Safety Science 61 (2014) 29–35

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

Multimodal warnings to enhance risk communication and safety

Ellen C. Haas [a,1], Jan B.F. van Erp [b,*]

[a] Multimodal Control and Displays Laboratory, Human Research and Engineering Directorate, U.S. Army Research Laboratory, Bldg. 459, Aberdeen Proving Ground, MD 21005, USA

[b] TNO, Department of Perceptual and Cognitive Systems, PO Box 23, NL-3769ZG Soesterberg, The Netherlands

ARTICLE INFO

Article history:
Available online 29 August 2013

Keywords:
Auditory warnings
Tactile warnings
Multimodal displays
Risk communication
Safety

ABSTRACT

Multimodal warnings incorporate audio and/or skin-based (tactile) cues to supplement or replace visual cues in environments where the user's visual perception is busy, impaired, or nonexistent. This paper describes characteristics of audio, tactile, and multimodal warning displays and their role in risk communications. The authors demonstrate that visual–auditory and visual–tactile displays can be significantly more effective than visual displays alone in enhancing user performance. The authors describe signal design guidelines, and illustrate the importance of knowledge of user attentional constraints and limitations in effectively using multimodal displays to communicate safety information. Finally, conclusions and recommendations for future multimodal warning display design and research are presented.

Published by Elsevier Ltd.

1. Background

Warnings are usually considered to be visual signs or product labels (Wogalter and Mayhorn, 2006). However, audio and skin-based (tactile) warnings can be used to supplement or replace visual cues to enhance risk communications. The ability to present warnings in alternate or multiple modalities is especially useful in situations where the user's visual perception in busy, impaired, or nonexistent. This paper will describe research related to the use of multimodal warnings, which can take the form of displays and signals, as new tools and approaches to enhancing risk communication and safety. First, characteristics of audio displays and their role in risk communications will be described. Next, tactile display warnings will be discussed. The third section will illustrate the use of combined audio and tactile warnings in multimodal displays. Finally, the authors will describe conclusions and recommendations for future research, and include signal design guidelines that describe how different sensory cues can be used, in isolation and in combination, for multimodal warnings.

1.1. Auditory warnings

Auditory warnings can be thought of as sound that is intentionally made and specifically designed not only to attract attention, but also to provide additional information about the nature of the warning event (Haas and Edworthy, 2006). Over the years, researchers and practitioners have found that auditory alarms can be useful in themselves or as supplements to visual signals because they are omnidirectional and cannot be involuntarily shut off. Durlach and Colburn (1978) stated "the fact that nature did not provide us with earlids is probably due to…the use of the acoustical channel for warning signal function to which it is exceptionally well matched."

Auditory warnings may include verbal cues that incorporate human speech in recorded, digitized, or synthesized form. Speech warnings are effective because they are highly redundant in the sense that a speech signal contains more information than necessary for sound identification (Shannon et al., 1995). However, speech cues may be at times conflicting or reinforcing of other auditory cues in the signal. Deatherage (1972) noted that speech warnings are most useful when the listener has no special training in coded signals, if workload or stress could cause them to forget the meaning of a coded signal, if there is a necessity for rapid two-way exchanges of information, or when the warning message deals with a future time requiring some preparation (i.e., a countdown to an event, in which tonal signals could be miscounted).

Auditory warnings may also consist of nonverbal sounds, which are most useful if the sounds can be easily associated with the warning condition itself (Haas and Edworthy, 2006). Nonspeech warning signals include tones or other nonverbal elements (i.e., tones, buzzers, klaxons, bells, or digitized tones) that use auditory elements to signal events to the listener. Tonal warnings can consist of single or multiple tones. Single-tone signals consist of one tone presented during the duration of the signal. This definition can be relaxed to include repetitions of the same tone, where the tone itself does not change, but simply repeats itself (with silences in between) and, therefore, functions as a single tone. Patterson (1982) suggested that warning signal duration (including onset

* Corresponding author. Tel.: +31 888 665 982; fax: +31 346 353 977.
E-mail addresses: ellenhaas@comcast.net (E.C. Haas), jan.vanerp@tno.nl (J.B.F. van Erp).
[1] Tel.: +1 410 688 7232; fax: +1 410 278 5997.

0925-7535/$ - see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.ssci.2013.07.011

and offset) should be at least 200 ms to allow the ear enough time to integrate the warning signal. Patterson also recommended that the signal duration not be overly long to prevent disruption of communication. Multiple tone signals consist of two or more different tones presented during one signal duration. Multitone signals can have a simple or complex harmonic structure, and because of their multitonal nature, may carry more information than single tonal warnings, although the burden is on the signal designer to choose sounds that can be easily associated with the warning condition (Wogalter, 1994). Patterson (1982) suggested that warning signal duration (including onset and offset) should be at least 200 ms to allow the ear enough time to integrate the warning signal. Patterson also recommended that the signal duration not be overly long to prevent disruption of communication.

There are advantages to both single- and multiple-tone auditory warnings. Multitone signals can be advantageous because they permit variations in signal pitch, loudness, and inter-tone spacing, so that the resultant warning has a distinctive temporal and pitch pattern, which may also make them easy to learn (Haas and Edworthy, 2006). This multitonal distinctiveness confers greater individuality on the warning, increases its potential to signal particular events, and also increases its resistance to masking from environmental noise (Patterson, 1982). However, single-tone auditory warnings may also be easily learned (Edworthy and Hellier, 2006). Traditional warnings of the single tone type, or of a repeating single tone (such as a repeating bell or klaxon) can be easily recognized and have particular associations with particular hazards. Lazarus and Hoge (1986) noted that tonal auditory warnings could be retained, not just learned, when they form well-established associations between signals and situations, stemming from the way those warnings have been used in the environment or in the workplace. They showed, for example, that sirens have an association to danger or threat.

Auditory icons, a special subset of nonspeech auditory signals, incorporate evocative sounds that can be used to describe the function they represent (Gaver, 1989), such as the sound of a "thunk" to indicate a document being delivered to a computer's recycle bin. Auditory icons have been used as emergency warnings, including in-vehicle collision warnings (Graham, 1999). Graham's in-vehicle auditory icons, which consisted of the sounds of a car horn and skidding tires, were compared with two conventional auditory warnings, which consisted of a simple tone and a voice saying "ahead." Graham found that the auditory icons produced significantly shorter braking reaction times than the conventional auditory warnings. However, he also found that auditory icons produced a greater number of inappropriate responses, such as braking in response to non-collision situations. As a result, Graham argued that signal designers should take care in assigning auditory icons the appropriate urgency and inherent meaning linked to the warning situation. Research indicates that effective auditory icons must be audible, identifiable, and interpretable, and have strong preexisting associations in order to facilitate learning and retention (Stephan et al., 2006).

Technological advances have made it possible for both verbal and nonverbal auditory cues to be presented spatially. With spatial audio displays, also known as three-dimensional (3-D) audio displays, a listener perceives spatialized sounds that appear to originate at different horizontal and vertical locations and distances outside the head, much like audio cues that naturally occur in the environment. Earphones are often used to present spatial audio cues. Although loudspeakers may be used, their use may be problematic because the listener may not be located in the precise location (relative to the loudspeakers) to allow the sound to be perceived as spatialized (Shilling and Shinn-Cunningham, 2002). Before the audio cues reach the earphones, they are filtered through computerized sound filter functions known as head-related transfer functions (HRTF's), which provide the sound with specific time, intensity, phase and reverberation cues that make the resulting sound appear to originate from different locations in space. A headtracker is often used to provide a stable audio cue reference point. Because each sound is presented in a different spatial location, listeners can use that spatial cue to selectively attend to more than one sound at a time as well as to sounds at designated locations. Relevant spatial audio safety applications include mitigating aircraft pilot spatial disorientation and providing meaningful direction-related system warnings (Endsley and Rosiles, 1995), as well as detecting target messages in continuous monitoring tasks (McAnally and Martin, 2007). However, Doll and Folds (1986) suggested that the use of nonspeech and speech signals requires caution in that nonverbal signals should be designed to enhance distinctiveness and masking resistance. In addition, the potential occurrence of concurrent speech and nonspeech warning signals should be avoided, if possible.

1.2. Tactile displays

Tactile displays present information through the user's sense of touch. For example, tactile labels can be presented as passive displays in the form of raised text, dots or symbols that the user has to explore through touch to perceive its meaning. Other passive tactile displays include tactile key identifiers (raised areas found on the F and J keys of keyboards to help identify finger position), raised areas on banknotes corresponding to note denomination, and notches on phone cards to help the user identify the card orientation. Passive tactile displays are also employed to present warnings on chemical products that contain raised line symbols as an important means to communicate warnings to the visually impaired. The exploration, processing and understanding of passive labels may require training.

More recently, a second category of tactile displays has emerged termed "active displays". This category presents a warning that the user can perceive without the need of exploration. A good example is the pager that became popular in the 1990s. In silent mode, the pager presented a warning through a vibration to the user's skin. Recently, active tactile warning displays are finding their way into everyday use, the vibration function of mobile phones being a well-known example (for an overview, see Jones and Sarter, 2008). Active tactile displays have numerous advantages, including the flexibility in warnings to be presented, and the option to combine or integrate tactile with visual and/or auditory warning signals. One factor that should be taken into account in interpreting early research is a potential novelty effect of active tactile displays. This effect may have affected experimental results, because this novelty increases the attention paid to tactile signals. However, now that tactile displays have become common in mobile devices, gaming pads, and other devices, recent work confirms earlier findings.

As with audio displays, it is important to distinguish between nondirectional and directional (spatial) tactile displays. Nondirectional tactile displays, such as those found in pagers and mobile phones, can consist of only one vibrator that delivers one or more vibrational pulses. For example, the stick shaker in a cockpit warns pilots that the aircraft is in danger of stalling. Several authors provide experimental evidence that indicates that nondirectional tactile displays are well suited to detect time critical events such as collision avoidance warnings (Tan et al., 2003; Martens and Van Winsum, 2001). Sklar and Sarter (1999) demonstrated that tactile cues are more effective than visual cues for indicating unexpected changes in status. Scott and Gray (2008) compared response times for tactile, visual, and auditory rear-end collision warnings in a driving simulator. They found that drivers using a tactile display had the shortest response time, and concluded that nondirectional tactile warnings show promise as effective rear-end collision