Journal of Experimental Psychology: Applied
2008, Vol. 14, No. 2, 165–178

1076-898X/08/$12.00 DOI: 10.1037/1076-898X.14.2.165

Auditory Warnings, Signal-Referent Relations, and Natural Indicators: Re-Thinking Theory and Application

Agnes Petocz
University of Western Sydney

Peter E. Keller
Max Planck Institute for Human Cognitive and Brain Sciences

Catherine J. Stevens
University of Western Sydney

In auditory warning design the idea of the strength of the association between sound and referent has been pivotal. Research has proceeded via constructing classification systems of signal-referent associations and then testing predictions about ease of learning of different levels of signal-referent relation strength across and within different types of auditory signal (viz., speech, abstract sounds, and auditory icons). However, progress is hampered by terminological confusions and by neglect of the cognitive contribution (viz., learning) of the person or user. Drawing upon semiotics and cognitive psychology, the authors highlight the indexical (as opposed to iconic) nature of so-called auditory icons, and the authors identify the cogniser as an indispensable element in the tripartite nature of signification. Classifications that neglect this third element, defining signal-referent relation strength only dyadically, yield results confounded by learning; classifications that correctly include the triadic relation yield research predictions that are redundant. These limitations of the standard method of constructing and testing classification systems suggest that auditory warning design must take the cognitive contribution of the user into account at an earlier stage in the design process.

Keywords: auditory icons, meaning, signal-referent relations, natural indication, learning

The differences between natural and artificial environments pose particular difficulties and challenges for the design and use of auditory warnings in artificial environments. In the natural environment, organisms have evolved to be sensitive to a wide variety of sounds, some of which indicate the existence or pending arrival of some threatening object or state of affairs. During development, learning results in the gradual increase of the subset of sounds that come to serve as auditory warnings, and in the elaboration of the innate reactions (i.e., the startle reflex and the automatic "fight-or-flight" response) into a diverse set of behavioral responses.

Artificial environments differ from the natural environment in several respects. They typically involve human-machine interfaces in purpose-built operating systems that are, simultaneously, embedded within the natural environment. In addition, artificial environments are relatively closed and controlled, and they are introduced at relatively late stages of the operator's development, when most of the signal-referent relations in the natural environment have already been learned. The warning signals in artificial environments do not evolve from the artificial environment, but are introduced into it by the designer. The success of these signals is usually measured in terms of how quickly and efficiently the operator can adjust to them.

It is gradually becoming clear that the most effective way to approach auditory-warning design for artificial environments is to combine acoustic and psychoacoustic knowledge with a general strategy to exploit as much as possible features of the natural environment and their psychological consequences (viz., perceived urgency, learned associations, etc.). However, the specific details of how best to do this have not yet been fully worked out.

According to Edworthy (1994), the acoustic and psychoacoustic problems are largely solved, albeit widely neglected in most auditory warning systems currently in operation. Clear guidelines have been developed (e.g., Patterson, 1982, 1989), and a substantial body of knowledge relevant to various design issues has been established (see Edworthy & Hellier, 2006a for a comprehensive review). In contrast, the *psychological* issues pertaining to auditory warning design have proven to be more difficult and, perhaps for that reason, have been generally neglected. It is fast becoming

Agnes Petocz and Catherine J. Stevens, School of Psychology & MARCS Auditory Laboratories, University of Western Sydney; Peter E. Keller, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig.

This research was supported by a University of Western Sydney Industry Links Grant awarded to Stevens and Keller working in collaboration with the Defense Science and Technology Organisation. The research was also supported by a 5-month period of research secondment awarded to Petocz by the College of Arts at the University of Western Sydney. We thank Simon Parker, Russell Martin, and Ken McAnally for seeding our interest in this area, and Glenn Newbery, Judy Edworthy, Siné McDougall, and an anonymous reviewer for helpful comments on an earlier draft.

Correspondence concerning this article should be addressed to Agnes Petocz, School of Psychology & MARCS Auditory Laboratories, University of Western Sydney – Bankstown, Locked Bag 1797, South Penrith DC NSW 1797 Australia. E-mail: a.petocz@uws.edu.au

evident that, if the design and implementation of auditory warnings is to move along efficiently and effectively, these issues can no longer be avoided. Edworthy and Hellier (2006b) note that the field currently suffers from a failure to "take account of the cognitive system of the user" (p. 12). They also observe that the auditory field appears to be lagging behind the visual, as evidenced by the lack of a generally agreed taxonomy for sound.

The broad aim of the present paper is to address this more neglected psychological side of the field and to show how the taxonomical problems and the role of the cognitive system of the user are interrelated: specifically, we aim (a) to review critically current theory and research on signal-referent relations in auditory warnings; (b) to identify and discuss a number of difficulties and confusions that we believe are hindering progress; (c) to offer clarifications and solutions by drawing upon material from related (but largely neglected) areas; and, as a result, (d) to present a more integrated, workable picture, leading to more fruitful directions for research and application.

Review of Theory, Research, and Application

The Use of Auditory Warnings in Artificial Environments

Interest in the use of auditory signals to convey warning messages in diverse types of human-machine interface was stimulated originally by the need to overcome problems associated with overloading the human operator's visual system (Cooper, 1977; DuRoss, 1978). However, inherent advantages of the auditory presentation of information have gained recognition gradually in human-machine interface design. Apart from the fact that reaction times to auditory stimuli are typically faster than to visual stimuli (e.g., Galton, 1899; Sanders, 1998; Woodworth & Schlosberg, 1954), sound is often the most effective medium for conveying information. Auditory signals can be attention-grabbing regardless of the operator's visual focus, accommodating operator mobility, allowing events both inside and outside the operator's field of view to be monitored (Burt, Bartolome, Burdette, & Comstock, 1995; Calhoun, Janson, & Valencia, 1988; Doll & Folds, 1986; Edworthy & Adams, 1996; Edworthy & Stanton, 1995; Gaver, 1989; Haas & Casali, 1995; McKinley & Ericson, 1997; Stokes & Wickens, 1988; Szalma et al., 2004). Furthermore, some dynamically changing events (e.g., approaching automobiles) are more readily represented in auditory than in visual displays (Gaver, 1989, 1993).

Acknowledgment of auditory advantages has led to widespread use of auditory warnings in a variety of domains. These include aviation (Begault & Pittman, 1996; Begault & Wenzel, 1992; Guillaume, Pellieux, Chastres, & Drake, 2003; M. W. Haas, 1995; McKinley & Ericson, 1997; Naish, 1990; Patterson, 1982), hospital alert systems (Momtahan, Hetu, & Tansley, 1993), computer interfaces (Blattner, Sumikawa, & Greenberg, 1989; Gaver, 1986, 1989, 1993; Jacko & Rosenthal, 1997), automotive safety (Dingus, McGehee, Manakkal, Jahns, Carney, & Hankey, 1997; Graham, 1999; Liu, 2001; Maltz & Shinar, 2004; McKeown, 2005), and equipment for users with disabilities (Edwards, 1995).

The variety in operating system characteristics necessitates the use of different types of sound as auditory warnings. Auditory signals have been grouped into three basic categories: (a) *speech*, (b) *abstract sounds* (e.g., simple tones or tone combinations) (Patterson, 1982, 1989), and (c) *auditory icons* (i.e., everyday environmental sounds, which rely on already-made associations) (Belz, Robinson & Casali, 1999; Edworthy & Adams, 1996; Gaver, 1986, 1989, 1993; Graham, 1999; Stephan, Smith, Martin, Parker, & McAnally, 2006).

Choosing Auditory Warnings on the Basis of Signal-Referent Relation Strength

In addition to operating system characteristics, one factor which is widely recognized to be important in selecting sounds for use as auditory warnings is what might be termed *signal-referent relation strength*, that is, the strength of the relationship, or association, between a given sound signal and the critical information to which it refers. Over the last two decades, the standard method has been to propose and test various classification systems, in an effort to establish the relative efficacy of different types of auditory signal that supposedly vary vis-à-vis the feature of signal-referent relation strength. Empirical studies have investigated ease of learning and recognition of various types of auditory signal (e.g., Belz et al., 1999; Edworthy & Hards, 1999; Graham, 1999; McKeown & Isherwood, 2007; Stephan et al., 2006), taxonomies, theories and processing models have been proposed and developed (e.g., Blattner, Sumikawa, & Greenberg, 1989; Edworthy & Hards, 1999; Familant & Detweiler, 1993; Gaver, 1986, 1993; Guillaume et al., 2003; Keller & Stevens, 2004; Maltz & Shinar, 2004; McAdams, 1993), and predictions derived from them have been further tested (e.g., Edworthy & Hards, 1999; Keller & Stevens, 2004).

Classification Schemes and Research Comparing Speech, Abstract Sounds, and Auditory Icons

In his seminal work on classification, Gaver (1986, 1989, 1993) proposed three types of perceptual (visual and auditory) mapping, based on the dimension of *articulatory directness* (Hutchins, Hollan, & Norman, 1986), that is, the *degree* to which the form of the signal or representation is constrained by its relationship to the represented object or event. In *nomic* (later, *iconic*) mappings, which involve the highest degree of articulatory directness, "meaning depends on the physics of the situation" (Gaver, 1986, p. 170). For example, the sound of a metal object being struck can be used to indicate the size of an object, because the pitch of the sound is a direct causal result of the size of the object. Here, it is the *causal* relation, rather than icon-object *similarity*, that is fundamental. *Metaphorical* mappings, in contrast, are less physically constrained and thus have less articulatory directness. They do not depend on physical causation. Instead, they "make use of similarities between the thing to be represented and the representing system" (p. 170). These include structure-mappings (such as that between genealogy and a tree) and metonymic mappings (such as the use of a hiss to stand for a snake). Finally, *symbolic* mappings are "entirely unconstrained in terms of their form" (Gaver, 1989, p. 88), and so involve the least articulatory directness. They are essentially arbitrary and rely on social convention for their meaning. Examples in the auditory domain are telephone bells and ambulance sirens—abstract sounds that acquire meaning through cultural learning.

According to Gaver, the stronger the signal-referent relation in terms of articulatory directness, the more intuitively obvious will it be, and the more easily will it be learned and recalled. Hence, the testable prediction is that nomic mappings should be relatively simple to learn, metaphorical mappings somewhat harder, and symbolic mappings the most difficult. Thus, Gaver recommends developing *auditory icons*, which are "natural sounds with a nomic or metaphorical mapping to the information to be represented" (Gaver, 1986, p. 172). Of course, speech seems to be an obvious counterexample, but we shall return to this later.

Several other classification schemes have been proposed following Gaver's line. For example, Edworthy and Hards (1999) replaced articulatory directness with level of *abstractness* of sound type. At one extreme, *concrete* sounds are everyday environmental sounds (e.g., the sound of waves, the sound of footsteps on leaves). *Semiabstract* sounds are "representative or "iconic" to the extent that they can be used to signify changes in a physical parameter by changes in their pitch, level, or some other acoustic parameter" (p. 608). Finally, *abstract* sounds are those whose physical attributes are entirely unconnected to their referents, (e.g., pure tones or alarms used in hospital medical apparatus). This concrete-abstract dimension is found also in the classification of visual icons (McDougall, Curry & de Bruijn, 1999, 2001; McDougall, de Bruijn, & Curry, 2000).

Others have developed somewhat different classifications. For example, Blattner et al. (1989) introduced a scheme that included both auditory icons and abstract sounds, but they focused on the latter, describing how abstract sounds (*earcons*) can be manipulated to form hierarchical families of music-like motifs for representing compound messages. Familant and Detweiler (1993) provided a more detailed taxonomy of signal-referent relations, in which the major focus was a distinction between *direct* and *indirect* forms of reference. In *direct reference*, the signal refers directly to the denotative referent, which is the event that is the target of the warning or message. This relation may be *iconic* (based on similarity, rather than, as for Gaver, causal connection) or *symbolic* (arbitrary, based on convention). In contrast to direct reference, *indirect reference* indicates the situation in which the signal refers to the denotative referent through an intermediate, or surrogate, *sign referent*, usually because the denotative referent may be difficult to portray. Although this taxonomy was developed for the *visual* domain, its distinction between direct and indirect referent relations has been considered promising for the auditory domain (Keller & Stevens, 2004; McKeown & Isherwood, 2007).

Investigations of the differences in terms of learning and retention between the three broad categories of auditory warning (speech, abstract sounds, and auditory icons) have generally shown an advantage for speech and auditory icons over abstract sounds. For example, Leung, Smith, Parker, and Martin (1997, reported later in Stephan et al., 2006) measured how many trials it took to learn the relationship between eight target objects or events (e.g., ship, missile-launch, unknown threat) and associated speech ("boat," "launch," "unknown"), auditory icons (the sound of a fog horn, a monkey screech, a bird territorial call), and abstract sounds. Abstract sounds were generated according to Patterson's (1982) recommendations. It was found that abstract sounds were learned and retained fairly poorly, whereas performance with speech and auditory icons was very good. In a more extensive study by Ulfvengren (2003), a similar advantage was found for speech and icons relative to abstract sounds and other arbitrary connections. The superiority in effectiveness of auditory icons has also been supported in studies using simulator-based automobile driver environments (e.g., Belz et al., 1999; Graham, 1999).

Within the set of auditory icons differences have also been found. For example, it was observed in the Leung et al. (1997) study that some of the icons were learned and retained more easily than others. The authors speculated, following Gaver (1986), that pairings of auditory icons and events vary systematically in terms of signal-referent relation strength.

Extending Research on Auditory Icons and Explaining Their Superiority

In a more recent investigation which focused solely on auditory icons, but which examined which types of icon are more easily learned and recognized, Keller and Stevens (2004) drew on a combination of the Gaver (1986) and the Familant and Detweiler (1993) taxonomies to present a conceptually reorganized classification scheme. This scheme includes, in addition to Gaver's three types of signal-referent mapping, a category of *ecological* mapping, where target and referent are related by virtue not of causal relation or physical similarity but of correlation in the natural environment (e.g., a *seagull* is used as a signal for *boat*).

In the Keller and Stevens three-tiered classification scheme (Keller & Stevens, 1999; Keller & Stevens, 2004, pp. 4), the lowest tier (Level I) specifies the three familiar types of signal (auditory icons, speech and abstract sounds). The second tier (Level IIa) specifies the types of *direct* relation between a signal and its immediate referent. These are assumed to exist along a dimension of iconicity, at the high end of which are sound signals that are related to their referents causally, as in sounds made by the referent event in the everyday world. At the lower end of the iconicity continuum, the signal-referent relation is manufactured, based either nonarbitrarily on some physical similarity or appropriateness, or purely arbitrarily, with no recognizable connection between signal and referent. The top tier (Level IIb) distinguishes two types of *indirect* relation, associations between a surrogate and target referent. These two types are ecological and metaphorical, the latter involving some equivalence of appearance and/or function (e.g., the sound of a mosquito for the target helicopter, based on the similarity of form of the two objects).

In a set of three experiments, Keller and Stevens expanded Gaver's concept of auditory icon to include their ecological and random signal-referent mappings, and tested various hypotheses regarding the relative learnability (as measured by associative strength ratings, amount of exposure required for learning, and target recognition response times) of four types of signal-referent relation (one direct and three indirect).

Keller and Stevens found that: (a) Rated association strength was significantly higher for ecological than for metaphorical relations, even though the latter elicited good agreement among participants in identifying an association between the target and surrogate referent; (b) when comparing the four levels of signal-referent relation (direct, indirect-ecological, indirect-metaphorical, and indirect-random), performance in general was best with direct relations, worst with random relations, and roughly equivalent for indirect-ecological and indirect-metaphorical relations, despite the higher ratings of association strength for ecological relations; (c) recognition of metaphorical and ecological relations, *once they had been learned*, was just as rapid and robust as recognition of direct relations, whereas random relations, even when learned as fully as possible within the time constraints of the experiment, continued to elicit significantly longer recognition latencies; and (d) metaphorical and ecological relations remained equivalent under conditions (viz., reversal of target and surrogate referents) designed to control for possible differences in iconicity and identifiability of sounds.

Explanations of the superiority and benefits of certain types of auditory icon vary. However, they all appeal to something analogous to Gaver's concept of signal-referent relation strength based on the intuitive links harnessed by articulatory directness, and they tend to agree that the strongest of such relations is that between sounds and their known direct causes. This relation is considered to be analogous to the visual concepts of *concreteness* (i.e., the extent to which the icon depicts real-world objects via shared features) and *semantic distance*, defined as the "strength of the icon-referent relation" (McDougall et al., 2001, p. 62).

Conflicting Findings, Reservations, and Mixed Recommendations

Despite the observed advantages of auditory icons, not all research has resulted in the same positive findings and the same level of enthusiasm. Here, we present some of these conflicting findings, without attempting to make sense of them. For example, Stanton and Edworthy (1998) compared the efficacy of new auditory icons to old abstract warnings in an intensive treatment unit (ITU) context, and confirmed their suspicion that recognition performance for ITU staff would be *worse* for the auditory icons relative to the abstract warnings, whereas the reverse would be the case for non-ITU staff. In the Graham (1999) automobile study using an in-vehicle collision-avoidance system, although auditory icons produced significantly faster reactions times than did abstract sounds, they also led to more inappropriate responses comprising "false positives," where drivers reacted with a brake press to noncollision situations.

Furthermore, Stanton and Edworthy (1999) note that one important feature of auditory warning signals is perceived urgency. Although the referents of auditory icons may be identified easily, the warnings may no longer convey a sense of urgency of the kind that can be conveyed via abstract sounds with the right acoustic properties. As Noyes, Starr, and Kazem (2004) point out, "there is a need for the warning system to depict accurately the nature and criticality of the problem" (p. 146). Some feel that perhaps concrete (i.e., real environmental) sounds serve informative functions better than they serve alerting or warning functions, so that "real sounds . . . may not work well as alarm sounds" (Edworthy & Hards, 1999, p. 617). This is consistent with the claim that "attempts to incorporate auditory icons as warning signals into complex systems have been largely unsuccessful" (Belz et al., 1999, p. 611).

In any case, the difference between auditory icons and abstract sounds is, just like "concreteness effects" for visual icons (McDougall, et al., 2000, p. 296; McDougall et al., 2001), only temporary, and it disappears completely when participants are allowed to generate their own cues. When users are allowed free rein to construct their own associations, auditory icons are not more easily learned and recalled than any other class of sound (Edworthy & Hards, 1999). This is useful information for the designer, especially given that many people are sceptical and uncomfortable with the introduction of auditory icons (Belz et al., 1999), seeing them as strange and somewhat gimmicky (Gaver, 1989; Graham, 1999).

Not surprisingly, the recommendations in the literature are as mixed as are these research findings. According to Edworthy and Hellier (2006b), "the broad category of sounds classed as "auditory icons" appears to be worth developing in clinical applications," because "there is considerable evidence that sounds which bear a closer relationship with their referent are easier to learn than ones which do not" (p. 14). With respect to the direction of this development, Stanton and Edworthy (1999) suggest, in accordance with their discussion of auditory *affordances*, that the design of auditory warnings should capitalize not on the link between the sound and its source, but, rather, on the link between the sound and its *potential for action*; referents should be the *action required* of the operator (e.g., a warning might signal a nurse to change a syringe pump, or a pilot to extinguish a fire). In contrast, Guillaume et al. (2003) argue that efficient alarms are those that "directly evoke a mental representation with an appropriate degree of urgency" (p. 211), and "ideally, all alarms should evoke a direct link with their cause," because these automatically elicit the appropriate behavior. They add that the common use of abstract warning signals need not be cause for despair, as a period of learning may create the appropriate mental representation eventually leading to automatic processing.

Indeed, with adequate training and involvement in design of warnings, the auditory icon advantage may disappear (Edworthy & Hards, 1999). Since new signal-referent relations *can* be learned, especially if the recommended "user-centered" approach (Edworthy & Stanton, 1995) is adopted, the opportunity exists to make use of sounds that exploit natural responses to acoustic properties contributing to perceived urgency (Stanton & Edworthy, 1999). Yet there may be limits to the effects of training. Keller and Stevens (2004) interpreted their findings to indicate that new signal-referent relations, if based on *random* connections, *cannot* be learned as effectively (vis-à-vis eventual speed of recognition) as those based on other types of relation. Either way, the design problem needs to be supplemented by appropriate methods of evaluation, for which a framework is yet to be developed (Edworthy, 1994; cf. Barker & van Schaik, 2000).

Making Sense of Theory and Empirical Research

This situation prompts a number of questions. What sense can be made of the wealth of (sometimes conflicting) material in the literature? Is it possible to explain the apparent anomalies? How adequate are the various classifications? If the aim of this field is to use research to develop guidelines for the effective design and implementation of auditory warnings in applied settings, is there any problem with the standard method of proposing and testing classification systems? How might the theories and empirical research results best be organized and, where appropriate, integrated so as to provide clearer lines for future development?

Of course, drawing general conclusions from so many different studies can be misleading, since some of the conflicting findings may be explained simply by the different methods, designs and choice of dependent measures. However, that is not the main obstacle. It is our contention that satisfactory answers to the questions raised require discussion and clarification of some problems and challenges, which have, as yet, received insufficient attention. As we show, these problems all converge onto a single factor: the neglect of the cognitive contribution (viz., learning) of the operator, and of its role in the concept of signal-referent relation strength. This deceptively simple point has widespread ramifications across the field.

Problems and Clarifications

Confusions in Terminology: Icon Versus Index and the Definition of Auditory Icon

An initial difficulty (that has been discussed briefly by Edworthy and Hellier, 2006a) is that the terminology in the relatively new field of auditory warnings has been developed largely in isolation from established definitions and classifications of signs and symbols in the field of semiotics (e.g., Alston, 1967; Eco, 1973; Hawkes, 1977; Langer, 1942; Morris, 1946; Peirce, 1932; Sapir, 1959; Skorupski, 1976). The semiotic literature is vast, complex, and disorganized, and sometimes carries an unwieldy amount of philosophical baggage. Yet, it does highlight a number of important features and distinctions, such as reference versus representation, indication versus substitution, denotation versus connotation, attributes/properties versus functions/relations, and metaphor versus metonymy. These tend to be neglected or confused in the auditory warning literature. Further confusions result from attempts to transfer to the auditory domain a set of terms that already have clear, but different, meanings in the visual domain.

To begin with, Gaver's classification of perceptual mappings parallels *conceptually*, but not *terminologically*, the influential classical tripartite division of signs offered by Charles Peirce (1932), the American founder of semiotics. Gaver (1989, p. 87) claims to be following Peirce's use of the term "icon," but he does not, in fact, do so in the strictest sense.

For Peirce, the genus *sign* is defined as "something that stands to somebody for something in some respect or capacity" (Peirce, 1932, 2.228). The sign is divided into three species, based on similarity, causal connection, or mere use.

An *icon* is "a sign which refers to the Object that it denotes merely by virtue of characters of its own" (2.247)—typically, similarity of shape or attributes (Peirce's earlier term for "icon" was "likeness"). Thus, an icon "has no dynamical connection with the object it represents; it simply happens that its qualities *resemble* [italics added] those of the object, and excite analogous sensations in the mind for which it is a likeness. But it really stands unconnected with them" (2.299). Examples are pictures and diagrams (however schematic), including the ideographs found in writing such as Egyptian hieroglyphics.

In contrast, an *index* is "a sign which refers to the Object that it denotes by virtue of being readily affected by that Object" (2.248)—that is, an indicator by virtue of a natural causal connection. Peirce offers the examples of "a rap at the door as indicative of a visitor" (2.92), or a piece of mold with a bullet-hole in it as a sign of a shot, "for without the shot there would have been no hole; but there is a hole there, whether anybody has the sense to attribute it to a shot or not" (2.304).

Lastly, a *symbol*, as implied by the original meaning of the term (i.e., "something thrown together"), is "a sign which is constituted a sign merely or mainly by the fact that it is used and understood as such" (2.307); that is, where the connection is arbitrary or conventional, as in the paradigmatic case of the symbols of language.

Consistent with Peirce's usage, the word *icon* (derived from the Greek *eikon*, meaning "likeness" or "image," such as a picture or a statue) has straightforward application in the *visual* domain. This is recognized by Familant and Detweiler (1993) in their visual taxonomy. An icon involves pictorial similarity, based on shared features. For example, a picture or a plastic model of a gun is said to have high iconicity owing to the high degree of visual resemblance. However, when the concept of an icon is applied in the *auditory* domain, the situation becomes much less straightforward. As we shall see, this may play some role in the perception that the auditory field lags behind the visual.

In the auditory warning literature at large, there are two anomalies that lead to confusion concerning the definition of an auditory icon. The first anomaly is that the term *icon* is used generally to refer to what Pierce describes as *index*. A high level of iconicity (Gaver's *iconic* or *nomic* mapping) involves a high degree of direct physical correspondence, that is, a *direct causal connection*. The second anomaly is that even this simple transposition is not maintained consistently. In auditory icons the signal-referent relation may be *iconic* (causal) or *noniconic* (*metaphorical* for Gaver, *ecological* and *random* for Keller and Stevens), so it is not the case that *all* auditory icons are, by definition, iconic. With respect to the *indexical* nature of auditory icons, semiotics can offer considerable further clarification. Semiotics distinguishes between *natural signs* and *conventional signs*.

Natural environmental sounds (whether biological or geophysical) are *natural indicators*. What Gaver (1993) refers to as "everyday listening" (pp. 1) is based on the hearer's learned expectation of informative connections between sound and its environment.

A naturally occurring sound is always *indicative of* some object, situation or event, which has typically been part of the antecedent causal conditions of the occurrence of that sound. As such, it may also be indicative of the *properties* of its immediate source, as in the case of discrete impacts, vibrating solids, aerodynamic events, liquid sounds, and so forth (Ballas, 1993; Gaver, 1993; Warren & Verbrugge, 1984), or indicative of the *nature* of the antecedent event, as in bouncing versus breaking objects (McAdams, 1993).

Here, a distinction can be made between *proximal cause* and *distal cause*. Crucially, what is indicated is not only, and not necessarily, the proximal cause, the sound-producing object or event; it may also be the distal cause, that is, whatever has led to the production of the sound. For example, a bird's warning cry may indicate the bird (as its direct or proximal causal antecedent), but it may also, more importantly, indicate the presence of a predator (as its indirect or distal causal antecedent).

For the hearer, the direct causal antecedent is not necessarily the more salient; sometimes, a warning cry may reliably indicate a predator even when the producer of the sound cannot be identified, and/or is irrelevant. At other times, the sound may indicate, in addition to its producer, the presence of something else that is temporally or spatially contiguous but causally much more remote. For example, the sound of a seagull may indicate the closeness of the sea, or the sound of an owl hooting may indicate that it is night-time.

Against this background clarified by semiotics, it becomes apparent that auditory icons are not icons at all, but natural indicators, exploited for purposes of conventional indication. In artificial operating system environments, all auditory warnings (unless they are sounds from accidental breakages or other unforseen events) are *designed* to serve as indicators, and, *ipso facto*, are *conventional indicators*. However, because *any* sound can be used, conventional indicators may also be natural indicators.

Thus, a preliminary descriptive classification of auditory warnings in artificial environments would involve the following three groups. The first group consists of any of the class of natural indicators that have been *adopted* to indicate (in order of the directness of the physical relation) (a) proximal cause, (b) distal cause, or (c) merely correlated object/event. This first group includes those indicators that are made by humanly manufactured objects (e.g., the sound of a car), and that may be legitimately considered part of the environment of natural indicators (i.e., the environment into which humans are born).

The second group consists of any of the class of 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 conventionally selected target. For example, in a warning based on a *metaphorical* signal-referent relation, the use of a mosquito sound might be used to indicate the target helicopter.

The third group consists of non-natural indicators, including abstract sounds (e.g., pure tones, random tone sequences, or earcons), any naturally occurring objects or events randomly connected to their referents, and speech. By speech, we mean spoken words arbitrarily connected to their meanings. However, some aspects of speech (such as expressive cries, screams, intonation) will belong to the class of natural indicators. Furthermore, some aspects of these nonnatural indicators will be arbitrary (as in the connection between a single word and its meaning), but others will be nonrandom (as in the systematic structures of earcons or the grammar of language).

Despite this variation in the third group, the term *auditory icon* can be clarified; it is used to denote the *set of conventional indicators which either adopt or adapt members of the set of natural indicators*. This clarification of the meaning of auditory icon allows us to go some way toward explaining the perceived discrepancies between visual and auditory domains.

Problems Arising from Visual-to-Auditory Transfer

If auditory icons are not really icons but indicators, that may partly explain why the attempt to transfer classifications from the visual to the auditory mode leads to difficulties, and why the auditory field lags behind the visual.

Visual icons are genuine icons, but auditory icons are not. Iconicity in the visual domain can be largely accommodated by focusing exclusively on the *dyadic* relation between signal and referent. This is because visual iconicity pertains to shared features between signal and referent. Perception of these shared features typically requires no more than a basic perceptual ability on the part of the person or user. Therefore, classification of visual icons can proceed in a relatively straightforward way. Degree or type of shared features can be systematically varied, and predictions for ease of learning and recall can be derived from them.

There is, of course, also pure iconicity within the auditory domain. Here, too, similarity is involved—specifically, shared acoustic features between two sounds. However, as we have argued, this is not what is generally meant by "iconic" in the auditory warning literature. Because auditory icons are based on *indication*, their perception draws much more upon experience and learning. It is not simply a matter of a person's ability to see or hear that two things share visual or acoustic features; rather, it is a matter of a person's recalling from experience that two things (sound and object/event) are associated causally.

Clearly, when the person or user enters the equation in such a way that the user's cognitive contribution (viz., prior learning) becomes a relevant variable, the situation immediately becomes more complex. Similar difficulties have been found in the area of sonification, where auditory display designers need to accommodate variations in the parameter of polarity. In the visual domain, polarity is usually positive (e.g., people judge an increase in the visual display to be appropriate for indicating an increase in the target). However, in the auditory domain, some people favor positive polarity (e.g., increase in pitch is appropriate for indicating increase in size) but others favor negative polarity (e.g., increase in pitch is appropriate for representing *decrease* in size). Also in sonification, *perceptual* magnitude estimations produce much more straightforward results than do *conceptual* magnitude estimations (e.g., Walker, 2002, 2007).

Naturally, the increased complexity introduced by the user's cognitive contribution will also apply to the visual domain, but only when attention shifts from shared perceptual features to more complex characteristics such as *meaningfulness* and *semantic distance* (McDougall et al., 2000).

Ambiguity in the Concept of Signal-Referent Relation Strength: The Role of Learning

When the cognitive contribution of the user becomes important, the field of semiotics is no longer sufficient. Semiotics now needs to be supplemented by principles of cognitive psychology. Psycho-semiotics has emerged as an attempt to combine psychology and semiotics and address the perceived neglect in semiotics of the person or subject (cf. Bouissac, 1998).

Psycho-semiotics brings to center stage the concept of signification as involving the full *triadic* relation between signal, referent, *and person*, rather than just the *dyadic* relation between signal and referent. Signification is, logically, a *three-term relation*, in which the subject term, the cogniser or user, is a crucial element (Petocz, 2001).

Of course, it is quite legitimate to describe the various types of relation that obtain between two of the terms, as in the *bases* for the semiotic distinctions between icon, index, and symbol. However, because something can only indicate something else *to someone* (or any cognizing organism), then only when the third element, the person or user, is brought into the picture is meaning or indication fully characterized. Without that, it is impossible to know whether a particular sign is to serve as an icon, index or symbol. The same sound (e.g., a mouse's squeak) may be an icon (for a baby's cry), an index (for the mouse), or a symbol (for an earthquake).

In the auditory warning literature, there is frequent reference to learning. With respect to the standard method of developing and testing classifications, learning is the dependent variable. The aim is to test which types of signal-referent relation are more easily learned and recognized than others. Although Gaver predicted that nomic mappings would be the easiest to learn and symbolic mappings the most difficult, he noted that "articulatory directness does not necessarily affect performance once a mapping has been well learned" (Gaver, 1986, p. 172).

Here, Gaver was noting the contribution of prior learning to the *independent* variable. The fact that symbolic mappings based on arbitrary connections, once learned, become *no different functionally* from those in natural indication (which themselves must also be learned) highlights a major problem that has significant conceptual and methodological implications. This problem is the ambiguous treatment in the literature of the key notion of *signal-referent relation strength*, a concept that not only lies at the center of existing classifications, but also forms the basis on which stimulus sets are designed and predictions are tested.

In brief, classification and research confounds and switches between the *dyadic* relation (articulatory directness between signal and referent) with the *triadic* relation (how well learned by the person is that signal-referent association).

On the one hand, the notion of signal-referent relation strength is understood, as the term implies, to refer to the strength of the *dyadic* relationship, or association, between a given sound signal and the critical information to which it refers (Gaver, 1986, 1989). Here, the third element, the person (including that person's prior learning) is omitted from the picture. Thus, the strongest relations are the direct, nomic, or iconic ones, where signal and referent are causally connected, the weakest relations are the abstract or symbolic ones, where there is no natural or physical connection between sound signal and its referent, and of intermediate strength are the indirect (metaphorical or ecological) relations (Keller & Stevens, 2004). Hence, asking the question *how strong* is the signal-referent relation is equivalent to asking how strong is the dyadic relation regardless of whether it is perceived, or how often it has been perceived, by anyone.

In this case, because the term is defined *independently of learning*, it might appear reasonable to offer predictions regarding ease of learning and remembering, based on the assumption that stronger relations are more easily learned and remembered. Much of the research has been based on such predictions and assumptions. However, this approach is flawed for two reasons.

First, it would be practically impossible to adhere to the definition based on the dyadic relation, and then test the predictions, without having the results confounded by *prior* learning. The test would need conditions in which the participants could be guaranteed not to have had the relevant prior associative learning with respect to the stimulus set (only neonates or Martians would qualify). If, for example, it is predicted (on grounds of causal directness) that the use of the sound of an elephant trumpeting to refer to the target "elephant" would involve a stronger signal-referent relation than would the use of an owl call to refer to the target "night," success in prediction for learning could only be *because* the causal directness of the former is *already learned*.

There is a second reason why testing predictions derived from the dyadic conception of signal-referent relation strength is flawed. The basic assumption underlying any such test, that is, the assumption that stronger relations (dyadically conceived) are more easily learned and remembered, is obviously false. If we do not know what sound an elephant makes, the directness of the causal relation per se does not help. Gaver (1986) does point out that articulatory directness is irrelevant once a mapping has been learned, but he stops short of stating that articulatory directness is of no use if the relevant causal connection has *not* yet been learned. Although we learn various causal connections before we learn the arbitrary connections between linguistic symbols and their referents, there is no obvious reason to believe that a child's learning of a new causal connection proceeds any more quickly than the learning of a new word, once that child has grasped the general concept of linguistic reference.

Therefore, of the two ambiguous treatments of signal-referent relation strength, this first is of little use. Predictions based on the relation defined dyadically are inevitably confounded by the prior learning of the person or user, and are in any case based on an assumption known to be false.

On the other hand, the concept of signal-referent relation strength is also used to refer to *how well learned* the connection between signal and referent is. This definition involves the full *triadic* relation between signal, referent and person. The best-learned connections, however arbitrary (as in speech), are acknowledged to comprise the *strongest* signal-referent relations. Indeed, the associations in speech are typically so well learned that the symbolic system as a whole becomes generally "transparent" and our engagement with it automatic. This explains why research that has compared speech warnings with other types of sound has generally found that the speech warnings are the most easily learned and recalled.

Clearly, this second (triadically conceived) definition of signal-referent relation strength entails a *different*, but much less fixed classification of sounds. Speech (typically) would be located at the strongest end, and unlearned abstract sounds or random connections would be located at the weakest end. However, the relative strength of the other categories would be variable, for it all depends on *how well learned* they are, and on the context of their occurrence.

The price of correctly defining signal-referent relation strength as a triadic signifying relation by including the cognitive contribution of the person is inevitable variation and uncertainty in classification. Ironically, however, this does not matter. For it is clear that testable predictions regarding ease of learning are vacuous, because they are already contained in the classification, such that the dependent and independent variables cannot be separated.

Thus, of the two ambiguous treatments of signal-referent relation strength, this second is also of little use as the basis of a classification system from which testable predictions about ease of learning are to be derived.

Figure 1 illustrates this confusion surrounding the concept of signal-referent relation strength. Figure 1a and 1b presents what might appear to be two separate, but overlapping, classification systems. Each of the systems contains the various types of conventional auditory indicator which we identified earlier in the discussion (encompassing the various signal-referent relations in existing classifications), and each represents auditory icons as adopted or adapted natural indicators. The only difference is the ordering of the items along the strong/weak dimension.

However, the aim of the figure is not to offer yet another classification for research purposes. The aim is to show what happens to classification when we identify and separate the two

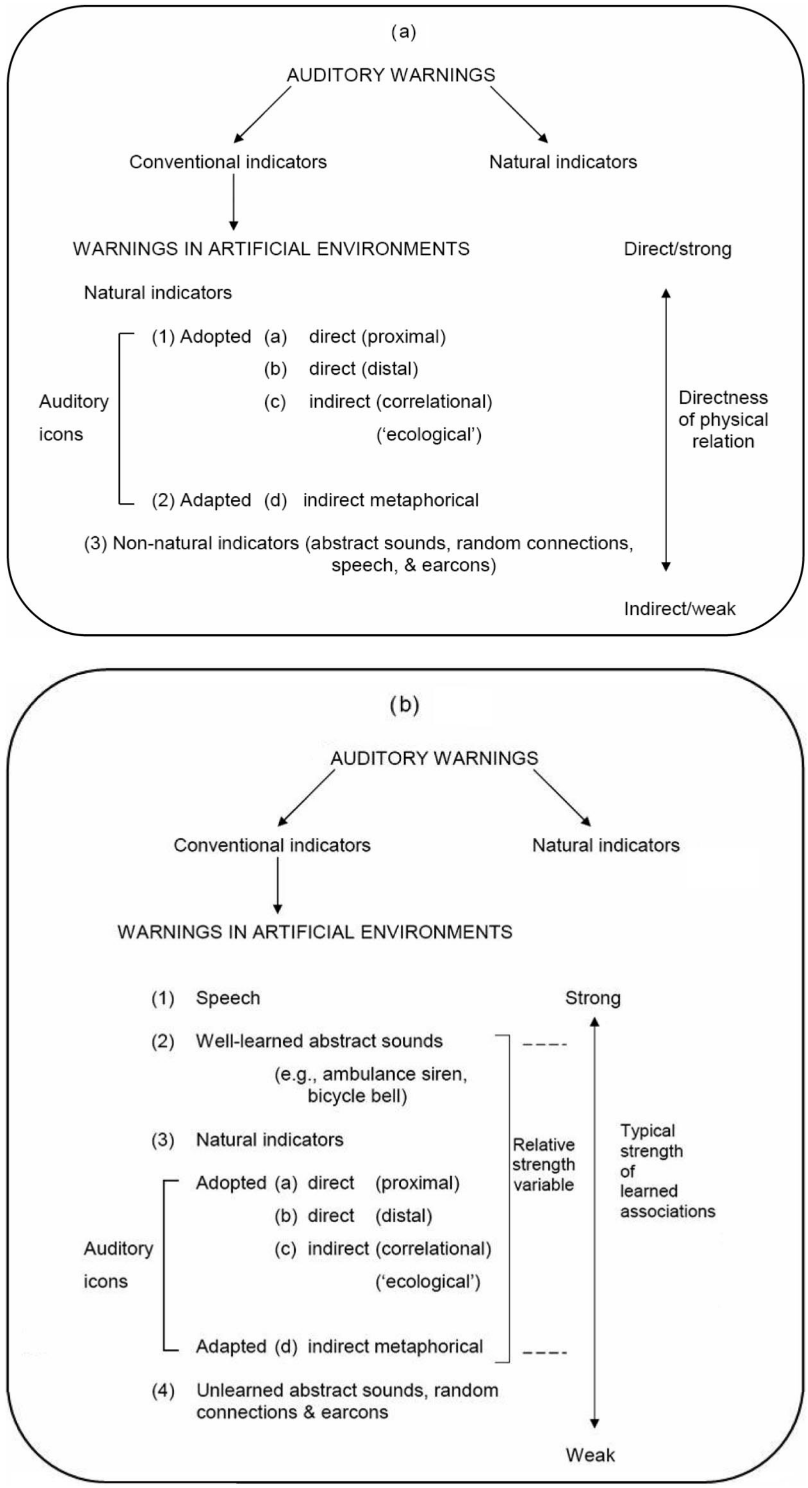

Figure 1. Illustration of the limitations of the standard method of proposing and testing auditory warning classification systems. Separation into two systems (1a and 1b) disentangles the ambiguous definition of signal-referent relation strength. Both systems (1a and 1b) include all types of signal-referent relation (including so-called auditory icons), and differ only in the ordering of items along the strong/weak dimension. In 1a signal-referent relation strength is defined only dyadically as the relation between signal and referent, independent of the person. In this system, research predictions regarding ease of learning will not be fulfilled because results will be confounded by the prior learning of the (neglected) person. In 1b signal-referent relation strength is defined correctly as the triadic relation between signal, referent, and person (i.e., as strength of learned association). In this system, research predictions regarding ease of learning are vacuous, because they are already contained in the classification.

different definitions of signal-referent relation strength that are conflated in the literature. In this illustration, we disentangle the classification of auditory warning sounds based on the dyadic conception of signal-referent relation strength (1a) from the classification based on the triadic relation (1b).

When the dyadic relation is used (1a), the strongest relations are those involving direct causality, and the weakest are those involving random or arbitrary connections, as in speech. However, predictions about ease of learning based on this ordering will not be fulfilled, because results will be confounded by learning (e.g., speech is already well-learned and so is likely to out-perform natural indicators).

When, on the other hand, the triadic relation is recognized (1b), the strongest relations are those (such as in speech) that are the most well-learned, and the weakest are those which have not been learned. There will be no fixed ordering for most items, because it will depend on how well-learned they are. However, predictions about ease of learning now become trivial.

Thus, Figure 1 illustrates why the standard method of proposing and testing auditory warning classification systems is seriously limited. We know beforehand that predictions derived from the first system will not be fulfilled, for results will be inevitably confounded by learning. We also know beforehand that predictions derived from the second system are already answered in the classification itself, rendering pointless their testing. Conflating the two systems serves only to disguise these problems.

Extensions of the Neglect of Learning and of its Context

We have argued that signification is a three-term relation in which the person plays a crucial role. Once the person is put back into the relation, to form the signal-referent-person triad, it becomes clear that the knowledge (prior learning) of that person forms a part of the relation that cannot be ignored.

Yet the neglect of learning is a theme that extends beyond the problem of classification and runs throughout the research in the auditory warning area. Insufficient weight has been given to what is already known concerning just how much prior learning is likely to override other factors, to inhibit or interfere with new learning, and even to intrude into the domain of abstract sounds. In addition, the sensitivity of learning to its context has been neglected. As a result, the usual efforts to maintain ecological validity in the research have been compromised.

Interference by prior learning can explain a number of experimental findings that appear to contradict the fact that new associations *can* be learned. For example, in the Keller and Stevens (2004) study, random signal-referent relations were never adequately learned, in the sense that recognition times remained relatively long. McKeown and Isherwood (2007) obtained similar results for the same type of signal-referent relation (e.g., footsteps to indicate low petrol level, baby sneeze to indicate speed limit being exceeded).

Crucially, however, when using sounds which already serve as natural indicators of their causal antecedents, the task becomes not one of learning random connections, but, rather, one of *un*learning *preexisting* connections. The difficulty of such a task is an indication of the strength of the already learned associations. In this respect, the case of stimulus-response compatibility may be relevant to the issue of bringing long-established everyday associations to a novel eccentric context (cf. Kornblum, Hasbroucq, & Osman, 1990). Dutta and Proctor (1992) observed that incompatible mappings between stimuli and responses (on spatial and symbolic dimensions, in separate experiments) interfered with performance on choice reaction time tasks even after 2,400 trials spread over eight experimental sessions. The robustness of prior learned associations can also explain why Stanton and Edworthy (1998) were able to demonstrate that new auditory icons would not fare as well for ITU (intensive treatment unit) staff who *had already learned the conventional warnings*. Likewise, the "inappropriate" or "risky" responses elicited by the auditory icons in the Graham (1999) study represent instances of interference by prior learning.

With respect to abstract sounds, classical conditioning can render our responses just as strong as they are for natural indicators. Language is the best example of this kind of strength of initially arbitrary signal-referent relation. Indeed, learning via classical conditioning is often strong enough to override prior natural responses to stimuli. Thus, for example, a sound that resembles that of a bicycle bell may well have *acoustic* features expected to be predictive of high-perceived urgency (Edworthy, Loxley, & Dennis, 1991; Hellier & Edworthy, 1999; Hellier, Edworthy, & Dennis, 1993). However, knowledge of learning would moderate that prediction, such that we should not be surprised to find that the sound is associated with a *low* level of threat, and even with pleasure (Guillaume et al., 2003).

Furthermore, neglect of learning has precluded critical scrutiny of the very notion of the abstractness of a sound, such as a siren to indicate an approaching ambulance, or applause to indicate approval. Yet there is a substantial body of literature relevant to this issue. It deals with the way our reactions to characteristics of sounds and symbols have been shaped by the combination of evolution, aspects of our physiology, and the natural environment.

Peirce (1932) commented on the probable original iconic and indexical aspects of language, and Sapir (1959) followed him. More recently, theoretical and empirical work on conceptual metaphor and on mappings across sounds and visual images presents evidence for the cross-cultural salience and consistent patterns of metaphorical thinking (Gallagher, 2005; Gibbs, 1994; Hopkins, 2000; Kövecses, 2005; Lakoff, 1993; Lakoff & Johnson, 1980; Lakoff & Turner, 1989; Ramachandran & Hubbard, 2001; Wheeler, 2005).

According to Edworthy and Hellier (2006a), "Warnings as metaphors is the main growth area in the brave new world of sound chips, but there are a great many issues to be resolved before clear guidance can be given about which types of metaphors might be best for which types of application" (p. 201).

The research in conceptual metaphor offers some promising and intriguing material concerning learning predispositions, associations, and the ability of users to draw upon these associations when using their own cues to attach meanings to abstract sounds. For example, it addresses the question why some sound-concept mappings (e.g., increasing pitch to represent increasing temperature) are consistently judged to be appropriate, whereas other mappings (e.g., decreasing pitch to represent increasing size) elicit more variation (Walker, 2002, 2007; cf. Gibbs, 1994).

There are important implications from this body of literature for the design of auditory warnings and for the predictability of specific findings in the research, such as those reported in Dalton and Lavie (2004), Edworthy and Hards (1999), Guillaume et al. (2003), Keller and Stevens (2004), Lemmens, Bussemakers, and de Haan (2000), McAdams (1993), and Walker (2002, 2007). As yet, however, the material on conceptual metaphor remains a relatively untapped resource.

Finally, in the auditory warning research on signal-referent relations, the sensitivity of learning to its context has been neglected. For example, unrealistically heterogeneous stimulus sets are likely to produce results that cannot be used reliably to inform the design of the relevant warning sets. In addition, the move out of the traditional experimental laboratory into more contextually relevant environments (e.g., aircraft cockpit or automobile simulators) or into the real-life environment (e.g., hospitals), brings its own problems. When these environments are distorted by the use of natural indicators [e.g., the skidding tires and car horn auditory icons in the Graham (1999) and the Belz et al. (1999) studies] to signal a restricted subset of their usual referents, or by the use of contrived speech signals (e.g., "ahead," instead of the usual "look out!"), then any gains made by the move toward greater ecological validity are likely to be forfeited. This is because the artificial environment is always embedded, both physically and psychologically, in the natural environment that includes learned associations. In the Graham (1999) study, what was experimentally defined as an "inappropriate response" (i.e., a brake press to the sound of car horn) was, according to the natural background setting of the participant's learned associations, an *appropriate* response. An experimentally appropriate response should not be operationally defined in such a way as to be inconsistent with what is naturally appropriate, particularly in an investigation of the efficacy of signals (auditory icons) which are, *by definition*, designed to harness natural responses. Ecological validity is compromised wherever the experimental set-up contravenes the normal environmental stimulus-response connections. Further, the more ecologically valid the research setting, the more likely it is to prime naturally learned associations.

Of course, while some associations are obvious and predictable, others are *not* and need to be collected. Belz et al. (1999) report on making use of their earlier attempts to "formalize the auditory icon design process" (p. 609) by requiring participants to "assign a meaning to the auditory icon, rate its perceived urgency, and rate its level of association with its intended (experimenter-selected) meaning" (p. 609). Keller & Stevens (2004) collected individual associations and ratings of association strength, and then used this information in their selection of stimuli for different categories of signal-referent relation. McDougall et al., (1999) collected judgments concerning a number of characteristics for 239 visual symbols, and recommended that "subjective ratings be applied in both the early and later stages of the design process" (p. 304). Despite these studies, the collecting of individual associations as a preliminary step does not appear to be part of accepted standard procedure.

Recommendations and Summary

Locating our Contribution Within the Standard Procedure for Auditory Warning Design

If classification-driven research of the type we have discussed is not useful for moving forward in the field of auditory warnings, it remains for us to offer some positive directions by locating the contributions of this paper within the standard procedure for auditory warning design.

In Figure 2, we present a generic flow diagram outlining the steps in the usual procedure for designing auditory warnings. Although these steps have not been set out explicitly in the literature, we believe that most researchers and designers would recognize the general framework as one that already implicitly informs their work. A useful analogy (and possibly helpful mnemonic) for this procedure might be the stages of film-making, involving preproduction (Steps 1–4), production (Steps 5–9) and post-production (Step 10). Within the production phase, there is also a specific process of "casting" potential signals (Steps 5–7).

We use this framework to highlight the particular steps relevant to our discussion and recommendations. These are the preproduction Step 1 and the production Steps 5 through 9.

At Step 1, our discussion has focused on the problems that may occur when new warnings are added to existing systems in an ad hoc fashion. Specifically, it is important not to adopt or adapt (as auditory icons) natural indicators that already exist in the relevant environment (such as the sound of screeching tires), if the use of these for conventional warnings will require restricting the conventional referents to only a subset of the natural referents.

At Steps 5 through 8, we have discussed a number of issues that might better inform the standard procedures. For example, the initial selection or design at Step 5 must be constrained by the external warning signals identified at Step 1. The iterative "casting" process involving Steps 5 through 7 should also largely be driven by knowledge of, and further investigations into, the cognitive system of the user. For example, at Step 7, individual associations both to the selected sounds and to the targets, and ratings of association strength, should always, where possible, be collected from a sample of the proposed user population. In addition, where possible, efforts to obtain *implicit* associations should be made. Here, the process must be sensitive to the difference between frequency of occurrence of certain events in a particular environment and frequency of exposure of an individual to those events. At Step 8, the final proposed warning signals should always be scrutinized carefully in the context of the stimulus set, the operating system environment, the relevant wider environment, and the individual learned associations (as collected).

At Step 9, the warning signals should, where possible, be tested for efficacy in an environment that is as close as possible to the proposed operating system environment, even if this requires an extended time-span to accommodate (where appropriate) the relatively rare occurrence of the warning sound.

Although the literature does reveal that some of these guidelines and steps have already been observed, they have clearly not been widely adopted. Rather, the evidence suggests that they have been unnecessarily (often inadvertently) contravened.

Summary

To summarize, we have focused on the psychological side—especially cognition and learning—and attempted to identify and clarify a number of problems which have emerged from a critical review of theory and research in the area of auditory warnings and signal-referent relations.

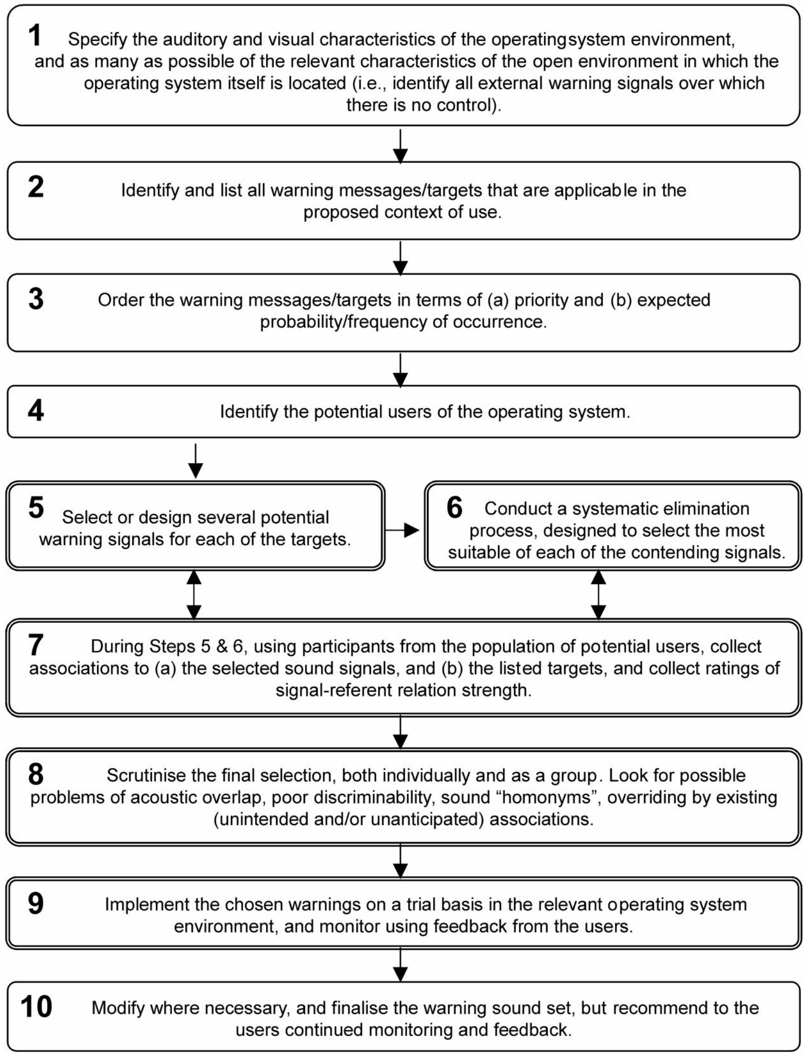

Figure 2. Flow diagram providing the context for locating the contributions of the paper within the standard procedure for designing auditory warning systems.

We have argued that there are conceptual confusions and inconsistency in terminology and definition, including those involved in visual-to-auditory transfer. Here, insights from the field of semiotics can help to clarify how the auditory warning literature has confused the concepts of *icon* and *index*. So-called auditory icons are not icons but conventional indicators that either adopt or adapt natural indicators.

This clarification of the meaning of auditory icon helps toward explaining some of the perceived discrepancies between visual and auditory domains. Visual icons are genuine icons, but auditory icons are not. Iconicity in the visual domain can be largely accommodated by focusing exclusively on the dyadic relation between signal and referent. Perception of shared features typically requires no more than a basic perceptual ability on the part of the person or user, such that the cognitive contribution of that user (in terms of prior learning) is likely to be negligible. Pure iconicity in the auditory domain would be comparable, involving similarity between two sounds. However, since auditory icons are based on indication, their perception draws much more upon experience and learning.

When the learning of the user enters the picture, insights from semiotics need to be combined with insights from psychology, as in the field of psycho-semiotics. Here, the theory of signification emphasizes the need to consider not just the dyadic relation between signal and referent, but the full triadic relation between signal, referent, and person. The importance of the triadic relation converges with the importance of the role of learning.

In the auditory warning literature, classification-driven research regarding signal-referent relation strength conflates the dyadic relation (articulatory directness between signal and referent) and the triadic relation (how well learned by the person is the signal-referent association). Disentangling these two (see Figure 1) shows why the method of proposing and testing classification systems is not useful. Predictions are either confounded by learning or vacuous.

In general, we have concurred with Edworthy and Hellier (2006a) in diagnosing confusions of terminology and failure to take account of the cognitive system of the user. Edworthy and Hellier identified the lack of an agreed taxonomy for sound as one of the major problems, and one of the reasons why the whole area of auditory cognition lags considerably behind that of visual cognition. They argued that this lag represents a void that researchers should be encouraged to fill.

We have moved further, into the *combination* of semiotics and psychology, and have highlighted the fundamental limitation of the standard method of proposing classification systems from which to derive testable predictions about relative ease of learning.

We are not suggesting that the descriptive task of classification should be abandoned; it is a legitimate and important part of theoretical work. We are suggesting that we should rethink how we *use* those classifications in the light of our empirical research aims. The practical implications of the issues we have discussed are that we may need to adjust the common starting point of our research, and redirect attention onto psychological factors and developing methods for accessing and accommodating earlier in the design process the cognitive contribution of the user.

It may fairly be objected that we *cannot* know in advance *all* of the learned associations of the end user. Even if we could, it would not help us, given that each end user is likely to have a different set of associations. However, the evidence from the research we have discussed suggests that we *can* know *some* of them, often those that are the most relevant, and these are also often shared across users. If even these had been taken into account, many of the problems in the research literature would not have arisen. In addition, there are developments in other areas of psychology and cognitive science (e.g., conceptual metaphor, as discussed above) which hold promise for expanding our knowledge base. Indeed, recent work in sonification has been drawing upon cognitive psychology in just the ways we are recommending be adopted more widely (cf. Walker, 2007). Therefore, the facts of learning and individual differences, while admittedly a challenge, need not be grounds for despair.

Finally, we have offered some guidelines and located them within the context of accepted procedures for auditory warning design. Naturally, some of the guidelines will appear to be unrealistic counsels of perfection, and precluded in many circumstances by practical constraints. Nor would we deny that the psychological problems are, in many respects, considerably more challenging than are the acoustic or psychoacoustic problems. Nevertheless, we believe that only by having a clear vision of what would be required under ideal conditions can we gradually work toward improvements. The greatest promise for progress in this field lies in recognizing how psychology contributes to the limitations of our taxonomies, and devising effective ways of further accessing and accommodating much earlier in the design process the cognitive contribution of the user.

References

Alston, W. P. (1967). Sign and symbol. In *The encyclopedia of philosophy* (Vol. 7, pp. 437–441). New York: Macmillan.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 250–267.

Barker, P., & van Schaik, P. (2000). Designing and evaluating icons. In M. Yazdani & P. Barker (Eds.), *Iconic communication* (pp. 161–177). Bristol, United Kingdom: Intellect.

Begault, D. R., & Pittman, M. T. (1996). Three-dimensional audio versus head-down traffic alert and collision avoidance system displays. *International Journal of Aviation Psychology, 6,* 79–93.

Begault, D. R., & Wenzel, E. M. (1992). Techniques and applications for binaural sound manipulation in human-machine interfaces. *International Journal of Aviation Psychology, 2,* 1–22.

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.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human-Computer Interaction, 4,* 11–44.

Bouissac, P. (1998). Converging parallels. Semiotics and psychology in evolutionary perspective. *Theory & Psychology, 6,* 731–753.

Burt, J. L., Bartolome, D. S., Burdette, D. W., & Comstock, J. R. (1995). A psychophysiological evaluation of the perceived urgency of auditory warning signals. *Ergonomics, 38,* 2327–2340.

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

Cooper, G. E. (1977). A survey of the status and philosophies relating to cockpit warning systems. *NASA Ames Contract Report No. NAS2–9117.*

Dalton, P., & Lavie, N. (2004). Auditory capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 180–193.

Dingus, T. A., McGehee, D. V., Manakkal, N., & Jahns, S. K., Carney, C., & Hankey, J. M. (1997). Human factors field evaluation of automotive headway maintenance/collision warning devices. *Human Factors, 39,* 216–229.

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

DuRoss, S. H. (1978). Civil Aircraft Warning Systems: A survey of pilot opinion within British Airways. *Royal Aircraft Establishment, Tech. Rep. No. 78056.*

Dutta, A., & Proctor, R. W. (1992). Persistence of stimulus-response compatibility effects with extended practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 801–809.

Eco, U. (1973, October 5). Looking for a logic of culture. In The Tell-tale sign: A survey of semiotics I. *Times Literary Suppl.,* 1149–50.

Edwards, A. D. N. (1995). *Extra-ordinary human-computer interaction: Interfaces for users with disabilities.* New York: Cambridge University Press.

Edworthy, J. (1994). The design and implementation of non-verbal auditory warnings. *Applied Ergonomics, 25,* 202–210.

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

Edworthy, J., & Hards, R. (1999). Learning auditory warnings: The effects of sound type, verbal labelling and imagery on the identification of alarm sounds. *International Journal of Industrial Ergonomics, 24,* 603–618.

Edworthy, J., & Hellier, E. (2006a). Complex nonverbal auditory signals and speech warnings. In M. S. Wogalter (Ed.), *Handbook of warnings* (pp. 199–220). Mahwah, NJ: Erlbaum.

Edworthy, J., & Hellier, E. (2006b). Alarms and human behaviour: Implications for medical alarms. *British Journal of Anaesthesia, 97,* 12–17.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors, 33,* 205–231.

Edworthy, J., & Stanton, N. (1995). A user-centered approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics, 11,* 2262–2280.

Familant, M. E., & Detweiler, M. C. (1993). Iconic reference: Evolving perspectives and an organizing framework. *International Journal of Man-Machine Studies, 39,* 705–728.

Gallagher, S. (2005). *How the body shapes the mind.* Oxford: Clarendon.

Galton, F. (1899). On instruments for (1) testing perception of differences of tint and for (2) determining reaction time. *Journal of the Anthropological Institute, 19,* 27–29.

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction, 2,* 167–177.

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.

Gibbs, R. W. Jr. (1994). *The poetics of mind. Figurative thought, language and understanding.* Cambridge: Cambridge University Press.

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

Guillaume, A., Pellieux, L., Chastres, V., & Drake, C. (2003). Judging the urgency of nonvocal auditory warning signals: Perceptual and cognitive processes. *Journal of Experimental Psychology: Applied, 9,* 196–212.

Haas, E. C., & Casali, J. G. (1995). Perceived urgency of and response time to multitone and frequency-modulated warning signals in broadband noise. *Ergonomics, 38,* 2313–2326.

Haas, M. W. (1995). Virtually-augmented interfaces for tactical aircraft. *Biological Psychology, 40,* 229–238.

Hawkes, T. (1977). *Structuralism and semiotics.* London: Methuen & Co.

Hellier, E., & Edworthy, J. (1999). On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics, 30,* 167–170.

Hellier, E., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors, 35,* 693–706.

Hopkins, J. (2000). Psychoanalysis, metaphor and the concept of mind. In M. P. Levine (Ed.), *The analytic Freud. Philosophy and psychoanalysis* (pp. 11–35). London: Routledge.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1986). Direct manipulation interfaces. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives on human computer interaction* (pp. 87–124). Hillsdale, NJ: Erlbaum, Inc.

Jacko, J. A., & Rosenthal, D. J. (1997). Psychology of computer use: XLVI. Age-related differences in the mapping of auditory icons to visual icons in computer interfaces for children. *Perceptual & Motor Skills, 84,* 1223–1233.

Keller, P., & Stevens, C. (1999). *A Taxonomy for Auditory Warnings (TAW): Learning associations between auditory icons and their referents.* Unpublished manuscript, MARCS Auditory Laboratories, University of Western Sydney.

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.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility: A model and taxonomy. *Psychological Review, 97,* 253–270.

Kövecses, Z. (2005). *Metaphor in culture: Universality and variation.* Cambridge, UK: Cambridge University Press.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.) *Metaphor and thought* (2nd ed.) (pp. 202–251). Cambridge: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor.* Chicago: University of Chicago Press.

Langer, S. K. (1942). *Philosophy in a new key.* Cambridge, MA: Harvard University Press.

Lemmens, P. M., Bussemakers, M. P., & de Haan, A. (2000). *The effects of earcons on reaction times and error-rates in a dual-task vs. single-task experiment*. Retrieved from http://www.icad.org/websiteV2.o/Conferences/ICAD2000/PDFs/BussemakersDualTask.pdf

Leung, Y. Y., Smith, S., Parker, S., & Martin, R. (1997). *Learning and retention of auditory warnings.* Unpublished manuscript, Defense Science & Technology Organisation, Australia.

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

Maltz, M., & Shinar, D. (2004). Imperfect in-vehicle collision avoidance warning systems can aid drivers. *Human Factors, 46,* 357–366.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in sound. The cognitive psychology of human audition* (pp. 146–198). Oxford: Clarendon Press.

McDougall, S. J. P., Curry, M. B., & de Bruijn, O. (1999). Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. *Behavior Research Methods, Instruments & Computers, 31,* 487–519.

McDougall, S. J. P., Curry, M. B., & de Bruijn, O. (2001). The effects of visual information on users' mental models: An evaluation of pathfinder analysis as a measure of icon usability. *International Journal of Cognitive Ergonomics, 5,* 59–84.

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.

McKeown, D. (2005, July). Candidates for within-vehicle auditory displays. *Proceedings of ICAD 05* – Eleventh Meeting of the International Conference on Auditory Display, Limerick, Ireland, July 6–9.

McKeown, D., & Isherwood, S. (2007). Mapping candidate within-vehicle auditory displays to their referents. *Human Factors, 49,* 417–428.

McKinley, R. L., & Ericson, M. A. (1997). Flight demonstration of a 3-D auditory display. In Gilkey, R. H., & Anderson, T. R. (Eds.). *Binaural and spatial hearing in real and virtual environments* (pp. 683–699). Mahwah, NJ: Erlbaum.

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.

Morris, C. (1946). *Signs, language, and behavior.* New York: Prentice Hall.

Naish, P. L. N. (1990). Simulated directionality in airborne auditory warnings and messages. In Life, A., Narborough-Hall, C. S., & Hamilton, W. I. (Eds.), *Simulation and the user interface* (pp. 127–141). London: Taylor & Francis.

Noyes, J. M., Starr, A. F., & Kazem, M. L. N. (2004). Warning system design in civil aircraft. In D. Harris (Ed.), *Human factors for civil flight deck design* (pp. 141–155). Hampshire: Ashgate.

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

Patterson, R. D. (1989). Guidelines for the design of auditory warning sounds. In *Proceedings of the Institute of Acoustics,* Spring Conference, Vol. II(5), Edinburgh: Institute of Acoustics, 17–24.

Peirce, C. S. (1932). In C. Hartshorne & P. Weiss (Eds.) (1960). *Collected papers of Charles Sanders Peirce* (Vol. II). Cambridge, MA: Harvard University Press.

Petocz, A. (2001). Psychology in the twenty-first century: Closing the gap between science and the symbol. In J. R. Morss, N. Stephenson & H. van Rappard (Eds.), *Theoretical issues in psychology* (pp. 367–377). *Proceedings of the International Society for Theoretical Psychology 1999 Conference.* Boston: Kluer Academic Publishers.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies, 8,* 3–34.

Sanders, A. F. (1998). *Elements of human performance: Reaction processes and attention in human skill.* Mahwah, NJ: Erlbaum.

Sapir, E. (1959). Symbolism. In E. R. A. Seligman (Ed.), *Encyclopaedia of the social sciences* (Vol. 13, pp. 493–494). New York: Macmillan.

Skorupski, J. (1976). *Symbol and theory.* Cambridge: Cambridge University Press.

Stanton, N. A., & Edworthy, J. (1998). Auditory affordances in the intensive treatment unit. *Applied Ergonomics, 29,* 389–394.

Stanton, N. A., & Edworthy, J. (1999). Auditory warning affordances. In N. A. Stanton & J. Edworthy (Eds.), *Human factors in auditory warning* (pp. 113–127). Aldershot, England: Ashgate Publishing.

Stephan, K. L., Smith, S. E., Martin, R. L., Parker, S. P. A., & McAnally, K. (2006). Learning and retention of associations between auditory icons and denotative referents: Implications for the design of auditory warnings. *Human Factors, 48,* 288–299.

Stokes, A. F., & Wickens, C. D. (1988). Aviation displays. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 387–431). San Diego, CA: Academic Press.

Szalma, J. L., Warm, J. S. Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., et al. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors, 46,* 219–233.

Ulfvengren, P. (2003). *Design of natural warning sounds in human-machine systems.* Unpublished doctoral thesis, Stockholm Institute of Technology, Sweden.

Walker, B. N. (2002). Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Applied Cognitive Psychology, 21,* 579–599.

Walker, B. N. (2007). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied, 8,* 211–221.

Warren, W. H., & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 704–712.

Wheeler, M. (2005). *Reconstructing the cognitive world. The next step.* Cambridge, MA: MIT Press.

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

Received May 29, 2007
Revision received February 13, 2008
Accepted February 25, 2008 ■