

The User Knows: Considering the Cognitive Contribution of the User in the Design of Auditory Warnings

Catherine Stevens and Agnes Petocz

School of Psychology & MARCS Auditory Laboratories,
University of Western Sydney, Locked Bag 1797,
South Penrith DC, NSW, 1797, Australia
{kj.stevens,a.petocz}@uws.edu.au

Abstract. An experiment that investigated effects of modality, warning type, and task demand on warning recognition speed and accuracy is reported. Using the experiment as a specific example, we argue for the importance of considering the cognitive contribution of the user (viz. prior learned associations) in the warning design process. Drawing on semiotics and cognitive psychology, we highlight the indexical nature of so-called auditory icons or natural indicators and argue that the cogniser is an indispensable element in the tripartite nature of signification.

Keywords: Auditory warnings, Workload, Modality, Icons, Semiotics.

1 Introduction

Caricatures of everyday sounds [1,2,3] have been considered as auditory icons. However, as we have argued elsewhere [4] the term icon, meaning likeness or image, while having straightforward application in the visual domain cannot be applied in a straightforward manner in the auditory domain. In audition, there are few true auditory icons – that is, where one sound is used to stand for another sound by virtue of resemblance between the two sounds. In the language of semiotics [5] what has been termed an auditory icon is more correctly an index. For clarity, we refer to what have been called auditory icons as *natural indicators* where these natural indicators have been adopted or adapted for purposes of conventional indication [4]. Abstract warnings where there is no prior systematic relation between signal and event are termed *symbolic indicators* (following Peirce) as their association is determined purely by convention.

In operational environments where the visual display is often very complex, auditory natural indicators have the potential to be effective warnings because they can be short, are not easily masked, are distinct from speech signals, can be used where the visual display is at risk of visual information overload, and when the critical event does not make a sound [3,6]. Reaction times are faster in response to auditory natural indicators compared with tonal and speech warnings [7] and auditory symbolic indicators [8,9].

There are two classes of natural indicators [4]. The first involves natural indicators that have been *adopted* to indicate cause or a correlated object/event. These can include indicators made by humanly manufactured objects (e.g., the sound of a car) but which may nevertheless be considered part of the environment of natural indicators (i.e., the environment into which humans are born). For example, the sound of a car failing to start is correlated with running out of fuel. The second class consists of those natural indicators that have been *adapted* to exploit naturally occurring shared features (particularly similarity of form or function) between what they naturally indicate and the selected target. For example, the sound of a sink draining (a whirlpool form) may be used to signal “tornado” [3].

Drawing upon semiotics and psychology, we argued earlier [4] that, because signification is a tripartite relation between signal, referent, and person/cogniser, auditory warning designers cannot afford to neglect the cogniser as an indispensable element. Indeed, the cognitive contribution of the user (viz. prior learned associations) appears to be the most significant factor in determining the effectiveness of warnings.

The cognitive processes involved in recognizing a natural indicator typically include recognition of the source of the sound or image and the activation of long-term memory. Symbolic indicators, on the other hand, need to be learned from the outset. For example, sound bursts with a unique set of frequencies and pause durations yielding a novel and unique pitch and temporal pattern. Features of such a sound or image need to be extracted and become associated with a particular critical event. The association can be strengthened with repeated exposure and directive training. In the context of warning design, both natural and symbolic forms of indication need to be learned but natural indicators have been learned previously and/or exploit some causal, correlational, or similarity of form or function relation. Symbolic indicators, on the other hand, are abstract and learned within an experimental session or within the idiosyncrasies of a particular operational system.

As we have noted [4], it is a truism that connections that have already been (at least partially) learned will be fully learned more easily than connections that have not been previously learned. The purpose of comparing natural and symbolic indicators here is to investigate whether this is the case when the warning set is small (i.e., within the capacity of adult working memory), and in both visual and auditory form.

If warnings are natural indicators for events that they signal, then, during training, accuracy should be greater and reaction time faster in response to natural indicators than to symbolic indicators. This should be the case for warnings presented in either the visual or the auditory modality. We would expect an advantage for natural indicators to hold not just for the learning phase, but also for a test phase in which there are additional and competing demands, as in a dual task.

1.1 Recognition of Warnings During High Workload

In studies of the effectiveness of different types of warning signals, task demand or workload is rarely investigated. Task demand is a crucial variable for the application and generalisability of results to real-world settings. An assumption, for example, that natural indicators are recognised more often and more quickly than symbolic

indicators when task demand is low and the operator is unstressed does not necessarily predict their efficacy during time pressured and/or critical situations. When critical incidents occur there are often many situations including alarms that need attention. To begin to investigate the recognition of natural indicators under more demanding conditions, in the present experiment, participants performed dual tasks – concurrent arithmetic calculations and warning recognition – with a systematic increase in the difficulty of the arithmetic task. Thus we investigate the effect of task demand on warning recognition speed and accuracy.

1.2 Aim, Design, and Hypotheses

The aim of the experiment was to investigate the effect of indicator type, modality and task demand on warning recognition speed and accuracy. The experiment consisted of a 2x2x2 factorial design: modality (auditory, visual), indicator (symbolic, natural), and task demand (high, low) with repeated measures on the latter two factors. The dependent variables were warning recognition accuracy and reaction time during learning and dual-task test phases. It was hypothesized that i) in both auditory and visual modalities, natural indicators compared with symbolic indicators elicit greater accuracy and faster reaction time during the learning phase; and ii) in both auditory and visual modalities, natural indicators compared with symbolic indicators elicit greater accuracy and faster reaction time during low and high demand conditions of the dual-task test phase.

2 Method

2.1 Participants





Forty adult participants (35 females and 5 males), from the University of Western Sydney took part in the study for which they received course credit. The mean age of the sample was 21.08 years, ($SD=5.27$, range 18-42 years). Twenty participants were presented with auditory signals and 20 with visual signals. During the test phase, 10 participants completed the low demand dual task first followed by high demand, another 10 participants completed the high demand dual task first, and 10 participants were presented initially with blocks of natural indicators followed by blocks of symbolic indicators, and 10 participants completed symbolic indicator blocks before natural indicator blocks. All participants had normal hearing and normal or corrected-to-normal vision.

2.2 Stimuli

The auditory and visual natural indicators used in the experiment were rated in a separate stimulus selection task as being highly related (means of 3.77-4.33 out of 5) with specified critical events. Ratings of association between symbolic indicators and events with which they were arbitrarily paired were low with a mean association rating of 1.56 ($SD=.72$) on a scale from *totally unrelated* (1) to *highly related* (5).

Auditory Natural Indicators. The auditory natural indicators were obtained from the websites www.sounddogs.com and www.findsounds.com. The four natural indicators were from the first class, that is, adopted to indicate cause or a correlated object/event. For example, the sound of coughing is correlated with an excess of a dangerous gas such as carbon monoxide. All sampled everyday sounds were 1 s, 16-bit mono, and standardized to a sample rate of 44.1 kHz, with normalized amplitude. Descriptions are given in Table 1.

Table 1. Visual and auditory natural indicators used as stimuli





Event	Visual Natural Indicator Description	Visual Natural Indicator	Auditory Natural Indicator Description
Low fuel	Petrol pump		Car failing to start
Carbon monoxide	Skull and crossbones		Coughing
Ground proximity	Plane diving into mountain		Explosion
Engine fire	Fire extinguisher		Fire engine siren

Auditory Symbolic Indicators. The auditory symbolic indicators consisted of tones with normalized amplitude, 44.1 kHz sample rate, 16-bit resolution, mono, and were all 1 s in duration. The sounds were designed based on guidelines set by Patterson [10] in which a burst of sound is first created and repeatedly played over the duration of the signal. Each sound burst had its own set of frequencies and pause durations, giving each warning a unique pitch and temporal pattern. The duration of the bursts varied from 0.19 s to 1 s. The tones and upper harmonics of periodic sounds were selected from frequencies in the range 150–3000 Hz. The auditory symbolic indicators stood in no obvious relation to the events with which they were paired.

Visual Natural Indicators. The visual natural indicators were obtained from the website www.clipart.com and depicted in Table 1. The visual natural indicators were designed to be similar to their auditory counterparts in the way that they related to their targets via causal/correlational based indication. For example, the image of a petrol pump was not used to indicate "petrol pump", which would have been a purely icon-based relation; instead, it was used to indicate something associated with a petrol pump (low fuel). All clipart images were shown in black and white.

Visual Symbolic Indicators. To ensure that image complexity was comparable across the set of visual natural and symbolic indicators, the visual symbolic indicators were obtained by enlarging a small section of visual images from clipart, other than those being used as visual natural indicators. All symbolic indicators were selected on the basis that there was no obvious relationship between the indicator and the event with which it was paired. All images were shown in black and white. Visual symbolic indicators for each of the four events are shown in Table 2.

Table 2. Visual symbolic indicators used as stimuli

Event	Visual Symbolic Indicator
Low fuel	
Carbon monoxide	
Ground proximity	
Engine fire	

Four critical aviation events that could potentially lead to an accident were selected. Symbolic auditory and visual indicators and natural auditory and visual indicators were designed for each of these. The critical events were presented as 'clickable' buttons on a computer screen, equidistant from one another. Warnings were presented either visually at the top of the computer screen for 1000 ms or auditorily through headphones, also lasting for 1000 ms.

A mathematical addition task, presented visually and concurrently on an adjacent computer, was constructed in the form of low and high demand conditions. The low demand version consisted of three numbers, all less than five (e.g., 1+2+3), presented in the middle of the computer screen. The high demand task consisted of two double figure numbers (e.g., 26+49). Participants were required to mentally add the numbers together and then say the answer aloud. The addition task was displayed on the screen for 2000 ms. Both low and high demand conditions consisted of a total of 16 additions. Warnings were presented intermittently with approximately four addition tasks presented to the occurrence of one warning.

2.3 Equipment

The experiment was programmed in PowerLaboratory version 1.0.3, and presented to participants on a Macintosh iBook G4. The concurrent addition task was programmed in SuperLab Pro 1.74 and presented to participants on a Macintosh Power Book G4. Auditory warnings were played through Koss stereo headphones.

2.4 Procedure

The procedure was approved by the University of Western Sydney Human Research Ethics Committee. An information sheet was distributed and participants provided written consent before the experiment began. Participants were randomly assigned to either the visual or the auditory condition and were provided with some context for each of the warnings by reading a Critical Aviation Events Information sheet. 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 randomly assigned to the indicators (symbolic or natural) that they would be exposed to and learn first. They were trained on the relation between each of the warnings and the corresponding event, then tested on the warning-event relation, until they had had a total of 16 presentations (two presentations of each of the four signals). Corrective feedback was

given after each response. This design contrasts with previous studies wherein participants have been trained to a certain criterion level of performance, for example [3,6,12]. We adopt a human factors approach involving minimal training and a stimulus set that should not exceed the capacity of adult working memory.

In the test phase, participants were required to perform a visual addition task in which numbers appeared briefly on the screen and they were required to add the numbers together as fast and as accurately as possible, while still responding to the warnings when they were presented. The visual addition task was presented through an adjacent computer, requiring that participants divide their attention between two computers. A second phase of the experiment (comprising both learning and test phases) involved the same procedure using a new set of signals from the warning type not yet tested. The experiment took 25-30 minutes.

3 Results

Descriptive statistics relating to arithmetic accuracy in the test phase are shown in Table 3. Warning recognition accuracy and reaction times on correct responses in learning and test phases are displayed in Tables 4, and 5, respectively.

Table 3. Mean accuracy on the arithmetic task in the dual task test phase (max. = 16)

Indicator	Modality	Low Demand		High Demand	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Natural	Auditory	13.75	2.05	5.60	4.64
	Visual	14.75	1.07	7.25	4.34
	Total	14.25	1.69	6.43	4.51
Symbolic	Auditory	14.30	1.84	6.00	4.35
	Visual	14.95	1.05	7.00	3.99
	Total	14.63	1.51	6.50	4.15

As a manipulation check, performance on the high and low demand versions of the concurrent arithmetic task was computed. There was a main effect of task demand, $F(1,38)=200.58$, $p<.05$, Cohen’s $d=2.44$, with a significant difference between arithmetic scores in the high demand ($M=6.46$, $SD=4.33$) and low demand ($M=14.44$, $SD=1.60$) conditions. There was no main effect of modality or indicator on arithmetic scores and no interactions between modality, indicator, or task demand factors.

Table 4. Mean warning recognition accuracy in the learning phase (max. = 16) and low and high demand test phases (max. = 4)

Indicator	Modality	Learning Phase		Low Demand		High Demand	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Natural	Auditory	14.65	1.79	3.45	0.83	3.65	0.59
	Visual	15.65	0.59	3.45	0.69	3.30	0.92
	Total	15.15	1.41	3.45	0.75	3.48	0.78
Symbolic	Auditory	10.45	2.84	2.60	1.23	2.35	1.04
	Visual	11.65	4.42	2.40	0.99	2.75	0.97
	Total	11.05	3.71	2.50	1.11	2.55	1.01

Table 5. Mean warning recognition reaction times (ms) in the learning phase and low and high demand test phases – correct responses only

Indicator	Modality	Learning Phase		Low Demand		High Demand	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Natural	Auditory	4253	357	2686	619	2119	332
	Visual	4023	407	2167	640	2261	743
	Total	4138	396	2427	675	2190	572
Symbolic	Auditory	4683	639	2353	674	2736	951
	Visual	4406	743	2104	722	2398	714
	Total	4544	698	2228	701	2567	848

3.1 Learning Phase

It was hypothesized that natural indicators compared with symbolic indicators elicit greater accuracy during the learning phase. Significantly greater accuracy was recorded during the learning phase in response to natural indicators ($M=15.15$, $SD=1.41$) than to symbolic indicators ($M=11.05$, $SD=3.71$), $F(1,38)=47.04$, $p<.05$, Cohen's $d=1.46$. There was no main effect of modality (Auditory $M=12.55$, $SD=2.32$, Visual $M=13.65$, $SD=2.15$) and no modality x indicator interaction.

It was hypothesized that natural indicators compared with symbolic indicators elicit faster reaction time during the learning phase. Reaction times (RT) of correct responses in the learning phase were significantly faster in response to natural indicators than to symbolic indicators, $F(1,38)=16.36$, $p<.05$, Cohen's $d=0.72$; natural indicators $M=4138.10$ ms ($SD=395.80$) and symbolic indicators $M=4544.40$ ms, ($SD=698.31$). There was no main effect of modality (Auditory $M=4469.08$, $SD=498.13$, Visual $M=4214.42$, $SD=575.31$) and no modality x indicator interaction.

3.2 Test Phase

The second hypothesis was that natural indicators compared with symbolic indicators elicit greater accuracy during high and low demand conditions of the dual-task test phase. In the high demand condition, accuracy was significantly greater in response to natural indicators than to symbolic indicators in both the auditory modality, $F(1,19)=24.51$, $p<.05$, Cohen's $d=1.54$ and visual modality, $F(1,19)=5.49$, $p<.05$, Cohen's $d=0.58$. Similarly, in the low demand condition, accuracy was significantly greater in response to natural indicators than to symbolic indicators in both the auditory, $F(1,19)=7.51$, $p<.05$, $d=0.81$, and visual modalities, $F(1,19)=12.72$, $p<.05$, $d=1.23$. There was a significant interaction in recognition accuracy between modality, indicator, and task demand, $F(1,38)=6.36$, $p<.05$.

It was hypothesized that natural indicators compared with symbolic indicators elicit faster reaction times during high and low demand conditions of the dual-task test phase. In the high demand condition, RTs were significantly faster in response to natural indicators than to symbolic indicators in the auditory modality, $F(1,19)=8.25$, $p<.05$, Cohen's $d=0.87$, but there was no significant difference between natural and symbolic indicator RTs in the visual modality. In the low demand condition and opposing the directional hypothesis, RTs were significantly faster in response to symbolic indicators than to natural indicators in the auditory modality, $F(1,19)=4.73$,

$p < .05$, $d = 0.51$. There was no difference between natural and symbolic indicators with respect to RT in the low demand, visual modality condition. There were three significant interactions involving RT: modality x task demand, $F(1,38) = 4.38$, $p < .05$; indicator x task demand, $F(1,38) = 10.08$, $p < .05$; and modality x indicator x task demand, $F(1,38) = 4.28$, $p < .05$.

4 Discussion

This experiment investigated the effect of indicator type, modality of presentation and task demand on warning recognition speed and accuracy during training and dual task test phases. As hypothesized, during the learning phase, natural indicators (or caricatures of everyday sounds and objects) relative to symbolic indicators elicited greater recognition and faster RTs. This pattern of responding was upheld in test phase conditions with respect to accuracy. The pattern was observed in RT only in the test phase involving high demand and the auditory modality.

Results from the learning phase provide support for the general hypothesis that in learning signal-event relations there is an advantage for those natural indicators that have been previously learned and are now being exploited for indication [4]. This pattern occurs even when a set of just four warnings is used, corroborating findings of others [6,8,9,11,12]. Natural indicators elicit recognition of the source of the sound or referent of the image, and activate associations in long-term memory. Symbolic signal-event relations need to be learned within an operational context. While RTs in the present experiment were relatively slow in response to warnings during training, they improved in the dual-task test phase.

The expected pattern of results has been obtained in accuracy but not fully in RT scores of the test phase. Natural indicators elicited significantly faster RTs in the high demand task with warnings in the auditory modality but not in the visual modality and not in the low demand condition. One explanation derives from scrutiny of the concurrent task arithmetic scores and an apparent weighting of tasks by participants. The mean arithmetic score in the low demand condition was approximately 14 out of 16 whereas in the high demand condition the maximum mean was around 7 out of 16. Participants performed well on simple addition tasks to the detriment of speed in recognizing auditory natural indicators. That this occurred only in the auditory modality may be attributable to the greater degree of possible ambiguity in the auditory natural stimuli. For example, the sound of coughing was used as a natural auditory indicator for carbon monoxide. However, it is a natural indicator of many other situations such as a person with a cold, a person choking, a person smoking heavily or (significantly for the present context) that there is a(n) engine fire. The same can be said for the sound of a car failing to start and the sound of an explosion. Of the four auditory natural indicators used in the experiment, the fire engine siren is perhaps the only one that is as closely connected to its target and as unconfounded with other possible connections as is its visual counterpart (the image of a fire extinguisher). In contrast, the visual natural indicator of a petrol pump is a typical fuel indicator in motor vehicles, and the skull and crossbones are a familiar indicator of poison. Visual natural indicators thus may already be better learned at the start of the experiment than are auditory natural indicators, and may also be more distinctive in

the sense of being less liable to confounding with other prior learned associations. Of course, there are several other factors which may have contributed to the visual advantage (e.g., temporal differences in stimulus registration). However, it is clear from other research [4] that, in general, the cognitive contribution (viz. prior learning) of participants and users must also be considered in warning design.

In the high demand condition, participants performed poorly on the mental arithmetic, possibly guessing, and leaving more cognitive resources for relatively fast warning recognition. The symbolic auditory indicators under low demand elicited more optimal performance with good arithmetic accuracy and, relative to the natural indicators, faster RTs. However, under conditions of high demand, arithmetic was again poor and auditory symbolic indicators were recognized slowly. The use of a demanding concurrent task has brought into relief the potentially complex and operationally important interaction between task load and indicator type.

A further advantage of the auditory modality may manifest when coupled with a visually-presented arithmetic task. However, the present results do not suggest interference from visually-presented indicators when the learning phase visual versus auditory modality accuracy scores are compared with test phase visual versus auditory modality scores. Similarly, in arithmetic scores there was no significant effect of modality.

The present experiment was designed specifically to contrast the design used by Perry et al [12] and, in the spirit of a human factors approach, to keep training to a minimum. Thus, rather than training all participants to 100% criterion level of performance as we have done in the past, all participants were exposed to a set number of training trials. The present results suggest that training to a criterion level of performance may be important especially in the case of symbolic indicators.

While there was no main effect of task demand in the test phase, demand did interact with modality in both indicator recognition and RT scores. This provides evidence of the need to examine the setting in which warnings will be used not only from the perspective of ambient noise, potential maskers, and the complexity of existing auditory and visual displays, but also the nature of the operational task and the load that it incurs. There is need also for effects of demand and indicator to be investigated in settings that approach the operational context such as an Advanced Aviation Training Device (AATD).

The importance of context in warning design is underscored by the present results. The artificial environment is always embedded, both physically and psychologically, in the natural environment that includes the learned associations of the user. Not only is it a truism that an already-learned association will be more easily and quickly learned than one which has not yet been learned, but it is also true that if the referent is just one of a large set of equally salient associations, such that learning the connection requires the learner to "unlearn" or ignore those other meanings, then a natural indicator may actually be *less* effective than a newly designed symbolic indicator that is free of "excess baggage". Thus, while exploiting natural indicators is a good idea, there are advantages and disadvantages. On the other hand, the user is likely to bring *some* association even to symbolic connections. In the present experiment, ratings of symbolic associations were typically higher than zero.

These observations suggest that the cognitive contribution of the user is not just one among many equally salient and cumulatively contributing factors to be taken

into account for the design of warnings. Instead, it infiltrates other factors which are often treated independently as in perceived stimulus complexity, meaning, semantic distance, perceived aesthetic appeal, and so on.

References

1. Ballas, J.A.: Common factors in the identification of an assortment of brief everyday sounds. *J. Exp. Psychol: Human Perception and Performance* 19, 250–267 (1993)
2. Gaver, W.W.: The Sonic Finder: An interface using auditory icons. In: *Human-Computer Interaction*, vol. 4, pp. 67–94 (1989)
3. Keller, P., Stevens, C.: Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *J. Exp. Psychol: App.* 10, 3–12 (2004)
4. Petocz, A., Keller, P., Stevens, C.: Auditory warnings, signal-referent relations and natural indicators: re-thinking theory and application. *J. Exp. Psychol: Applied* 14, 165–178 (2008)
5. Peirce, C.S.: In: Hartshorne, C., Weiss, P. (eds.) *Collected Papers of Charles Sanders Peirce*, vol. II. Harvard University Press, Cambridge (1932/1960)
6. Stephan, K.L., Smith, S.E., Martin, R.L., Parker, S.P.A., McAnally, K.: Learning and retention of associations between auditory icons and denotative referents: implications for the design of auditory warnings. *Human Factors* 48, 288–299 (2006)
7. Graham, R.: Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics* 42, 1233–1248 (1999)
8. Belz, S.M., Robinson, G.S., Casali, J.G.: A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors* 41, 608–618 (1999)
9. McKeown, D., Isherwood, S.: Mapping candidate within-vehicle auditory displays to their referents. *Human Factors* 49, 417–428 (2007)
10. Patterson, R.D.: *Guidelines for Auditory Warning Systems on Civil Aircraft* (Paper No. 82017). Civil Aviation Authority, London (1982)
11. Gaver, W.W., Smith, R.B., O’Shea, T.: Effective sounds in complex systems: The ARKola simulation. In: *Conf. on Human Factors in Computing Systems (CHI 1991)*, pp. 85–90. Association for Computing Machinery, New York (1991)
12. Perry, N., Stevens, C., Wiggins, M., Howell, C.: Cough once for danger: An experimental investigation of auditory icons as informative warning signals in civil aviation. *Human Factors* 49, 1061–1071 (2007)