Journal of Experimental Psychology: Applied
2004, Vol. 10, No. 1, 3–12

1076-898X/04/$12.00 DOI: 10.1037/1076-898X.10.1.3

Meaning From Environmental Sounds: Types of Signal–Referent Relations and Their Effect on Recognizing Auditory Icons

Peter Keller and Catherine Stevens
University of Western Sydney

This article addresses the learnability of auditory icons, that is, environmental sounds that refer either directly or indirectly to meaningful events. Direct relations use the sound made by the target event whereas indirect relations substitute a surrogate for the target. Across 3 experiments, different indirect relations (ecological, in which target and surrogate coexist in the world; metaphorical, in which target and surrogate have similar appearance or function, and random) were compared with one another and with direct relations on measures including associative strength ratings, amount of exposure required for learning, and response times for recognizing icons. Findings suggest that performance is best with direct relations, worst with random relations, and that ecological and metaphorical relations involve distinct types of association but do not differ in learnability.

Auditory warnings are used in diverse types of human–machine interfaces, including personal computers (Gaver, 1986), hospital equipment (Momtahan, Hetu, & Tansley, 1993), and high performance aircraft (Doll & Folds, 1986). These operating systems differ in terms of the nature of the information being transmitted, the number of potential warnings, concurrent task demands, and environmental conditions. Such variety in operating systems necessitates the use of different types of sounds as auditory warnings. These traditionally include speech and abstract sounds (e.g., simple tones; Patterson, 1982) and, more recently, auditory icons (i.e., environmental sounds; Edworthy & Adams, 1996; Gaver, 1989; Graham, 1999; Leung, Smith, Parker, & Martin, 1997). In addition to operating system characteristics, the strength of the relationship or the association between a given sound signal and the critical information to which it refers (i.e., signal–referent relation strength) is germane to the business of selecting sounds for use as auditory warnings (Gaver, 1986, 1989).

The applied focus of research into auditory warnings has meant that choice of sound signal type has been influenced by operating system characteristics more so than by signal–referent relation strength, which has been treated as a more tacit consideration perhaps because it is difficult to quantify. Adequate theoretical background is lacking about the nature of the associations formed between various types of auditory signals and their referent events. Although it seems clear that stronger signal–referent relations can be established with speech sounds rather than with abstract sounds, the relative efficacy of auditory icons is unclear at present. Leung et al. (1997) found that accuracy in learning and retaining associations was similarly high with speech and auditory icons but much lower with abstract sounds. However, they also observed that some auditory icons were learned and retained more easily than others. They speculated that pairings of auditory icons and events vary systematically in terms of signal–referent relation strength. This variation highlights the need for an organizing framework that includes not only broad categories of auditory warnings, such as speech, auditory icons, and abstract sounds (e.g., Blattner, Sumikawa, & Greenberg, 1989) but also finer resolution of categories that reflect differences in auditory icon type. The aim of this study was to investigate part of a new classification scheme that distinguishes between different types of auditory icons (Keller & Stevens, 1999). Specifically, signal–referent relation type was manipulated in three experiments to measure the effects of type of association on learnability.

Taxonomies of Signal–Referent Relations

Existing taxonomies of auditory warnings are piecemeal—a consequence of their development for specific applications, notably computer interfaces. Gaver (1986) described a taxonomy that included only auditory icons, classifying these into three categories according to the degree to which a sound has direct physical correspondence to its referent event. In nomic mappings, meaning depends on the physics of the situation. For example, the sound of a metal object being struck to indicate the size of an object (Coward & Stevens, in press). Metaphorical mappings rely on

Peter Keller and Catherine Stevens, MARCS, University of Western Sydney—Bankstown, New South Wales, Australia.

Peter Keller is now at Max Planck Institute for Psychological Research, Munich, Germany.

This research was supported by Foundation and Industry Linkage Grants 03329/7001 and 03329/7503, respectively, from the University of Western Sydney in collaboration with Simulation and Human Factors, Air Operations Division, Defence Science and Technology Organisation, Department of Defence, Melbourne, Australia.

We thank Philip Barker for helpful comments; Simon Parker, Ken McAnally, Russell Martin, Sean Smith, and Karen Stephan for their assistance with the design of the experiments; and Cameron Cutts (data collection), Melinda Gallagher (data analysis), and David Green (programming).

Examples of the sound stimuli may be downloaded from the MARCS Auditory Labs Web site: http://sites.uws.edu.au/research/marcs/research/human/icons.htm

Correspondence concerning this article should be addressed to Catherine Stevens, MARCS Auditory Laboratories, University of Western Sydney—Bankstown, Locked Bag 1797, South Penrith DC NSW 1797, Australia. E-mail: kj.stevens@uws.edu.au

similarities between the thing to be represented and the representing system, such as descending pitch to represent a falling object. Finally, symbolic mappings are essentially arbitrary, relying on social convention for meaning. Symbolic mappings involve abstract sounds that acquire meaning through cultural learning, for example, telephone dial tones and sirens.

A more exhaustive taxonomy of signal–referent relations has been provided by Familant and Detweiler (1993). Although formulated in the context of visual icons, Familant and Detweiler's scheme has the potential to be modified to accommodate a broad spectrum of auditory icons as well as speech and abstract sounds. Its main advantage lies in acknowledgment of distinctions between direct and indirect forms of reference and iconic and symbolic relations. When adapted to the auditory domain, Familant and Detweiler's notion of *direct reference* indicates a signal–referent relation in which there is only one referent involved, the denotative referent (e.g., machine gun fire sound for firing machine gun). The denotative referent is the event that is the target of the warning or message. *Indirect reference* occurs when there are at least two referents. Typically the signal refers to the denotative referent through an intermediate sign referent. In other words, sign referents in indirect relations serve as surrogates for denotative referents that may be difficult to portray. For example, on the desktop of a Macintosh computer, the file removal program (the denotative referent) is represented by the visual image of a trash can (the sign referent). An example from the auditory domain would be use of the sound of rattling trash cans—the sign referent—to signal that a computer file has been successfully deleted—the denotative referent.

A New Classification of Auditory Warnings

The classification scheme outlined here organizes potential auditory warnings according to type of signal–referent relation (Keller & Stevens, 1999). The structure is borrowed from Familant and Detweiler's (1993) taxonomy. However, modifications and extensions have been made to their visuocentric framework to bolster suitability for describing auditory signals and to enable generalization to contexts beyond computer interfaces. Psychological processes involved in the recognition of auditory information motivate these changes.

Conceptual structure. The classification of auditory warnings is based on a three-tiered interactive structure (see Figure 1). The lowest tier (Level I) specifies type of signal, including speech, abstract sounds, and auditory icons (which span both natural and artificial environmental sounds; e.g., geophysical and mechanical sounds, respectively). Immediately above the signal type level is a tier (Level IIa) that specifies type of sign relation (i.e., association between a signal and its immediate referent). Sign relations are assumed to exist along a dimension of iconicity. At the high iconicity end of the continuum are sound signals that are related to their referents causally in a "principled, systematic way (described by physics)" (Gaver, 1986, p. 173). Typically, the signal is the sound made by the referent event in the everyday world. At the low end of the iconicity continuum, signal and referent have a manufactured relationship that may be either nonarbitrary (based on some form of physical similarity or appropriateness) or arbitrary (artifactual and random).

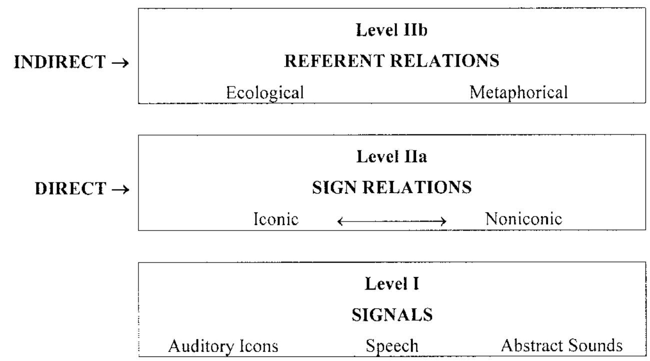

Figure 1. Schematic structure of the classification scheme for auditory warnings.

The two tiers of the organizing framework described so far are sufficient to account for Familant and Detweiler's (1993) notion of direct reference, in which the event identified with the sound signal is the target referent. The top tier (Level IIb) specifies type of referent relation (association between surrogate and target referents) and only comes into operation in indirect signal referent relations. The main division in the indirect reference branch of the hierarchy is between ecological and metaphorical relations. These relations were not explicitly represented in the nomenclature of Familant and Detweiler's taxonomy. In ecological relations, target and surrogate referents are distinct and not necessarily physically similar yet are identifiable with one another because they tend to coexist in the world. For example, a *seagull* surrogate can be used to refer to the target *ship.* In metaphorical relations, similarity between surrogate and target is based on their equivalence in some respect, such as appearance and/or function (Barker & van Schaik, 2000). For example, *camera* surrogate can be used to refer to the target *eye.* Drawing on Gentner's conception of analogical forms (Gentner, 1983; Gentner & Markman, 1997), the proposed scheme also recognizes several subtypes of metaphorical relation (based on whether the referents themselves are related in terms of auditory, visual, or functional features, or a mixture of such features), but these are not discussed here. Some examples of signal–referent relations from different categories are shown in Table 1.

Psychological processes in learning and recognizing signal–referent relations. The tiered structure of the organizing framework stems from the assumption that both learning and recognizing a warning takes place in stages. The first stage involves identifying the source of the sound signal (sound source identification), and the second stage involves forming or recalling the association between the sound source and the warning message (associative learning and memory). It is proposed that success in recognition depends on the type of signal at the sound source identification stage and the type of signal–referent relation at the associative memory stage. The link between the two recognition stages becomes automatic when they are integrated through learning.

With regard to signal type, exclusive attention is given here to the efficacy of different types of environmental sounds, as previous research has investigated the effectiveness of speech and abstract sounds relative to auditory icons in general (Leung, et al., 1997; Smith, Parker, Martin, McAnally, & Stephan, 1999). Sound

Table 1
Examples of the Three Categories Under Investigation

		Indirect	
Target referent	Direct	Ecological	Metaphorical
HELICOPTER	HELICOPTER	GUNFIRE	MOSQUITO
SNAKE	SNAKE HISS	FRIGHTENED MONKEY	TRAIN
TORNADO	TORNADO	TREE BREAKING	SINK DRAINING

source identification with environmental sounds is influenced by factors such as causal uncertainty (whether the signal is easily confused with other signals) and sound typicality (how typical the signal is of a particular source; Ballas, 1993). However, the prime focus of the current study is associative memory—the second stage of recognition in our framework—specifically the effect of the type of signal–referent relation on the relative ease with which associations are learned and recognized.

Direct signal–referent relations should require little—if any—learning, as they are composed solely of sign relations based on associations encountered in everyday interactions in the world: The sound simply needs to be identified for one to recognize the referent. On the other hand, indirect signal–referent relations are more complex than their direct counterparts in the sense that they consist of both sign and referent relations. The sign relation aspect of indirect relations operates in a similar fashion to sign relations in direct forms of reference, but recognizing the referent relation requires an additional step in processing. The first aim of this study was to investigate whether this extra step is costly in terms of both the amount of exposure required to learn the meaning of a new auditory icon and the amount of time required to apprehend this meaning once it has been learned. The second aim of this study relates to the learnability of ecological versus metaphorical indirect relations.

Target and surrogate referents in ecological relations are associated with one another because of their coexistence in a specific environment: They belong to a common conceptual category. Research on conceptual category formation by Rosch and others indicated that cognitive representation of conceptual categories is organized hierarchically (Medin & Coley, 1998; Rosch, 1973; Rosch, Mervis, Gray, Johnsen, & Boyes-Braem, 1976; Smith & Medin, 1981). In general, in these conceptual hierarchies, superordinate-level categories (e.g., vehicle) subsume basic-level categories (e.g., car), which, in turn subsume subordinate-level categories (e.g., sports car). Within-category members are similar in some crucial respect, whereas between-category members are dissimilar in this respect (Medin & Coley, 1998). Therefore, stronger associations exist within rather than between categories.

In metaphorical relations, target and surrogate referents come from distinct conceptual categories but share common physical and/or functional features. Here it is assumed that such feature overlap enhances the perceived similarity of the target and the surrogate referent and thus facilitates the learning and recognition of their relationship. This assumption follows from Tversky's (1977) feature-based theory of perceived similarity, which postulates that such similarity is high to the extent that the objects or events being compared have many common, but few distinctive, features (e.g., size, shape, acoustic properties). It is important that these features are relevant in the sense that they inform the process of discovering a relationship that is meaningful in the context of the particular comparison being made. In most cases, relevant features are those that are most typical (Rosch & Mervis, 1975) and most salient (Tversky, 1977) in this context. Therefore, referent relations should be most easily learned and recognized when there is a high degree of overlap between the most typical and salient features of the surrogate and target referents. The foregoing raises the question of whether (a) membership within a common conceptual category (in ecological relations) or (b) featural overlap (in metaphorical relations) is more important in determining the learnability of target-surrogate–referent relations.

Hypotheses

Our main hypothesis was that direct relations are easier to learn than indirect relations because they are simpler and they capitalize on associations already formed during prior real-world experiences. Furthermore, we expected that—within the indirect branch of the scheme—ecological relations would be learned more readily than metaphorical relations because target and surrogate referents in ecological relations reside within a common real-world domain and hence are (presumably) linked within a preexisting conceptual hierarchy. For instance, in the *seagull–ship* example mentioned earlier, both are members of a superordinate-level category that could be labeled *sea-related objects.* In metaphorical relations, however, new links between distinctive conceptual categories need to be established through processes such as feature matching and analogical reasoning (Barker & van Schaik, 2000; Gentner & Markman, 1997; Tversky, 1977). It is an open question whether the new links in metaphorical relations are as easily recognized as the preexisting links in ecological relations once they are learned.

Experiment 1

Aim, Design, and Hypotheses

The aim of Experiment 1 was to identify surrogate referents as candidates for indirect relation categories. The experiment examined the strength of conceptual links between targets and referents and used pictures and words to control for any confound with the identifiability (typicality and causal uncertainty) of sounds. Picture and word pairs also enabled simultaneous presentation of items. The independent variable, referent relation, consisted of two levels, ecological relations and metaphorical relations. The dependent variable was ratings of strength of association. We hypothesized that associations between surrogate referents and target events are perceived to be stronger in the ecological category than in the metaphorical category. Within each category, the surrogate example that is perceived by participants to have the most features in common with and be most closely related to its target event will be retained for use in subsequent experiments.

Method

Participants. The sample consisted of 90 student volunteers (61 women, 29 men; median age = 19 years, M = 20.87 years, SD = 5.18) enrolled in Psychology I at the University of Western Sydney. Students received course credit for participating. Participants were blind to the hypotheses.

Stimuli and equipment. Stimuli consisted of labeled pictures of target and surrogate referents. As an example, the stimulus item associated with the target *helicopter* included an image of a helicopter (*Art Explosion 250,000 for Macintosh,* 1996) with the label *HELICOPTER* beneath it. Similarly, the metaphorical surrogate for helicopter, *mosquito,* was represented by an image of a mosquito with the label *MOSQUITO.* The visual image was used only to disambiguate the target. For example, the word *glasses* (used as a practice item) could have referred either to eyeglasses or drinking glasses. Several examples of surrogate referents were tested for each target event. There were three ecological surrogate examples and three metaphorical surrogate examples associated with each target referent. This required nine Target Referents × Two Categories (ecological, metaphorical) × Three Surrogate Referent examples, yielding a total of nine target event stimulus items and 54 surrogate referent stimulus items. The set of targets and referents is listed in Table 2. The basis for each metaphorical association is shown in parenthesis. The experimental trials were sequenced in SuperLab version 1.74 (Haxby, Parasuraman, Lalonde, & Abboud, 1993) and presented to participants using a Macintosh Powerbook computer and an Epson datashow projector.

Procedure. Each surrogate referent and its related target event item were presented simultaneously and remained on-screen for 40 s. There was a 3-s intertrial interval, and a doorbell sound was used to signal the start of a new trial. Trials were presented in random order with different stimulus sequences to distribute serial order effects. Participants were exposed to both the ecological and the metaphorical sets of relations; condition order was counterbalanced. For each surrogate–target pair, participants (a) listed the main features shared by the surrogate and target event and (b) rated on a 5-point, 3-label rating scale the degree of relatedness between surrogate and target event. The rating scale extended from *unassociated* (1) to *moderately associated* (3) to *highly associated* (5). Participants were tested in two groups of 45. The experiment lasted 50 min.

Table 2
Targets and Surrogate Referents for Experiment 1

Target	Ecological	Metaphorical
BABY	Rattle	Mouse (small)
	Cradle	Rooster (noisy in the morning)
	Toy	Ball (bouncy)
HORSE	Cow	Race car (fast)
	Brush	Tractor (strong worker)
	Saddle	Spring (leaps)
SNAKE	Tree	Poison (fatal)
	Monkey	Apple bite (Biblical allusion)
	Rock	Train (shape)
FIRE	Fire engine	Tiger (color)
	Match	Bull (aggressive)
	Hose	Light bulb pop (heat)
THUNDER	Rain	Bomb (explosive)
	Skid	Truck (loud)
	Branch	Hand clap (sound quality)
TORNADO	Glass	Earthquake (natural disaster)
	Waves	Sink draining (spinning motion)
	Door slam	Stampede (destructive)
BOAT	Water	Bus (transport)
	Seagull	Balloon (requires wind)
	Bell	Bowl (hollow vessel)
HELICOPTER	Machine gun	Fan (spinning blades)
	Winch	Mosquito (shape)
	Seat belt	Spinning top (motion)
CAGE	Bird	Jar (container)
	Bear	Zipper (closure)
	Whip	Whistle (law enforcement)

Note. The basis for each metaphorical association is shown in parenthesis.

Results

It was hypothesized that ecological relations would be given higher strength of association ratings than would metaphorical relations. A one-way analysis of variance (with α = .05) revealed that as predicted, association strength was significantly higher for ecological (M = 2.47, SD = 1.35) than for metaphorical relations (M = 1.44, SD = 1.25), $F(1, 88) = 327.4$, MSE = 0.14, $p < .05$, Cohen's f = 0.3. There was no significant effect of the order in which conditions were presented, $F(1, 88) = 0.85$, MSE = 0.33, $p > .05$, Cohen's f = 0.032.

The most common shared features of surrogate referents and target events noted by participants are summarized in Table 3. The mean ratings in Table 3 refer to the most highly rated of the three surrogates paired with each target, hence means differ from those reported above (maximum rating is 5—*highly associated*). The open-ended nature of this part of the experiment yielded a large amount of data (up to five descriptors for each of 54 associations by 90 participants). Two raters were provided with basic categories and subcategories suggested by the icon designer in light of concepts of similarity and association (Ballas, 1993; Tversky, 1977). The raters undertook the task of entering descriptive terms into a spreadsheet containing the categories, creating additional subcategories as required, and tallied the occurrence of each term. The categories and subcategories of similarity or association were as follows: environment, structure (configuration or shape, size, colour/texture, material, sound, duration, and temperature), function (role, potential, and movement), and quality (conceptual, emotional). Synonyms were treated as similar items. The raters were naive to the purpose of the categorization.

There was good consensus among responses from participants. For example, for each of the most highly rated ecological and metaphorical pairs there was a dominant response with similar wording used. Table 3 shows that the mean agreement (or proportion of the 90 participants who gave the same written response) ranged in the ecological condition from a minimum of 62.2% for *fire–fire engine* to a maximum of 98.9% for *horse–cow.* In the metaphorical condition, minimum agreement on the most common description was 67.0% (*helicopter–fan*) whereas maximum agreement of 92.0% occurred in response to the pair *cage–jar.* The mean agreement for the most highly rated pairs was 81.5% and 82.8% for ecological and metaphorical relations, respectively. One can also see in Table 3 that the most common descriptions for ecological items pertained to the relationship between the targets

Table 3
Mean Ratings of Association Strength, Standard Deviation, Most Common Description of Association, and Percentage Agreement in Experiment 1

Target	Referent	Mean rating	*SD*	Most common description	% agreement
		Ecological			
BABY	Rattle	3.40	.89	Baby uses rattle	84.44
HORSE	Cow	3.78	.81	Farm animals	98.89
SNAKE	Tree	2.11	1.11	Bush/jungle	91.11
FIRE	Fire engine	3.69	.61	Engine extinguishes fire	62.22
THUNDER	Rain	3.62	.66	Storm	84.44
TORNADO	Glass	2.46	1.25	Tornado breaks glass	65.56
BOAT	Water lapping	3.70	.68	Boat floats on water	80.00
HELICOPTER	Machine gun	1.81	1.09	War	83.33
CAGE	Bird	3.20	1.04	Cage houses bird	84.44
M		3.02	.90		81.60
		Metaphorical			
BABY	Mouse	1.0	1.03	Small	82.0
HORSE	Racing car	2.3	1.91	Transport/racing	82.0
SNAKE	Poison	2.7	.99	Deadly/poisonous	76.0
FIRE	Tiger	1.2	1.02	Dangerous	74.4
THUNDER	Bomb	2.3	1.10	Loud	86.0
TORNADO	Earthquake	3.3	.92	Natural disaster	90.0
BOAT	Bus	2.3	1.06	Transport	95.6
HELICOPTER	Fan	2.1	1.10	Blades	67.0
CAGE	Jar	1.9	1.08	Both hold things	92.0
M		2.1	1.05		82.8

and their surrogates in the everyday world, whereas the descriptions for metaphorical items related more to common features (with the possible exception of the *horse–racing car* and *tornado–earthquake* items).

Discussion

As hypothesized, strength of association between target and surrogate referent was higher in the ecological than in the metaphorical relation condition. This corroborates the assumption that the target and surrogate referent pair in an ecological relation comes from a common domain, conceptual hierarchy, or category of events. The mean ratings indicate that in general, there is a plausible association between target and surrogate referents forming a metaphorical relation. However, new links between distinctive category frameworks may need to be established. It is important to note that the ratings and free-response descriptions indicate that metaphorical relations are not purely idiosyncratic on the part of the icon designer, as there was good agreement among participants in identifying an association between the target and surrogate referent in metaphorical relations. A question remains as to whether there is any difference in the rate at which ecological and metaphorical relations can be learned and the speed with which they are recognized. Experiment 2 examines the ease with which indirect ecological and metaphorical relations can be learned relative to each other and relative to direct relations. The ecological and metaphorical target–referent relations that were given the highest ratings of associative strength in Experiment 1 were used as stimuli in Experiment 2.

Furthermore, given the proclivity of humans to seek meaning and our capacity to learn symbolic relations such as written language, a condition was included in Experiment 2 that involved referent relations in which the target and surrogate were unrelated. In these random relations, target and surrogate referents from the metaphorical condition were scrambled such that there was no discernible relation (in the experimenters' minds) between target and referent. The random relations were used to control for any general ability to form associations between sounds and events irrespective of whether they are related in a meaningful way.

Experiment 2

Aim, Design, and Hypotheses

The aim of Experiment 2 was to investigate the effect of referent relation type on the relative ease of learning auditory target–referent relations. The one-way between-subjects design consisted of the independent variable, relation, with four levels: direct, indirect–ecological, indirect–metaphorical, and indirect–random metaphorical. The first dependent variable was the number of trials required to reach criterion performance, with the criterion being 100% accuracy on two consecutive blocks of test trials. Response time as the second dependent variable held the potential to shed light not only on the learnability of icons but also on the speed of recognition once relations are learned.

We hypothesized that direct relations are learned in fewer training trials than are indirect relations and that reaction time is fastest in response to direct relations, followed by ecological and then metaphorical relations. Additionally, we hypothesized that learning and recognition of random relations is poor relative to intended metaphorical target–referent relations.

Method

Participants. The sample comprised 64 participants (53 women and 11 men). Participants were students and staff of the University of Western Sydney and volunteers from the community. Sixteen participants were randomly assigned to each of the four conditions (direct: $M = 21.13$ years, $SD = 4.59$; ecological: $M = 22.94$ years, $SD = 6.27$; metaphorical: $M = 21.88$ years, $SD = 3.74$; random: $M = 22.56$ years, $SD = 6.96$). All participants had self-reported normal hearing and were blind to the hypotheses.

Stimuli and equipment. The sounds used as stimuli were obtained from *sfx, Sound Effects for Movies and Videos* (1992) CDs, and the Internet site www.sounddogs.com. All sounds were 4–6 s in length. The relatively long duration was needed to ensure clarity of the concept or event being conveyed, for example, starting the engine of a bus or the complete roar of a tiger. The mean length of sounds in the direct, ecological, and metaphorical conditions was 5,306 ms ($SD = 790.81$), 5,029 ms ($SD = 631.35$), and 5,388 ms ($SD = 842.91$), respectively. There was no significant difference between sound lengths across conditions, $F(2, 24) = 0.584$, $MSE = 578{,}158.3$, $p > .05$, Cohen's $f = 1.54$. The ecological and metaphorical referents used as stimuli and descriptions of three sets of sounds are shown in Table 4. Sounds were sampled at 44.1 kHz with 16-bit resolution, recorded in mono, and normalized for amplitude across conditions.

The experiment was programmed in C++ and presented on a Gateway PC through headphones. Nine clickable buttons were arranged in an arc on the computer screen such that each button was equidistant from a centrally located *ready* button. Each button had a label that named a unique target. Random–metaphorical pairs were created by pairing targets and referents that had no obvious relationship; these referents are shown in the final column of Table 4.

Procedure. Participants were told that they would be presented with a series of items and sounds and that their task would be to learn the item–sound pairs. Blocks of nine training trials alternated with blocks of nine test trials. Each item–sound pair was presented once in random order in each trial. In the training trials, a word appeared on the screen (the target) and a sound was played (the referent).

The blocks of alternating training and test trials were presented to participants with a new random order of stimuli in each block. Participants indicated their responses using a mouse click on a labeled target button, and response time was recorded from the onset of the sound. Participants were instructed to respond as quickly and as accurately as possible—they could respond at any time, and the sound would continue to play. After each trial, participants returned the mouse cursor to the central "*ready*" position. The ordinal position of the labeled buttons was randomized at the start of each block of training and test trials. Training and test trial blocks alternated until criterion performance was achieved—100% accuracy on two consecutive test-trial blocks. An accuracy score out of 9 was presented to participants on the computer screen at the end of each training block. The experiment lasted 30 min.

Results

Three planned orthogonal contrasts were used to test the hypotheses that (a) direct relations are learned more readily than are ecological relations, (b) ecological relations are learned more readily than are metaphorical relations, and (c) metaphorical relations are learned more easily than are random relations. Results based on the number of trials required to reach criterion performance are reported first, followed by response time analyses.

The average number of trials required to reach criterion performance in each of the four conditions was as follows: direct, $M = 2.63$ ($SD = 0.89$); ecological, $M = 4.00$ ($SD = 0.97$); metaphorical, $M = 4.18$ ($SD = 1.17$); and random, $M = 8.81$ ($SD = 3.33$). The distinction between frequency distributions for trials required to reach criterion for each of the four signal-referent relations is shown in Figure 2. In accordance with the first hypothesis, significantly fewer trials were needed to reach criterion in the direct condition relative to the ecological condition, $F(1, 60) = 4.27$, $MSE = 3.54$, $p < .05$, Cohen's $f = 0.27$. Figures 3a and 3b illustrate the number of trials required to reach criterion in the ecological and metaphorical conditions, respectively. The second hypothesis, that ecological relations are learned more readily than are metaphorical relations, was not supported, $F(1, 60) = 0.08$, $p > .05$, Cohen's $f = 0.04$. However, the third hypothesis received support: Random relations took significantly more trials to learn than did metaphorical relations, $F(1, 60) = 48.29$, $p < .05$, Cohen's $f = 0.90$.

Response times were analyzed for correct responses only, using the same three planned orthogonal contrasts as before. The mean number of errors recorded during training trials, collapsed across the four conditions, was 8.92 trials, $SD = 2.04$. Mean number of errors per condition and mean response times per condition on the first and final test trials are shown in Table 5. Contrary to the first two hypotheses, there was no significant difference in response times in the first test trials when comparing direct and ecological relations, $F(1, 60) = 2.21$, $MSE = 3{,}126.5$, $p > .05$, Cohen's $f = 0.19$, or ecological and metaphorical relations, $F(1, 60) = 2.23$, $p > .05$, Cohen's $f = 0.19$. The third hypothesis was supported: Response time was significantly longer for random than for metaphorical relations on both the initial test trial, $F(1, 60) = 8.7$, $p < .05$, Cohen's $f = 0.12$, and the last test trial (i.e., where 100% accuracy had been reached), $F(1, 60) = 22.61$, $p < .05$, Cohen's $f = 0.61$.

Table 4
Targets, Surrogate Referents, and Sounds for Experiment 2

Target	Ecological	Metaphorical	Random
BABY–cry	Rattle–shake	Mouse–squeaking	Earthquake
HORSE–whinny	Cow–moo	Race car–engine	Poison
SNAKE–hiss	Tree–rustle	Poison–aerosol spray	Race car
FIRE–crackle	Fire engine–siren	Tiger–roar	Mouse
THUNDER–roll	Rain–falling	Bomb–explosion	Jar
TORNADO–wind	Glass–breaking	Earthquake–ground breaking	Fan
BOAT–sails flapping	Water–lapping	Bus–engine, horn	Bomb
HELICOPTER–rotor	Machine gun–firing	Fan–starting	Tiger
CAGE–door slam	Bird–squawking	Jar–lid unscrewing	Bus

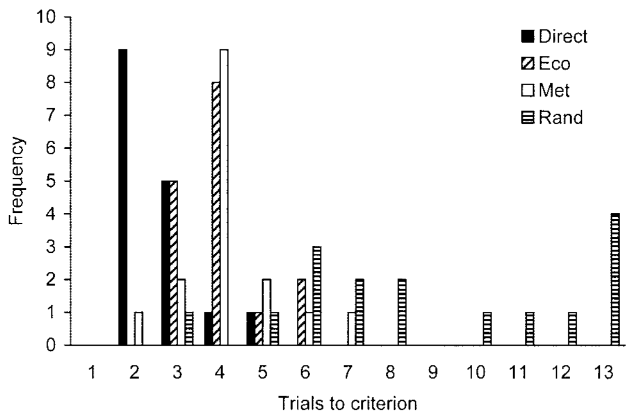

Figure 2. Experiment 2: Frequency distribution (based on number of participants) of trials required to reach criterion as a function of the four types of signal–referent relations. Eco = ecological; Met = metaphorical; Rand = random.

Discussion

The results from Experiment 2 support the hypothesis that direct relations involve immediate and rapid recognition of a target from its sign relation. Direct relations are simple and capitalize on associations recognized or learned previously in everyday interactions in the world. Indirect relations resulted in longer acquisition, consistent with greater conceptual distances or levels in a cognitive hierarchy of long-term knowledge (Barker & van Schaik, 2000; Medin & Coley, 1998; Rosch, 1973). Regardless of the differing types of associations in ecological and metaphorical relations, both kinds of relation convey information and can be learned. Moreover, once learned ecological and metaphorical relations are recognized as rapidly as are direct relations.

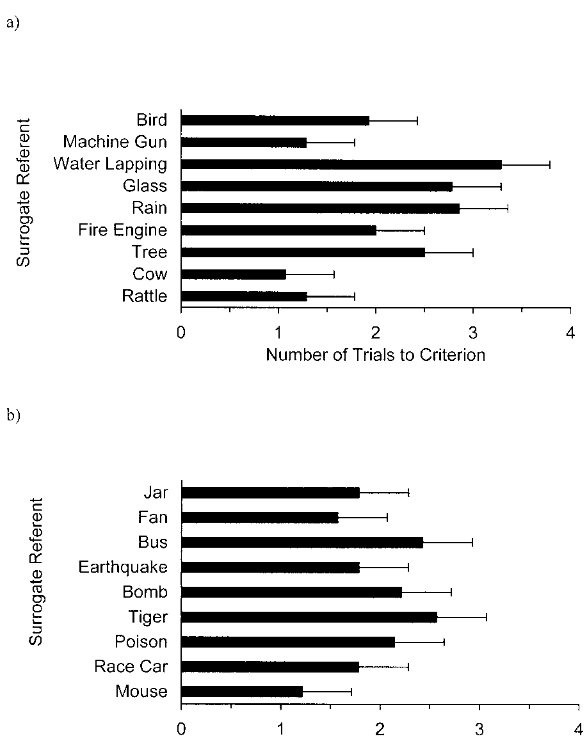

Figure 3. Experiment 2: Mean trials to criterion for surrogate referents in (a) ecological and (b) metaphorical conditions. Errors bars refer to standard error of the mean.

Table 5
Response Times (in s) on First and Final Test Trials and Error Rates During Training in Experiment 2

Relation type	First test trial RT *M*	First test trial RT *SD*	Final test trial RT *M*	Final test trial RT *SD*	Mean errors *M*	Mean errors *SD*
Direct	2.87	0.79	2.88	0.76	0.56	0.73
Ecological	3.80	0.99	3.22	0.76	2.94	1.73
Metaphorical	4.72	1.16	3.26	0.91	3.94	0.77
Random	6.57	3.09	4.62	0.81	28.25	4.94

Note. RT = response time.

Significantly poorer performance in response to random versus intended metaphorical relations illustrates the potential for using ecological or metaphorical referent relations as an informative or communicative device—performance is accurate only when there is some discernible and memorable link between target and surrogate referents.

The results of Experiments 1 and 2 reflect humans' ability to recognize direct relations and their capacity to recognize or learn target–referent relations on the basis of shared ecological cues or more conceptual, metaphorical relations. However, in designing auditory icons and target–referent relations there are two issues that further complicate the matter. The first is equating the iconicity and identifiability of the sound signals—to what extent is the sound imageable (Paivio, 1969, 1971), able to be labeled (Bartlett, 1977; Bower & Holyoak, 1973), and what is its articulatory directness or degree of physical correspondence to the referent event (Gaver, 1986; Hutchins, Hollans, & Norman, 1986)? The second issue relates to the relative salience or meaningfulness of particular target–referent relations.

At first blush, it seems reasonable to attempt to address the iconicity–identifiability issue by conducting a speeded perceptual identification task and to address the salience–meaningfulness issue by conducting an item analysis of target–referent pairs using, for example, the data collected in the present experiment. However, it is plausible—indeed likely—that both sound iconicity and relation salience interact. For example, a sound with good iconicity–high identifiability may still be a component of a weak signal–referent relation or vice versa. Hence, a preferable solution is to contrive a context in which the set of sounds is common across conditions (ecological and metaphorical relations) and only the type of relation differs. Such a fully balanced design formed the basis of Experiment 3—target and surrogate referents of indirect relations were reversed so that the target became the referent and the referent became the target. All surrogate referent sounds were taken from the original set of direct relation stimuli (used in the present experiment) to ensure that the sounds were equally identifiable and only relation type differed. Thus, the fully balanced design neutralized any differences in sound iconicity to precipitate any potential differences in the learnability of ecological versus metaphorical relations that may have been obscured in the present experiment. Although the main focus of Experiment 3 was to investigate whether indirect target–referent relation type (ecological vs. metaphorical) affected learnability, the use of sounds from the direct condition in the present experiment allowed the learnability of direct and indirect relations to be observed across the two experiments.

Experiment 3

Method

Participants. The sample consisted of 40 female and male participants with 20 randomly assigned to each condition (ecological: $M = 20.40$ years, $SD = 3.44$, 3 men and 17 women; metaphorical: $M = 23.5$ years, $SD = 7.99$, 5 men and 15 women). Participants were students enrolled in Psychology I at the University of Western Sydney, and they received course credit for participating. All participants reported normal hearing and were blind to the hypotheses.

Stimuli and equipment. The nine sounds were taken from the direct condition set used in Experiment 2, and this time identical sounds were used in the ecological and metaphorical conditions. The on-screen button label (referent) paired with the sounds varied in the ecological and metaphorical relation conditions. These pairs and the direct sounds are listed in Table 6.

The experiment was programmed in PowerLaboratory Version 1.0.3 (Chute & Westall, 1996) and presented on a Macintosh 7300 through headphones. Nine clickable buttons were arranged in an arc on the computer screen such that each button was equidistant from a centrally located *ready* button. Each button had a label that named a unique target.

Procedure. As in Experiment 2, blocks of alternating training and test trials were presented to participants using a new random order of stimuli in each block. Participants indicated their response with a mouse click, and response time was recorded from the onset of the sound. Participants were asked to respond as quickly and as accurately as possible—they were told that they could respond at any time and that the sound would continue to play. After each trial, participants returned the mouse cursor to the central *ready* position that was equidistant from the nine labeled target buttons. The order of the buttons was randomized in each block of training and test trials. Training and test trial blocks alternated until criterion performance was achieved—100% accuracy on two consecutive test trial blocks. The experiment lasted 30 min.

Results

The number of trials required to reach criterion and latency from onset of sound to response were recorded and analyzed using one-way analyses of variance. Consistent with results obtained in Experiment 2, there was no significant difference in the number of trials required to reach criterion for ecological ($M = 6.85$, $SD = 3.62$) or metaphorical conditions ($M = 6.90$, $SD = 2.43$), $F(1, 38) = 0.003$, $MSE = 9.48$, $p > .05$, Cohen's $f = 0.03$. There was also no significant difference in response time recorded either on the first or last test trials in the ecological versus metaphorical conditions: first test block, ecological, $M = 4.96$ s, $SD = 2.16$; metaphorical, $M = 5.46$ s, $SD = 2.06$, $F(1, 38) = 0.56$, $MSE = 4{,}452.5$, $p > .05$, Cohen's $f = 0.01$; last test block, ecological, $M = 4.30$ s, $SD = 1.17$, metaphorical, $M = 4.01$, $SD = 0.84$, $F(1, 38) = 0.80$, $MSE = 1{,}040.3$, $p > .05$, Cohen's $f = 0.01$.

Table 6
Targets, Surrogate Referents, and Sounds for Experiment 3

Target referent		
Ecological	Metaphorical	Surrogate referent
Rattle	Mouse	BABY–cry
Cow	Race car	HORSE–whinny
Tree	Poison	SNAKE–hiss
Fire engine	Tiger	FIRE–crackle
Rain	Bomb	THUNDER–roll
Glass	Earthquake	TORNADO–wind
Water lapping	Bus	BOAT–sails flapping
Machine gun	Fan	HELICOPTER–rotor
Bird	Jar	CAGE–door slam

Discussion

The results of Experiment 3 confirm those of Experiment 2, showing that there is no significant difference in recognition accuracy or speed of response to ecological versus metaphorical relations. This result has been upheld when a common set of sounds has been used to neutralize effects of sound identifiability. These two experiments contribute both to theory and application. The theoretical contribution is the demonstration that although the type of target–referent association differs and association strength of ecological relations was generally rated more highly than that of metaphorical relations (see Table 3), there was no discernible effect on performance. A practical outcome is to legitimize the use of a range of target–referent relations and to expand the base of sounds that can be used for communicative purposes. An important caveat is that training with feedback is required to ensure that ecological and metaphorical relations are memorized effectively.

As indicated in the classification scheme, the design of auditory icons is a two-step process requiring selection of iconic and identifiable sounds as well as salient and meaningful target–surrogate referent relations. Recognition of direct relations (see Experiment 2) requires little training. Beyond that, careful selection and piloting of sounds and relations is needed. A second applied outcome of these experiments is to recognize the benefit of using particular experimental designs to test the effectiveness of sounds and target–referent relations. The fully balanced design of Experiment 3 enables investigation of relation type on performance, whereas the Experiment 2 design enables scrutiny of sound and relation-type interactions.

General Discussion

Three experiments tested the psychological validity of a new conceptual framework for classifying auditory icons according to whether their constituent environmental sound signals are related to referent meaningful events either directly or indirectly (through common ecological or metaphorical associations). In general, our findings demonstrate the potential of a broad variety of auditory icons to convey information efficiently. It was hypothesized that signals that capitalize on preexisting knowledge about associations—direct relations—are relatively easy to learn. The number of trials required to reach criterion suggests that direct relations are acquired more readily than are indirect–ecological relations. These findings are consistent with the notion that direct relations are recognized accurately because they are relatively simple, being composed exclusively of sign relations in which there is a close physical correspondence, or nomic mapping (Coward & Stevens, 2004; Gaver, 1986), between a sound and a referent event (e.g., the sound of a crackling fire to signal fire). Response time was unaffected by relation type. Differences in response times may appear in contexts involving greater cognitive demand and/or where sounds are compressed.

The failure to find a difference in performance with ecological and metaphorical relations can be seen as positive in the sense that it demonstrates that a range of different indirect target–referent relations can be learned. Specifically, our results do not distinguish between the amount of exposure required to learn, and the psychological processing speed for, ecological relations—wherein target and surrogate referents are associated through coexistence within a specific environment—and metaphorical relations—wherein target and surrogate referents are similar in appearance and/or function. This null finding is not trivial, as the results of Experiment 1 indicate that a distinction between ecological relations and metaphorical relations has psychological validity. The ratings of associative strength of ecological signal–referent relations were, in general, higher than those of metaphorical relations. Moreover, when ratings of associative strength were high there was strong concordance among participants' descriptions of the features shared by the target and referent making up the relation. Nevertheless, the findings of Experiments 2 and 3 provide no support for the idea that types of association based on conceptual categories (ecological) or featural similarity (metaphorical) affect learnability differentially. Such differences, if they exist, may emerge only under more demanding, time-limited, or complex task conditions or when associative strength of the two types of relation is manipulated more precisely.

The relatively poor performance in the random condition included in Experiment 2 indicates that performance is optimal only when there is a discernible and memorable relationship between surrogate referent to target: Random pairs of targets and referents were never learned effectively. Indeed, in Experiment 2 it was found that random relations took over three times as much exposure to learn than direct relations and over twice as much exposure as ecological and metaphorical relations; even when learned, response was still approximately 1.5 times slower for random relations than for the other types of relations. Such large differences are noteworthy because they could translate into problems in applied settings wherein auditory icons are used as warnings. In contrast, although there may be some initial costs with ecological and metaphorical relations compared with direct relations (approximately twice as many training trials were needed to reach 100% accuracy), once learned recognition of such systematic indirect relations is robust. More important, this finding invites the recommendation that they should be considered—along with direct relations—as viable candidates for use in warning systems. The encouraging results with metaphorical relations are particularly interesting because they highlight a human predisposition for discovering and exploiting even quite obscure relationships between environmental contingencies. This predisposition to "find meaning" effectively widens the potential database of sounds that can be used as auditory warnings. Indeed, extraordinary relations may elicit deep processing (Craik & Lockhart, 1972), resulting in good retention.

The real-world motivation for these experiments was to examine, principally, whether auditory icons could be used to convey information in complex environments in which the visual system was at risk of overload. The scrutiny of relevant literature revealed a paucity of theories about auditory cognition and the way in which humans learn about auditory events and their meanings. The first contribution then has been the development of a theory and classification scheme relating to the learning and recognition of auditory icons. Furthermore, the empirical work has significance for applied settings. It has been demonstrated that direct relations are recognized immediately and that nine direct target–referent relations are identified as effective auditory icons. We have also demonstrated that indirect relations—both ecological and metaphorical—are recognizable after a period of exposure and training. Eighteen indirect relations were designed and tested. They involved concrete images and concepts, and their image-ability is likely to assist with memorization. We speculate that the concrete and distinctive quality of indirect relations may also facilitate recognition of a relatively large set of icons in demanding conditions and may outperform abstract sounds such as earcons (Blattner et al., 1989).

The purposes to which the current set of auditory icons may be put include their use to convey information in which the visual system is overloaded, speech is masked by background noise, the language of operators varies, or auditory events enhance explanation and understanding in an educational context. For example, in an aviation context the approach of a helicopter can be signaled by the sound of the rotor, a machine gun, or a fan; changes of weather can be communicated using a thunder roll, rain fall, or wind; and extreme weather changes can be communicated using the sound of glass breaking or ground quaking. The use of the current set of auditory icons as warning signals may be possible after further investigation of the ease of recognition when tasks are complex, response time is limited, and sounds have been shortened.

Because we have demonstrated that a variety of target–referent relations or auditory icons can be learned, researchers can begin to work with shorter versions of the current sounds. At present, sound signal length is similar to that of a spoken utterance—as warning signals, shorter sounds and concomitantly faster response times are desirable. One possibility is to compress and meaningfully manipulate parameters of these sounds. It also remains to consider the strength or fragility of direct and indirect relations with increases in cognitive workload, which can be manipulated using dual-task paradigms, limiting response time, and manipulating the size of the set of auditory icons to be learned. The results of the present experiments highlight the importance of sound iconicity and signal–referent relation potency and provide the basis for a set of direct and indirect relation norms. The conceptual framework described here and the initial findings serve as a foundation on which future investigations of auditory icons can be built.

References

Art Explosion 250,000 for Macintosh [Computer software]. (1996). Aurora, CO: Hallogram Publishing.

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.

Bartlett, J. C. (1977). Remembering environmental sounds: The role of verbalization at input. *Memory & Cognition, 5,* 404–414.

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.

Bower, G. H., & Holyoak, K. (1973). Encoding and recognition memory for naturalistic sounds. *Journal of Experimental Psychology, 101,* 360–366.

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

Coward, S. W., & Stevens, C. J. (in press). Extracting meaning from sound: Nomic mappings, everyday listening, and perceiving object size from frequency. *The Psychological Record.*

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

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

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

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

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.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155–170.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52,* 45–56.

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

Haxby, J. V., Parasuraman, R., Lalonde, F., & Abboud, H. (1993). Superlab: General-purpose Macintosh software for human experimental psychology and psychological testing. *Behavior Research Methods, Instruments & Computers, 25,* 400–405.

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.

Keller, P., & Stevens, K. (1999). *A taxonomy for auditory warnings: Learning associations between auditory icons and their referents.* Unpublished manuscript, University of Western Sydney, MARCS Auditory Laboratories, Sydney, Australia.

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

Medin, D. L., & Coley, J. D. (1998). Concepts and categorization. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 404–439). San Diego, CA: Academic Press.

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.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76,* 241–263.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

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

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 4,* 328–350.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnsen, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–440.

sfx Sound Effects (CD). (1992). Fort Lauderdale, FL: RTV.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, S., Parker, S., Martin, R., McAnally, K., & Stephan, K. (1999). Evaluation of sound types for use as auditory warnings. [abstract] *Australian Journal of Psychology, 51*(Suppl.), 42.

Sound effects for movies and videos (CD). (1995). Amsterdam, Holland: Digimode.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Received July 23, 2001
Revision received September 30, 2003
Accepted September 30, 2003 ■