Cough Once for Danger: Icons Versus Abstract Warnings as Informative Alerts in Civil Aviation

Nathan C. Perry, Catherine J. Stevens, Mark W. Wiggins, and **Clare E. Howell,** University of Western Sydney-Bankstown, South Penrith, Australia

Objective: An experiment investigated the efficacy of auditory icons as warning signals in an aviation context. **Background:** Iconic signals, such as a cough to signal dangerous levels of carbon monoxide, convey information about the nature of an incident and alert the operator that there is a problem, whereas signals that are arbitrarily associated with a critical incident provide relatively less information. Warning recognition speed and accuracy are likely to be influenced by modality of presentation (visual, auditory, auditory + visual) and task demand (low, high). **Methods:** The 172 participants completed a computer-based training session and test task that involved responding to abstract or iconic auditory (1 s), visual, or auditory + visual warnings associated with seven critical incidents while performing low- and high-demand concurrent tasks. **Results:** Significantly fewer training trials were required to learn iconic warnings than abstract warnings. An advantage for iconic warnings persisted into the test phase, evident most consistently as greater warning recognition accuracy. The effect was observed in both high- and low-demand conditions. Auditory abstract warnings, in particular, elicited slow reaction times and poor accuracy. **Conclusion:** Associations between a small number of meaningful environmental sounds and critical incidents can be learned with ease relative to more abstract associations, although training is required and response times are relatively slow. **Application:** Sets of distinctive auditory iconic warnings can be designed to alert and inform pilots about non-time-pressured events. Potential applications of language-neutral icons as informative warnings include civil, commercial, and defense aircraft.

INTRODUCTION

Research has highlighted significant constraints in the design and use of abstract sounds as warnings. Alarms that blare or beep involve an arbitrary mapping between a signal and the event to which it refers. A litany of the weaknesses of abstract alarms notes that they are loud and repetitive and may mask other communication; they annoy rather than inform; they do not convey urgency; they may elicit a startle response that interferes with the necessary reaction; they go unrecognized 40% of the time when there are more than seven or so different alarms; and they require training and retraining (Begault, 1994; Momtahan, Hetu, & Tansley, 1993; Patterson, 1982).

Patterson (1982) recommended ways to select the frequency components of a warning sound's spectral profile to lessen the likelihood of masking by noise and other warnings and recommended the use of four to nine warning signals. A system with four signals and two *attensons* (initial, nonverbal signals) that signaled different levels of urgency was regarded as optimal, the purpose of the attenson being to inform the operator of the type of warning and that more information is available.

Auditory icons or caricatures of everyday sounds (Belz, Robinson, & Casali, 1999; Calhoun, Janson, & Valencia, 1988; Gaver, 1989, 1993; Graham, 1999; Ho & Spence, 2005) circumvent a number of the problems associated with abstract alarms as they involve a relation between a signal and referent – for example, the sound of coughing to signal dangerous levels of gas. Auditory icons can be short, are not easily masked, and are generally recognizable and distinguishable (Keller &

Address correspondence to Catherine J. Stevens, MARCS Auditory Laboratories, University of Western Sydney-Bankstown, Locked Bag 1797, South Penrith DC NSW 1797, Australia; kj.stevens@uws.edu.au. ***HUMAN FACTORS,*** Vol. 49, No. 6, December 2007, pp. 1061–1071. DOI 10.1518/001872007X249929.

Stevens, 2004). This ease of recognition of iconic signals has been put to use in interface design, where visual icons are used extensively based on the assumption that icons can transcend language barriers and present meaning in a condensed form (Caplin, 2001; Gittens, 1986; McDougall, de Bruijn, & Curry, 2000).

When designing sets of alarms for a particular context, there are open questions about the nature of the warning (abstract, iconic); relative effectiveness of auditory, visual, and auditory + visual modalities; and potential interactions among warning type, modality, and task demand (low and high cognitive load).

Recognition of Abstract and Iconic Warnings

Graham (1999) demonstrated that learning and response time were reduced using auditory icons relative to abstract warnings (tones). In Graham's (1999) experiment, two auditory icons (car horn, screeching tires) were tested together with tonal and speech warnings in a collision warning context. Without training, participants recorded significantly faster reaction times in response to auditory icons as compared with other signals. Belz et al. (1999) compared conventional auditory warnings and auditory icons to warn commercial motor vehicle operators of impending collisions. Auditory icons elicited significantly better driver performance over conventional auditory warnings. Gaver, Smith, and O'Shea (1991) reported good performance of novice operators using up to 14 auditory icons in the simulation of a bottling plant. The advantage of auditory icons over abstract warnings is the availability of a large number of distinguishable sounds.

Studies of visual icon characteristics – semantic distance, concreteness, familiarity – provide a theoretical basis for the present experiment. McDougall et al. (2000) have demonstrated that the importance of icon characteristics changes with experience. For example, semantic distance, operationalized in the present experiment as an iconic-abstract comparison, is crucial initially and while icon-function relations are being learned. Icon familiarity is important later and reflects access to information in long-term memory.

If warnings are iconic for the events they represent, then few training trials relative to abstract warnings will be needed for participants to reach a criterion level of recognition accuracy. Reaction time (RT) in response to iconic warnings might also be expected to be fast, relative to abstract warnings, during training trials. The cognitive processes involved in learning iconic warnings will include recognition of the source of the sound or image and long-term memory activation. By contrast, abstract warnings will need to be learned from scratch. Features of the sound or image will need to be extracted and associated in memory with the particular critical event. The association will be strengthened with repeated exposure and directive training.

An open question is whether greater accuracy and faster RT in response to iconic warnings, relative to abstract warnings, persist once participants reach a criterion level of accuracy. Even after a period of training to recognize abstract warnings, the acquired familiarity and long-term memory connections that exist for icon-function relations may lead to better accuracy and RT for iconic warnings. To approximate real-world environments, cognitive load should be manipulated.

Warning Recognition in High-Workload Situations

Many situations require people to perform tasks simultaneously. Competition for cognitive resources, such as attention, memory capacity, and communication channels, impacts performance on concurrent tasks (Norman & Bobrow, 1975). Generally, there is less interference when a visual and an auditory task are performed concurrently than when tasks are performed that involve the same modality (Wickens & Liu, 1988).

In a multiple resources model of attention (Wickens, 2002), separate resources exist for each of four resource categories: modalities (auditory, visual), codes (spatial, verbal), stages (perception, response), and visual channels (focal, ambient). According to this model, there should be greater interference between concurrent tasks competing for resources with the same categories, such as two visual tasks (operationalized here as a visual warning recognition task and a visual mathematical task), than there will be for tasks that use resources from different categories, such as concurrent audio and visual tasks.

Recognition of Auditory and Visual Warnings

Text-based warnings are common visual warnings; they convey information easily and with little ambiguity (Edworthy & Adams, 1996). However, physical degradation and difficulties in interpretation have led to the wider use of icons. Ideally, visual icons allow rapid and accurate communication of information without extensive training in the relation between signal and target event.

As auditory warnings are omnidirectional in nature (requiring less scanning of instruments by operators) and can reach farther distances than can visual warnings (Doll & Folds, 1986), they are often used to convey information in emergency situations. In a simple RT task, reactions are generally faster to an auditory than to a visual stimulus (Woodworth & Schlosberg, 1954). Auditory and multimodal displays have been found to be effective in reducing RT and ratings of subjective workload, as compared with visual-only displays (Liu, 2001). In heavily loaded visual displays, an auditory display can improve time-sharing performance (Wickens, Sandry, & Vidulich, 1983). Combining auditory and visual signals may enhance task performance (e.g., Belz et al. 1999).

Aim, Design, and Hypotheses

Based on the preceding discussion of alarm presentation and resource theory, there is merit in determining the effects of the following variables and combinations of them on learning and recognizing warning signals. The first factor, warning type, contrasted iconic and abstract warnings. The second, modality, consisted of three levels: auditory, visual, and auditory + visual (A+V). The third factor, task demand, consisted of two levels: low and high. The first two variables were between subjects. The dependent measures were learning rate, measured as the number of trials taken to reach a criterion level of accuracy during the training phase, and, in both the training and test phases, warning recognition accuracy and RT. It was hypothesized that

1. on training trials, participants would require fewer trials, make fewer errors (number incorrect), and record faster RTs in response to iconic warnings than to abstract warnings in order to achieve a criterion level of accuracy (Belz et al., 1999; Gaver et al., 1991; Graham, 1999; McDougall et al., 2000);
2. in high- and low-demand conditions of a dual-task test phase, accuracy (number correct) would be greater and RTs faster in response to iconic warnings as compared with abstract warnings (Belz et al., 1999; Gaver et al., 1991; Graham, 1999; McDougall et al., 2000); and
3. in the high-demand dual-task test phase, involving a visual mathematical task and warning recognition task, there would be less interference such that accuracy would be greater and RTs faster in response to warnings presented in a nonvisual modality (i.e., auditory warnings), as compared with warnings presented visually (Norman & Bobrow, 1975; Wickens, 2002).

METHOD

Participants

The participants were 215 students at the University of Western Sydney and received course credit for taking part in the experiment. Of these, 34 were excluded from analysis as they failed to reach criterion (they did not learn all signals within 12 training-test blocks). An additional 9 participants were excluded as outliers (i.e., z score ±3.29; Tabachnick & Fidell, 2001). The final sample consisted of 172 participants (135 women and 37 men, mean age = 20.55 years, SD = 4.38 years, range 17–39 years). There were 82 participants in the abstract condition (auditory = 24, visual = 29, A+V = 29) and 90 in the iconic condition (auditory = 29, visual = 28, A+V= 33).

Stimuli

A stimulus selection rating task was conducted prior to the experiment to determine the most representative warning for each event from the four warning types: iconic auditory, iconic visual, abstract auditory, and abstract visual (between subjects). In the rating task, 40 participants (12 men, 28 women, mean age = 22.46 years, range = 18–32 years) rated the degree of association between each event and warning on a 5-point scale ranging from *totally unrelated* (1) to *highly related* (5).

Participants also listed up to five ways in which the warning was related to the event. For example, a possible relationship between the visual image of an elephant and the critical event of an aircraft being overweight could be that both an elephant and an overweight plane are heavy. Seven events were used as stimuli in the experiment. The seven events and their mean ratings are shown in Table 1. Participants rated all abstract auditory and abstract visual warnings and events as unrelated. The mean rating for abstract warnings was 1.56, SD = 0.72.

Iconic auditory warnings. Auditory icons were obtained from the Web sites www.sounddogs.com and www.findsounds.com. Indirect (e.g., metaphorical, ecological) relationships between the

TABLE 1: Highest-Rated Visual and Auditory Iconic Warnings for Events in the Stimulus Selection Phase

Event	Visual Icon and Mean Rating		Auditory Icon and Mean Rating	
Low fuel		3.77	Car failing to start	4.17
Conflicting traffic		4.78	Car brakes screeching	4.33
Carbon monoxide		3.77	Coughing	4.75
Ground proximity		4.33	Explosion	3.67
Electrical failure		3.67	Zapping sound	4.50
Aircraft icing		3.66	Cold wind blowing	3.50
Aircraft overweight		3.89	Elephant trumpeting	3.67

Note. Maximum rating = 5, *highly related.* Icons reproduced with permission from Clipart.com.

warning and the event were chosen because not all events have a direct relationship that can easily be portrayed (Keller & Stevens, 2004). For example, there is no sound that signifies an aircraft being overweight. Although not directly related to the event, auditory icons were chosen on the basis that they still shared a strong indirect relationship. For example, the sound of an explosion is indirectly related to the critical event of ground proximity. All sampled everyday sounds were 1 s in duration, 16-bit mono, and standardized to a sample rate of 44.1 kHz, with normalized amplitude.

Abstract auditory warnings. The auditory abstract warnings consisted of 1-s mono tones with normalized amplitude, 44.1-kHz sample rate, and 16-bit resolution. The abstract sounds were designed based on guidelines set by Patterson (1982), in which a burst of sound is created and played repeatedly over the duration of the signal. Each sound burst had its own set of frequencies and pause durations, giving each warning unique pitch and temporal patterns. The duration of bursts varied from 0.19 s to 1 s. The tones and upper harmonics were selected from frequencies in the range 150 to 3000 Hz. The features of each of the abstract sounds, and the events to which they were assigned, are shown in Table 2. The abstract sounds stood in no obvious relation to the events with which they were paired.

Iconic visual warnings. The visual icons were obtained from the Web site www.clipart.com and are depicted in Table 1. The visual icons were selected on the basis of their relationship to each of the critical aviation events, but not directly so. For example, for the low-fuel event, a petrol pump was selected as one of the visual icons. All Clipart images were black and white.

Abstract visual warnings. To ensure that image complexity was comparable across the set of visual icons and abstract images, the abstract visual warnings were obtained by enlarging a small section of iconic images from Clipart, other than those being used as iconic visual warnings. In abstract warnings, there was no obvious relationship between the signal and the event with which it was paired. All images were black and white. Abstract visual warnings for each of the seven events are shown in Table 2.

TABLE 2: Abstract Visual and Abstract Auditory Warnings Used as Stimuli

		Abstract Auditory Warning Features		
Event	Abstract Visual Warning	No. of Bursts	Burst Duration (s)	Frequencies (Hz)
Low fuel		3	0.19	1000
Conflicting traffic		3	0.35	500, 2000
Carbon monoxide		2	0.50	689, 1033, 2067, 1722
Ground proximity		1	1.00	600, 850
Electrical failure		2	0.50	1200, 950
Aircraft icing		2.5	0.40	1300, 860, 1600, 730
Aircraft overweight		6	0.083	3000

Note. Icons adapted with permission from Clipart.com.

Bimodal warnings. The multimodal iconic and abstract warnings were A+V combinations of the seven auditory and visual warnings.

In the training phase of the experiment, the seven critical events were presented as "clickable" buttons on a computer screen, equidistant from one another. Warnings were presented either visually, within a square at the top of the computer screen for 1,000 ms, or auditorily, again lasting for 1,000 ms.

The test phase of the experiment included a mathematical addition task, presented visually and concurrently, designed to form the low- and high-demand conditions of the dual task. The low-demand version consisted of three single digits, all less than 5 (e.g., 4, 1, 3), and the high-demand task consisted of four single digits, greater than the number 5 (e.g., 7, 9, 6, 8). Participants were required to mentally sum the numbers and say the answer aloud. Each trial of digits was displayed in a 6.5- × 4.5-cm box centered in the middle of the screen for a period of 500 ms. Both low- and high-demand dual tasks consisted of 36 additions. The ratio of addition items to warning signal presentation was 4:1 to simulate an operator performing another task when an alarm occurs.

The order of low- and high-demand conditions was counterbalanced across participants to distribute serial order effects. Each warning was presented once, in random order, in high- and low-demand conditions of the test phase. Visual warnings were presented in the upper half of the screen within a 6.5- × 4.5-cm box.

Equipment

Abstract warnings were constructed using CoolEdit Pro 2.1. The stimulus selection rating task was presented to small groups using MS PowerPoint on a laptop computer (AMD Athlon XP2000), a data projector, and speakers. The training and test phases of the subsequent experiment were programmed in PowerLaboratory Version 1.0.3 (Chute & Westall, 1996) and presented to participants individually on a Power Macintosh 7300/200, with a 15-inch (38-cm) monitor and built-in speakers. Speaker volume was set to a constant, comfortable listening level with sound pressure level of 65 to 70 dB. The computer recorded accuracy and RT from the mouse click made to identify the critical event. The experimenter recorded the mathematical addition responses uttered by participants.

Procedure

At the beginning of the training phase, participants were provided a context for each of the warnings by reading an information sheet about critical aviation events – for example, "Carbon monoxide is a colorless, odorless gas that can be produced through the burning of fossil fuels. If sufficient levels of carbon monoxide enter the cockpit, the pilot can be rendered unconscious."

Participants were trained on the warning-event pairs through successive presentations of each warning and directions to click the on-screen button labeled with the paired event. They were tested as to how well they had learned the association between warning and event and were alternately trained (via presentation of the warning and direction to the labeled button) and tested until they reached 100% accuracy on 21 test trials.

Once this criterion level of accuracy was attained, participants proceeded to the test phase, in which they performed visual addition tasks. Numbers flashed on the screen and participants were required to add the numbers together as quickly and accurately as possible, while still identifying the warnings via a mouse click. The experiment took 30 to 40 min.

RESULTS

Data were analyzed using planned comparisons, guided by the hypotheses, for each dependent variable. An ANOVA is not required where planned comparisons are conducted, and neither ANOVA nor MANOVA would allow the scrutiny of specific planned comparisons involving interactions that were required to evaluate Hypotheses 2 and 3. Where multiple comparisons were conducted, the significance level was adjusted accordingly.

Hypothesis 1: Iconic Versus Abstract Warnings During Training

It was hypothesized that to recognize iconic warnings, as compared with abstract warnings, participants would take significantly fewer training trials to reach criterion, record a lower error rate, and record faster RTs. Alpha rate was set at .05.

Trials to criterion. Hypothesis 1 was supported, $F(1, 170) = 115.38$, $p < .05$, $\eta^2 = .40$. Iconic warnings took significantly fewer training trials to reach criterion, with a mean of 30.31 ($SD = 13.38$) training trials recorded, compared with a mean of 81.62 ($SD = 43.11$) required for abstract warnings.

Error rate. Hypothesis 1 was supported in terms of error rate, $F(1, 170) = 109.08$, $p < .05$, $\eta^2 = .39$, with significantly fewer errors being made in the learning phase in response to iconic warnings ($M = 1.42$, $SD = 2.13$) than to abstract warnings ($M = 16.01$, $SD = 13.09$).

Reaction time. Hypothesis 1 was not supported with regard to RT, $F(1, 172) = 3.80$, $p > .05$, $\eta^2 = .02$: for iconic warnings $M = 4949.96$ ms ($SD = 728.97$ ms), and for abstract warnings $M = 5173.44$ ms ($SD = 775.32$ ms).

Hypothesis 2: Iconic Versus Abstract Warnings under High and Low Demand

Manipulation check: Cognitive load. In the test phase, participants completed a concurrent visual task (simple or complex mathematical addition items) while identifying iconic or abstract auditory, visual, or A+V warning signals. To confirm that the mathematical task manipulated cognitive load, we analyzed warning recognition accuracy and RT in two separate one-way ANOVAs. Significant main effects were found for accuracy, $F(1, 171) = 5.34$, $p < .05$, and RT, $F(1, 171) = 54.77$, $p < .05$.

As anticipated, warning recognition scores were higher and RTs were faster during the concurrent low-demand task than during the high-demand dual task: Low-demand accuracy $M = 6.30$ ($SD = 0.94$) and RT $M = 4695.59$ ms ($SD = 1295.44$ ms), whereas high-demand accuracy $M = 6.11$ ($SD = 1.07$) and RT $M = 6002.12$ ms ($SD = 2686.33$ ms). As for accuracy on the mathematical addition task, participants made fewer errors during low demand ($M = 1.38$, $SD = 2.36$) than in the high-demand condition ($M = 24.34$, $SD = 10.42$).

High demand. It was hypothesized that in the high-demand condition of the dual-task test phase, accuracy would be greater and RTs faster in response to iconic warnings as compared with abstract warnings. Planned comparisons relating to Hypotheses 2 and 3 were conducted with a Bonferroni-adjusted alpha of .0025. Descriptive statistics for high-demand condition accuracy and RT are presented as histograms for Warning Type × Modality in Figures 1 and 2, respectively.

As hypothesized, accuracy was significantly greater, $F(1, 166) = 41.15$, $p < .0025$, Cohen's $d = 1.72$, and RT was faster, $F(1, 166) = 10.75$, $p < .0025$, $d = 0.75$, in response to iconic auditory warnings than to abstract auditory warnings. Iconic visual versus abstract visual warnings did not differ with respect to accuracy or RT. Iconic bimodal warnings elicited significantly greater accuracy than did abstract bimodal warnings, $F(1, 166) = 14.91$, $p < .0025$, $d = 1.21$. There was no significant difference between iconic bimodal and abstract bimodal warnings with respect to RT.

Low demand. In the low-demand condition, accuracy was also significantly greater, $F(1, 166) = 23.39$, $p < .0025$, $d = 1.11$, and RT was faster, $F(1, 166) = 27.35$, $p < .0025$, $d = 1.30$, in response to iconic auditory warnings as compared with abstract auditory warnings. Figures 3 and 4 depict the low-demand condition results as a function of Warning Type × Modality for accuracy and RT, respectively. Although iconic visual warnings elicited significantly greater accuracy than did abstract visual warnings, $F(1, 166) = 27.49$, $p < .0025$, $d = 1.45$, they did not differ with respect to RT. Similarly, iconic bimodal warnings elicited significantly greater accuracy than did abstract bimodal warnings, $F(1, 166) = 16.41$, $p < .0025$, $d = 1.21$, but they did not differ with respect to RT.

Hypothesis 3: Auditory Versus Visual Warnings, High Demand

The third hypothesis was that accuracy would be greater and RTs faster in response to auditory warnings as compared with visual warnings, as the concurrent mathematical task was visual in nature, with the effect manifesting in the high-demand condition. Planned comparisons revealed that there was no significant difference in either accuracy or RT recorded in response to iconic auditory versus iconic visual warnings in the high-demand condition. There was a significant difference in both accuracy and RT in response to abstract auditory and abstract visual warnings, $F(1, 166) = 17.98$, $p < .0025$, $d = 1.06$, and $F(1, 166) = 11.75$, $p < .0025$, $d = 0.77$, respectively. However, the abstract auditory warnings elicited poor accuracy and very slow RT (M = 8213.29 ms, SD = 3457.30 ms) relative to abstract visual warnings; see Figures 1 and 2.

The possibility that bimodal presentation facilitates accuracy and RT was not borne out in the data. As Figures 1 to 4 show, and contrasts of

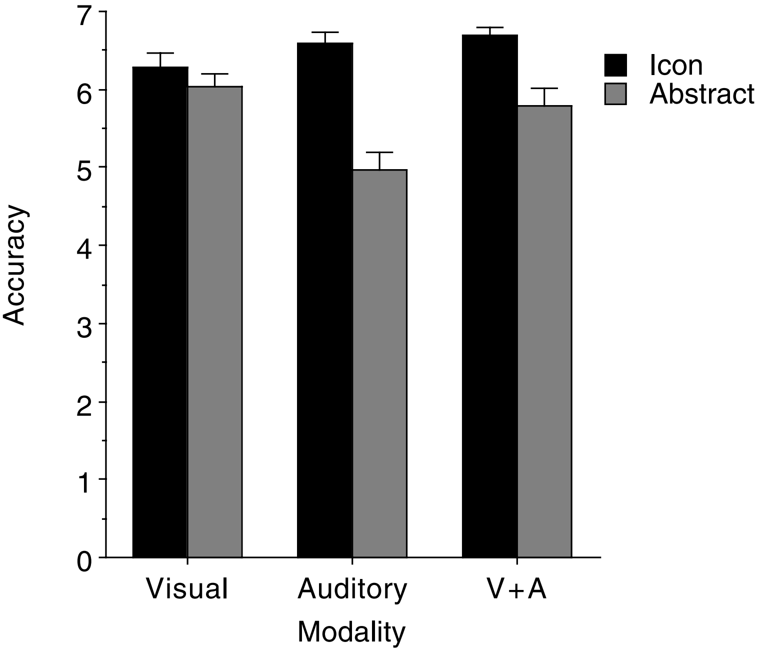

Figure 1. Test phase mean accuracy scores (maximum = 7), high-demand condition: Warning Type × Modality. Error bars refer to standard error of the mean.

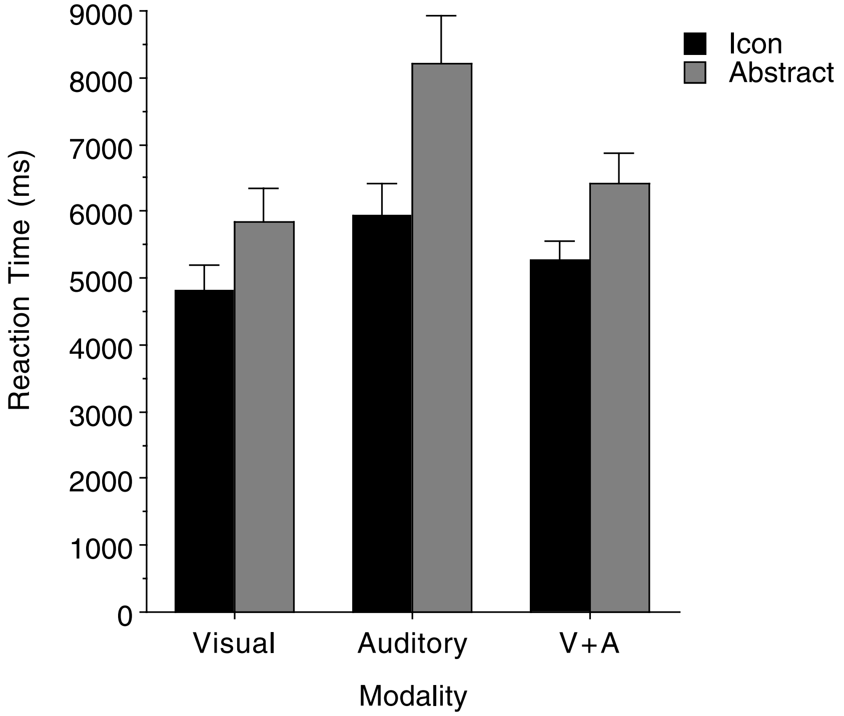

Figure 2. Test phase mean reaction times (in milliseconds), high-demand condition: Warning Type × Modality. Error bars refer to standard error of the mean.

Modality × Iconicity × Demand interactions reveal, bimodal presentation does not elicit superior performance relative to unimodal presentation.

DISCUSSION

The aim was to compare the efficacy of iconic and abstract warnings within training and dual-task test settings. Participants reached the criterion level of accuracy with significantly fewer training trials and a lower error rate in response to iconic warnings than to abstract warnings, in keeping with findings of Graham (1999), Gaver et al. (1991), and Belz et al. (1999). Abstract warnings required, on average, more than twice the number of training trials.

An explanation of the difference in learning warning-event relations in the iconic and abstract conditions is the greater familiarity and reduced semantic distance involved in the former type of signal-event relation (Isherwood, McDougall, & Curry, 2007). The source of the icon sound or image was recognized as having been heard or seen before or as being an instance of a particular category of event or object. In the iconic conditions, the semantic distance between the icon and event was small, and a semantic connection between the icon and the event was formed or a preexisting association recognized. With repeated presentation of the icon and directive instruction in the training trials, participants recognized the correct event from the choice of events relatively quickly.

By contrast, learning an association and accurately recognizing that association between an abstract warning and a critical event is an effortful process. The abstract warning sound or image needed to be analyzed and paired in memory with an arbitrarily associated event. Even with repetition and reinforcement of the association, semantic distance was great, demanding that participants use mnemonic or conscious strategies to recognize the associated event. Despite the challenge of learning abstract associations, most participants reached the criterion level of accuracy, and test phase results were generally good. The learning process, however, was slow and arduous.

A second hypothesis investigated whether the advantage for iconic over abstract warnings persisted in conditions of high and low cognitive demand. Iconic warnings, particularly auditory and bimodal, yielded a processing advantage relative to abstract warnings in both high- and low-demand

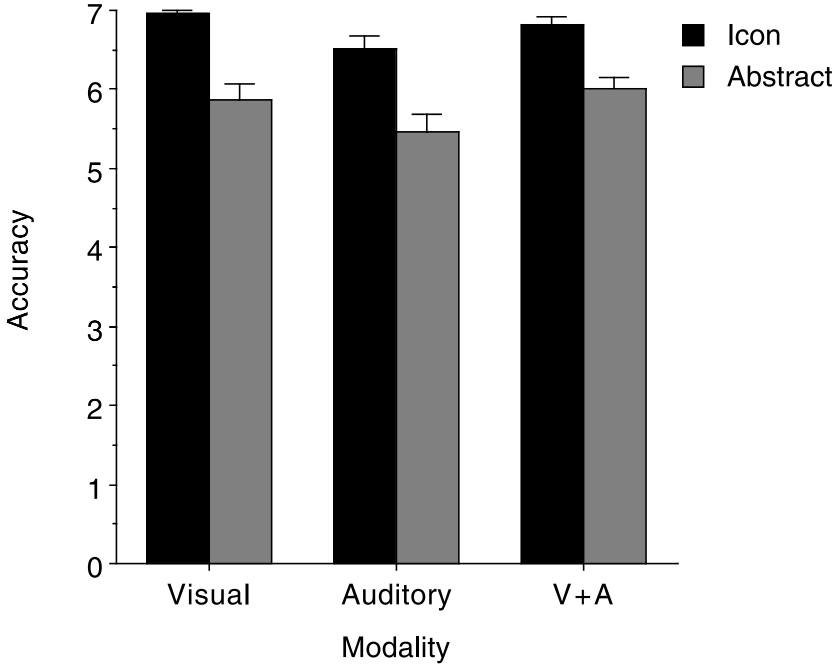

Figure 3. Test phase mean accuracy scores (maximum = 7), low-demand condition: Warning Type × Modality. Error bars refer to standard error of the mean.

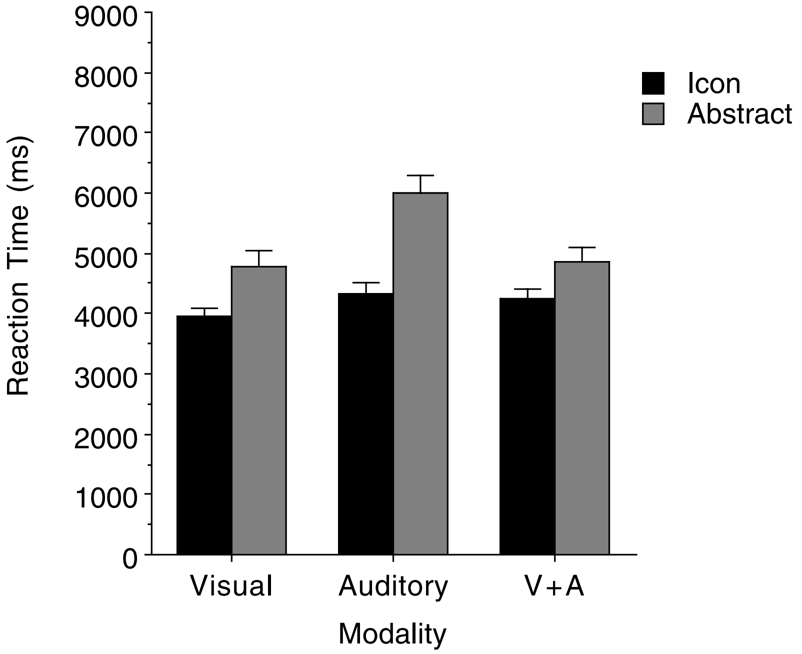

Figure 4. Test phase mean reaction times (in milliseconds), low-demand condition: Warning Type × Modality. Error bars refer to standard error of the mean.

conditions. The processing advantage is likely to be the result of familiarity (meaningful associations in long-term memory) for icon-event relations. The set of iconic warnings may have had greater variability in its visual and acoustic features than the set of novel abstract warnings, although variability was a goal in designing the abstract warning sets.

Whether there is an advantage for auditory warnings as compared with visual warnings when a concurrent task involves the visual modality was examined in Hypothesis 3. Although the concurrent task involved the visual presentation of digits, it required mental arithmetic and utterance of the answer aloud. These tasks may have involved modalities other than vision alone (e.g., subarticulation of digits, digit addition, a verbal and motor response) such that there was less definitive separation between modalities for stimulus processing and response than that required for the predictions of multiple resources theory to be upheld (Wickens & Liu, 1988).

RTs recorded in response to all warning signals in the experiment were long, from 4 to 8 s. This length of response reflects processing: The participant needs to detect the warning, identify the source (e.g., coughing), bring the association to consciousness (dangerous level of carbon monoxide), and press a button to indicate the response (Keller & Stevens, 2004). Such processing time prohibits the use of iconic signals in situations where immediacy of response and appropriate action are of the essence. However, in situations where there is a need to inform an operator of a non-time-pressured event, iconic warnings can play a role. There is also a need for iconic warnings to be investigated using realistic tasks in a simulated operational environment.

In addition to realism, the relationship between operational experience and performance in response to iconic warnings needs to be investigated. Of particular significance is the extent to which the strength of a warning-event relation is dependent on exposure to the warning, the perceived potential consequence of the warning, and/or the stage of the operation during which the warning occurs. For example, a warning for cabin depressurization may be less salient to an operator when the aircraft is on the tarmac than when it is climbing through 14,000 feet. However, it is a failure with serious consequences and an event to which an operator would have relatively less exposure. This combination of factors may impact responses to iconic warnings in an operational context.

Implications of the results include that a small set of critical event-icon mappings can be learned and that recognition of abstract auditory warnings is slow and inaccurate. Warnings, whether iconic or abstract, need to be clearly distinguishable. They require design, construction and evaluation as a set. The design must account for the particular operational environment, including ambient noise, speech, and other signals. If informed by distinctive and context-sensitive design, iconic warnings have a potential application in aviation to alert and inform operators about non-time-pressured incidents or events.

ACKNOWLEDGMENTS

The authors acknowledge the funding support provided by the Australian government, through the Australian Transport Safety Bureau's Aviation Safety Research Program (B2005/0120). We thank Clipart.com for permission to use and reproduce the seven original images and modifications.

REFERENCES

Begault, D. R. (1994). *3-D sound for virtual reality and multimedia.* Boston: AP Professional.

Belz, S. M., Robinson, G. S., & Casali, J. G. (1999). A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors, 41,* 608–618.

Calhoun, G., Janson, W. P., & Valencia, G. (1988). Effectiveness of three-dimensional auditory directional cues. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 68–72). Santa Monica, CA: Human Factors and Ergonomics Society.

Caplin, S. (2001). *Icon design: Graphic icons in computer interface design.* London: Cassell.

Chute, D. L., & Westall, R. F. (1996). Fifth generation research tools: Collaborative development with Powerlaboratory. *Behavior Research Methods, Instruments, and Computers, 28,* 311–314.

Doll, T. J., & Folds, D. J. (1986). Auditory signals in military aircraft: Ergonomic principles vs. practice. *Applied Ergonomics, 17,* 257–264.

Edworthy, J., & Adams, A. (1996). *Warning design: A research prospective.* London: Taylor & Francis.

Gaver, W. W. (1989). The Sonic Finder: An interface using auditory icons. *Human-Computer Interaction, 4,* 67–94.

Gaver, W. W. (1993). What in the world do we hear: An ecological approach to auditory event perception. *Ecological Psychology, 5,* 1–29.

Gaver, W. W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: The ARKola simulation. In *Proceedings of the Conference on Human Factors in Computing Systems: Reaching Through Technology (CHI 1991)* (pp. 85–90). New York: Association for Computing Machinery.

Gittens, D. (1986). Icon-based human-computer interaction. *International Journal of Man-Machine Studies, 24,* 519–543.

Graham, R. (1999). Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics, 42,* 1233–1248.

Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing driver's visual attention. *Journal of Experimental Psychology: Applied, 11,* 157–174.

Isherwood, S., McDougall, S., & Curry, M. (2007). Icon identification in context: The changing role of icon characteristics with user experience. *Human Factors, 49,* 465–476.

Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied, 10,* 3–12.

Liu, Y.-C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics, 44,* 425–442.

McDougall, S. J. P., de Bruijn, O., & Curry, M. B. (2000). Exploring the effects of icon characteristics on user performance: The role of icon concreteness, complexity, and distinctiveness. *Journal of Experimental Psychology: Applied, 6,* 291–306.

Momtahan, K., Hetu, R., & Tansley, B. (1993). Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics, 36,* 1159–1176.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.

Patterson, R. D. (1982). *Guidelines for auditory warning systems on civil aircraft* (Paper No. 82017). London: Civil Aviation Authority.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science, 3,* 159–177.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30,* 599–616.

Wickens, C. D., Sandry, D., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors, 25,* 227–248.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology* (Rev. ed.). New York: Henry Holt and Company.

Nathan C. Perry is a Ph.D. student at the University of Western Sydney, where he received his bachelor's degree (honors) in psychology in 2006.

Catherine J. Stevens is an associate professor in the School of Psychology and the MARCS Auditory Laboratories at the University of Western Sydney. She obtained a Ph.D. in experimental psychology from the University of Sydney in 1993.

Mark W. Wiggins is an associate professor of psychology and Head of the School of Psychology at the University of Western Sydney. He received his Ph.D. in psychology from the University of Otago, Dunedin, New Zealand, in 2001.

Clare E. Howell is a research assistant at MARCS Auditory Laboratories at the University of Western Sydney, where she graduated with a postgraduate diploma in psychology in 2003.

Date received: August 14, 2006
Date accepted: May 21, 2007