Multimodal Cueing: The Relative Benefits of the Auditory, Visual, and Tactile Channels in Complex Environments

Carryl L. Baldwin[1], Charles Spence[2], James P. Bliss[3], J. Christopher Brill[3], Michael S. Wogalter[4], Christopher B. Mayhorn[4], and Thomas K. Ferris[5]

[1]George Mason University, Fairfax, VA
[2]University of Oxford, Oxford, UK
[3]Old Dominion University, Norfolk, VA
[4]North Carolina State University, Raleigh, NC
[5]Texas A&M University, College Station, TX

Abstract

Determining the most effective modality or combination of modalities for presenting time sensitive information to operators in complex environments is critical to effective display design. This panel of display design experts will briefly review the most important empirical research regarding the key issues to be considered including the temporal demands of the situation, the complexity of the information to be presented, and issues of information reliability and trust. Included in the discussion will be a focus on the relative benefits and potential costs of providing information in one modality versus another and under what conditions it may be preferable to use a multisensory display. Key issues to be discussed among panelists and audience members will be the implications of the existing knowledge for facilitating the design of alerts and warnings in complex environments such as aviation, driving, medicine and educational settings.

INTRODUCTION

It is currently possible to provide information to pilots, drivers, and other professional operators simultaneously or concurrently in the visual, auditory, and tactile modalities. While this capability presents many opportunities for keeping operators informed, engaged, and even entertained, it also presents potential drawbacks. Excessive information may be distracting, confusing, and may result in excessive mental workload. In complex, safety critical situations the effective use of display modality is essential.

Providing too much information may lead to confusion and distraction (Fitch, Hankey, Kleiner, & Dingus, 2011; Meredith & Edworthy, 1995; Patterson, 1990). Identifying the appropriate match between display modality and situation involves resolving a number of key issues. The latest research in this area and key current issues will be highlighted and discussed.

This panel brings together leading international experts in the use of visual, auditory, and tactile modalities and will examine the key issues that must be considered when choosing which modality or combinations of modalities to utilize. These issues along with the relative benefits of each modality will be briefly presented and then ample time for audience discussion and interaction will be provided.

Multisensory warning signal design – Insights from cognitive neuroscience – Keynote

(Charles Spence)

The last decade or two has witnessed impressive cognitive neuroscience insights in terms of our growing understanding of multisensory information processing in humans. I will review evidence from multiple laboratory and driving simulator studies showing that multisensory warning signals, while not necessarily being any more effective than unimodal cues under conditions of low perceptual load, retain their capacity to capture an operator's spatial attention under high load conditions (Ho & Spence, 2008; Spence, 2010). This, at least, is the result if the component unisensory signals are presented from the same spatial location (or at least from the same direction). Cucially, though, multisensory warning signals lose their attentional capture advantage if the component unisensory cues are presented from different spatial locations/directions (Ho, Santangelo, & Spence, 2009), as tends to be the case in many real world multisensory warning systems.

I will also highlight recent research showing how warning signals presented from near rear peripersonal space (i.e., the space just behind a driver's head) can be particularly effective in terms of automatically breaking through and eliciting orienting responses (see Ho & Spence, 2009). Taken together with a number of other recent findings, I will make the case that contemporary cognitive neuroscience research has a great deal to offer in terms of enhancing the design of both unisensory and multisensory warning signals in the coming years (Spence, in press).

The Connection between Emergency Signal Multimodality and Reliability (James P. Bliss)

On the basis of the available research literature, there is a clear distinction between multimodal problem detection and multimodal operator notification. Published systems for multimodal problem detection include a recent system for seizure detection in epileptics (Conradsen, Beniczky, Wolf, Terney, Sans, & Sorensen, 2009) and a surveillance system that combines video sources with auditory sources (e.g., breaking glass, screaming) to produce more reliable signals (Dedeoglu, Toreyin, Gudukbay, & Catin, 2008). The logic behind multimodal signals for notifying operators is based on the fact that operators frequently process a variety of data streams to ascertain the existence of a problem. Alarm respondents frequently rely on the presence of redundant or additional information to confirm signal validity (Bliss, 2003) and first responders seem to prefer alarms that engage multiple modalities (Herring & Hallbeck, 2010). As noted by Herring and Hallbeck, highly consequential task environments such as radiation detection benefit greatly from multimodal alarm presentation but also suffer greatly from a lack of signal reliability. Therefore, studying perceived reliability of multimodal signals is prudent, even if there is still much to understand about the perceived reliability of unimodal signals.

Of course, a lack of understanding has not prevented designers from creating multimodal signaling systems; many exist and are preferable in the minds of operators (Herring et al., 2010). Some are “redundant” systems, where a combination of visual, auditory, olfactory, tactile (and in some cases even gustatory) stimuli occur together in time. Examples include aviation cockpit systems where linked auditory and visual signals occur together; cell phones, where certain auditory signals are accompanied by a visual display; or neutron (radiation) detectors, where visual and tactile signals co-occur. Other systems are “user selectable”, where control of multimodality is granted to the human operator. Cellular telephones fall into this category (tactile vibration and auditory stimuli are controllable). Our own results have suggested that visual signals evoke faster reactions, whereas accuracy rates did not change across modalities (Bliss, Liebman, & Brill, 2012). Given the variety of designed and naturally occurring multimodal systems, however, several questions remain open for discussion. Many of these questions will be discussed in this panel.

Advantages of Multimodal Displays
(J. Christopher Brill)

Multimodal displays offer several potential advantages over unimodal displays. Shifting information from a visual display to an alternative display modality can reduce the processing demands on visual attention while mitigating structural interference, the limitation associated with monitoring multiple displays with a visual system capable of only a single focal point (Wickens, 1984, 2002). Moreover, situational or environmental constraints may preclude the usage of a particular display modality, thereby requiring multimodal adaptation. For example, visual displays are undesirable during covert night operations, as they could give away one’s position. Auditory displays may be ineffective in the presence of noise, particularly if they require intact binaural hearing for the accurate perception of spatial cues. Vibrotactile display effectiveness may be compromised by low-frequency whole body vibration, such as while driving in rough terrain. Each of these scenarios poses a situation in which migration of information to alternative display modalities would yield performance benefits.

Although situational constraints may help guide the selection of display modality, additional factors must be considered, such as the attentional demands and information-processing costs of multimodal displays, as well as response competition for multimodal signals (Brill & Ferguson, in press). The data suggest that vibrotactile cues can serve as highly intuitive signals imposing lower information-processing demands, facilitating faster response times, and yielding less subjective workload, as compared to spatial auditory signals, and in some circumstances, visual signals. However, when multimodal cues are implemented poorly (e.g., low stimulus-response compatibility), the performance “costs” are the greatest for vision and touch. Additional factors related to the implementation of multimodal displays will be discussed, including redundancy gain and the impact of multimodal distractors on performance.

More is better than one: Research and an application for warning about gas leaks
(Michael S. Wogalter)

Research generally indicates that warnings presented both visually and auditorily are more effective (e.g., detection, compliance) than warnings given through either modality alone (Cohen, Cohen, Mendat,

& Wogalter, 2006). Multimodal warnings provide a way to alert people who are occupied with a task involving one or the other modality. The presented information can be the roughly the same in both modalities (e.g., same words in speech and print) or can be different for each modality (auditory alert combined with more extensive information given visually). The combination of modalities frequently shows linear additive effects suggesting separate benefits of modality specific information. Task complexity also plays a role (e.g., Cohen et al., 2006).

Multimodal Gas Leaks. Beyond the main modalities of vision and audition, other modalities may be incorporated into warnings. One is olfactory. Some lessons can be learned with respect to the warnings involved with natural and propane gas delivered to homes and businesses. Gas leaks can lead to explosions and fires. Both types of gases are odorless. Commonly added to these gases before delivery to end-users is a chemical odorant with a skunk-like smell (usually comprised of mercaptan compounds). Gas companies rely on people smelling leaked gas as a signal to evacuate the area and avoid the danger. However, people may not smell the odorized gas for a variety of reasons (e.g., dispositional, congestion, being asleep; see Wogalter & Laughery, 2011). One practical method is to inform people about alternative ways that they might detect a gas leak. One currently available way to detect gas leaks is to use electronic gas detectors. Like smoke detectors, they produce an auditory signal when the gas is "sensed" by the device. Unfortunately most natural and propane gas users do not have gas detectors (Kim & Wogalter, 2012). Also information on cues from other modalities should be communicated to consumers: hearing the sound of gas escaping, and/or visually seeing bubbling, and discolorations.

Key Issues for Warnings (Christopher B. Mayhorn)

The vast majority of warnings are communicated via the visual modality as words or symbols and via the auditory modality as sounds or verbal messages (Wogalter, 2006). If such warnings are considered as part of a warnings system, the nature of the target audience and the environment where the warning will be received must be considered. Some populations, such as hearing impaired individuals, possess personal characteristics that preclude the use of auditory warnings. Likewise, vision impaired individuals are not likely to benefit from hazard information that is communicated visually. In other instances, some populations such as older adults suffer from degradation of the visual and auditory channels (see Kline & Scialfa, 1997). These issues may be complicated further by the nature of the environment.

As will be discussed in the panel, it is understood that warnings can include multiple sensory modalities but each modality has important characteristics (in the absence of the aforementioned situational factors) that must be considered during warning design. For instance, previous research suggests that visual warnings are important in certain complex situations where people need repeated or continual access to safety information (Barlow & Wogalter, 1993). In the case of a complex multi-step safety procedure where working memory might be taxed, the ability to re-read information on a printed warning is essential. By contrast, an auditory warning in the form of speech is fleeting in nature and may possibly be unintelligible due to interference from background environmental noise(Edworthy & Hellier, 2000). Such situations could conceivably require users to wait for the message to be repeated and this could delay warning compliance.

In other situations, where the temporal demands of a warning require immediate attention, auditory warnings are often recognized for being dynamic and omni-directional (Wogalter & Mayhorn, 2005). Where someone must actively focus attention in the general area where a visual warning is located to be alerted to its presence, an auditory warning effectively serves to direct attention to itself. In situations that require the use of a visual warning, it must be noted that visual warnings that include symbols are more likely than those that do not contain symbols to capture attention (Laughery, Young, Vaubel, & Brelsford, 1993). Likewise, the use of other visual design factors such as color might further enhance warning conspicuity. Previous research within the HF/E warning literature clearly indicates that colors such as red and yellow suggest greater levels of hazard than other common colors such as green and blue (Braun & Silver, 1995; Chapanis, 1994).

While the choice of warning modality is often complex, it is generally recognized that the use of multiple modalities to disseminate warnings is often beneficial because the hazard information is presented redundantly (Paivio, 1990). Several studies indicate that the use of two modalities to deliver a warning is more effective than using a single modality (Barlow & Wogalter, 1993; Wogalter & Young, 1991) perhaps because people may believe that the warning is important if it is being disseminated in more than one modality. Interestingly, there is also evidence that multimodal warning is important in vehicles because of the high risk of driver distraction (Lerner, Kotwal, Lyons, & Gardner-Bonneau, 1996).

Tactile displays: supporting simple and complex communications via the sense of touch
(Thomas K. Ferris)

The sense of touch offers a number of display advantages for use in complex environments. First, in most domains of interest, touch is often generally less involved in task-related processing than vision or audition; so interpreting tactile cues may interfere less with ongoing task demands. It also offers a unique combination of display affordances in that the signal is obligatory (i.e., not easily ignored), capable of being sensed regardless of the orientation of attention, and supports privatized messaging since tactile signals cannot be seen or heard (or easily felt) by others.

While several dimensions of touch have been explored for communicating encoded messages – including pressure, texture, temperature, and even pain – vibrotactile presentations (coded vibrations) support the greatest amount of expressiveness with current technologies, and have been emphasized to date in tactile display designs. The simplest forms of vibrotactile messaging are cues or notifications corresponding to a binary event, for example, when a cell phone/pager vibrates to announce the presence of an incoming call. Researchers have found simple vibrotactile cues can be very effective for communicating a state change in automated systems, for example, mode changes in flight deck automation (Sklar & Sarter, 1999). By modulating one or multiple dimensions of a vibrotactile signal, more complex communications can be achieved. The spatial dimension can be used to direct attention to different areas of space by presenting vibrations to different locations on the body, which is an effective technique when there are many relevant visual data sources, such as in the driving environment (e.g., Ho, Tan, & Spence, 2005). Body locations can also be used to naturally associate the display signal with the represented data, for example, by presenting coded vibrations representing patient blood pressure data on an anesthesiologist's upper arm, in the same location where a blood pressure cuff is commonly applied (Ferris & Sarter, 2011). Other dimensions that have been used to encode data include the frequency and gain/intensity of vibration, waveform, and temporal dynamics such as rhythm. Iconic patterns that modulate one or more of these dimensions – "tactons" (e.g., Brewster & Brown, 2004) or "haptic icons" (e.g., MacLean & Enriquez, 2003) – have been effectively demonstrated in a number of HCI and interpersonal communication applications.

As more information is encoded into a vibrotactile display signal, several considerations must be made for which and how many dimensions to utilize. First, tactile signals are subject to masking effects and forms of change blindness in the spatial dimension (Gallace, Tan, & Spence, 2006), and also in nonspatial dimensions such as intensity (Ferris, Stringfield, & Sarter, 2010), which can inhibit the signals from being reliably perceived. Cognitive limitations in interpreting the signal must also be considered in the choice of encoding method. For example, the spatial and symbolic working memory demands of concurrent tasks can greatly impact the interpretability of tactile messages when the same cognitive resources are required to decode the messages (Ferris & Sarter, 2010). Even without interference from concurrent demands, the limited perceptual resolution and processing bandwidth of the tactile channel require that display signals remain relatively simple, compared to auditory or visual displays. However, when display signals are well-mapped to the represented data and their complexity is properly managed within a given set of concurrent task demands, tactile displays offer much promise for supporting communications with a human operator in complex environments.

Summary. The panel of design experts, in conjunction with key note speaker, Charles Spence, and audience participation will examine the relative benefits of each signal modality for operator cueing in complex environments and highlight areas where additional research is needed.

References

Barlow, T., & Wogalter, M. S. (1993). Alcoholic beverage warnings in magazine and television advertisements. *Journal of Consumer Research, 20*, 156-174.

Bliss, J.P., Liebman, R., & Brill, J.C. (2012). Alert characteristics and identification of avatars on a virtual battlefield. Intelligent Decision Technologies, 6(2), 151-159.

Braun, C. C., & Silver , N. C. (1995). Interaction of signal word and colour on warning labels: Differences in perceived hazard and behavioural compliance. *Ergonomics, 38* (11), 2207-2220

Brewster, S.A., & Brown, L.M. (2004). Tactons: Structured tactile messages for non-visual information display. Proceedings of the 5th Australasian User Interface Conference, 15-23.

Brill, J. C., & Ferguson, E. L. (in press). Performance differences on a multimodal target selection task. Abstracts of the Annual Meeting of the American Psychological Association, Orlando, FL.Chapanis, A. (1994). Hazards associated with three signal words and four colours on warning signs. Ergonomics, 37(2), 265-275.

Cohen, H. H., Cohen, J., Mendat, C. C., & Wogalter, M. S. (2006). Warning channel: Modality and media. In M. S. Wogalter (Ed.) *Handbook of Warnings* (pp. 123-134). Mahwah, NJ: Lawrence Erlbaum Associates.

Conradsen, Beniczky, Wolf, Terney, Sans, & Sorensen (2009). Multimodal intelligent seizure acquisition (MISA) system: A new approach towards seizure detection based on full body motion measures. Proceedings of the IEEE Engineering Medicine and Biological Sciences Conference.

Dedeoglu, Y., Toreyin, B.U., Gudukbay, U. & Catin, A.E. (2008). Surveillance using both video and audio. In P. Maragos et al. (Eds), *Multimodal Processing and Interaction*. New York: Springer.

Edworthy, J., & Hellier, E. J. (2000). Auditory warnings in noisy environments. Noise and Health, 6, 27-39.

Ferris, T.K. & Sarter, N. (2010). When content matters: The role of processing code in tactile display design. IEEE Transactions on Haptics, 3(3), 199-210.

Ferris, T.K. & Sarter, N. (2011). Continuously informing vibrotactile displays in support of attention management and multitasking in anesthesiology. Human Factors, 53(6), 600-611.

Ferris, T., Stringfield, K., & Sarter, N. (2010). Tactile "change blindness" in the detection of vibration intensity. Proc HFES 54th Annual Meeting, 1316-1320.

Fitch, G. M., Hankey, J. M., Kleiner, B. M., & Dingus, T. A. (2011). Driver comprehension of multiple haptic seat alerts intended for use in an integrated collision avoidance system. *Transportation Research Part F-Traffic Psychology and Behaviour, 14*(4), 278-290Gallace, A., Tan, H.Z., & Spence, C. (2006). The failure to detect tactile change: A tactile analogue of visual change blindness. Psychonomic Bulletin & Review, 13(2), 300-303.

Herring, S., & Hallbeck, S. (2010). Conceptual design of a wearable radiation detector alarm system: A review of the literature. *Theoretical Issues in Ergonomics Science, 11*(3), 197-219.Ho, C., Santangelo, V., & Spence, C. (2009). Multisensory warning signals: When spatial correspondence matters. Experimental Brain Research, 195, 261-272.

Ho, C., & Spence, C. (2008). The multisensory driver: Implications for ergonomic car interface design. Aldershot: Ashgate Publishing.

Ho, C., & Spence, C. (2009). Using peripersonal warning signals to orient a driver's gaze. Human Factors, 51, 539-556.

Ho, C., Tan, H. Z., & Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. Transportation Research Part F: Traffic Psychology and Behavior, 8(6), 397-412.

Kim, S., & Wogalter, M. S. (2012). Human factors considerations for a warning system to detect gas leaks. *Proceedings of the Human Factors and Ergonomics Society*, *56,* in press.

Kline, D. W., & Scialfa, C. T. (1997). Sensory and perceptual functioning: Basic research and human factors implications. In A. D. Fisk and W. A. Rogers (Eds.), Handbook of human factors and the older adult (pp. 327-328). San Diego: Academic Press.

Laughery, K. R., Young, S. L., Vaubel, K. P., & Brelsford, J. W. (1993). The noticeability of warnings on alcoholic beverage containers. Journal of Public Policy and Marketing, 12, 38-56.

Lerner, N. D., Kotwal, B. M., Lyons, R. D., & Gardner-Bonneau, D. J. (1996). Preliminary human factors guidelines for crash avoidance warning devices. (Report DOT HS 808 342). Washington D.C.: National Highway Traffic Safety Administration, Office of Crash Avoidance Research.

MacLean, K. & Enriquez, M. (2003). Perceptual design of haptic icons. Proceedings of Eurohaptics '03, 351-363.

Meredith, C. & Edworthy, J. B. (1995). Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of Advanced Nursing January, 21*(1), 15-20.

Paivio, A. (1990). *Mental representations: A dual coding approach*. New York: Oxford University Press

Patterson, R. D. (1990). Auditory warning sounds in the work environment. In D. E. Broadbent, J. T. Reason & A. D. Baddeley (Eds.), *Human factors in hazardous situations.* (pp. 37-44): Clarendon Press/Oxford University Press.

Sklar, A. E. & Sarter, N. B. (1999). Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. Human Factors, 41(4), 543-552.

Spence, C. (2010). Crossmodal spatial attention. Annals of the New York Academy of Science, 1191, 182-200.

Spence, C. (in press). Drive safely with neuroergonomics. The Psychologist.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), Varieties of attention (pp. 63-101). New York: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomic Science, 3(2), 159-177.

Wogalter, M. S. (2006). Handbook of warnings. Mahwah, NJ: Lawrence Erlbaum Associates.

Wogalter, M. S., & Laughery, K. R. (2011). Failure to detect gas leaks: Forensic human factors considerations. *Ergonomics in Design*, January, 21-23

Wogalter, M. S., & Mayhorn, C. B. (2005). Providing cognitive support with technology-based warning systems. Ergonomics, 48(5), 522-533.

Wogalter, M. S. & Young, S. L. (1991). Behavioral compliance to voice and print warnings. Ergonomics, 34, 79-89.