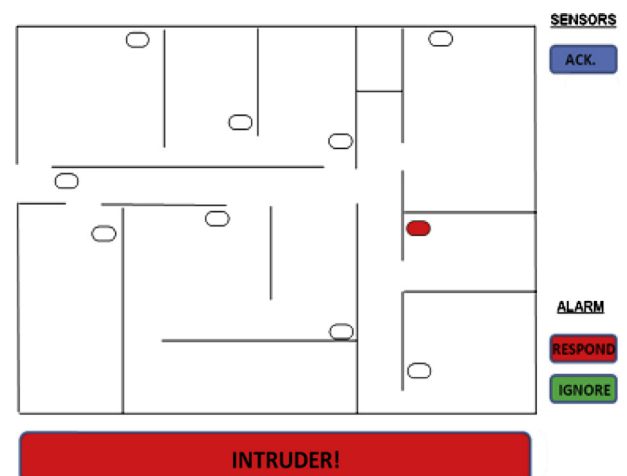


**Fig. 1.** The Sensor and Alarm Reaction Task.