Specific Evacuation Instructions Enhance Spoken Fire Warnings

Jesseca R. I. Taylor and Michael S. Wogalter
Department of Psychology
North Carolina State University
Raleigh, North Carolina 27695-7650 USA

Most building fire alarms are presented auditorily and almost all are simple non-verbal sounds, such as bells and buzzers—despite the availability of inexpensive, feasible, voice-presentation technology. One advantage of speech or vocal signals is that they can convey specific avoidance instructions. This study examined the content (wording) of spoken fire alarms. Sixty-five participants rated 90 spoken fire evacuation warnings plus two non-speech sounds (white noise and a recording of an actual [simple, nonverbal] fire alarm) on a 9-point scale (0=not at all, 8=extremely) on their acceptability as a building fire alarm. Significantly higher ratings were assigned to spoken than to non-spoken warnings, except the recorded fire alarm had significantly higher ratings compared to the shortest speech warning ("Fire, Fire"). Speech warnings that provided evacuation instructions (e.g., "Use Stairs") and communicated the need to evacuate right away (e.g., "Exit Now") were rated significantly higher than those lacking that information. Implications of these results are discussed.

INTRODUCTION

In the past three decades, a large body of research has been conducted on warnings. Most research has focused on variables associated with visual warnings that benefit or detract from their effectiveness on various measures (see e.g., Laughery, 2006). There has been relatively less research on warnings in other modalities such as auditory warnings. Most auditory warning research has been involved in examining one or more different types of simple auditory sounds such as bells, buzzer alarms, etc. There has been less research on speech (spoken, voiced, vocal, oral, aural—all are used interchangeably) warnings despite the technology being available to use it.

There are several advantages of auditory warnings over visual warnings. Auditory warnings are omnidirectional. Receivers do not have to have their heads directed toward the source to hear the warning, unlike visual warnings in which the eyes' focus must be oriented on a visual stimulus. In situations that involve complex visual tasks with high workload, they might not notice a visual warning. Auditory warnings might be more effective in those situations because they would be less competition for limited capacity during a high visual workload task (Kahneman, 1973; Wickens, 1984).

There are several types of auditory warnings. Auditory warnings can be categorized as non-speech versus speech. Further, non-speech sounds can be simple or complex. Simple non-speech warnings are most typical for auditory warnings, e.g., beeps, buzzers, bells, horns, etc. They are frequently used as indicators, i.e., the possible presence of something but they lack specific information. Complex non-speech sounds can be multifaceted in form (varying over time in frequency, amplitude, and timbre) and thus could potentially convey more information than a simple indicator-type sound. Complex non-speech sounds can convey or cue more information when the coded sounds become associated with specific kinds of events. For example in most large aircraft, different nonverbal sounds are coded to convey different events, such as a flight attendant call or left engine fire. However, complex auditory warnings require specific training so that the sounds cue their intended meaning. Pilots need to be retrained every so often so they will not forget the meaning of less frequently presented sounds such as the one assigned to left engine fire.

In the last several decades, technology has enabled the use of digitized speech warnings. The cost of voice chips is inexpensive (e.g., used in greeting cards and answering machines) and could be adapted to deliver warning information (Conzola & Wogalter, 1999). With spoken warnings, extensive training is not needed because people can use pre-existing knowledge of language to comprehend the spoken messages. Speech warnings should not be excessively lengthy as they generally take more time to transmit than people can read the same information in print. As a consequence, lengthy speech warnings could delay or reduce comprehension and postpone potentially necessary emergency responses. Duration is not the only consideration in choosing speech warnings. It should be balanced with content. Good warnings have all the safety information necessary and nothing extraneous. Thus, while brevity is an important concern, the message needs to be long enough to provide the most important content information (Wogalter, DeJoy, & Laughery, 1999).

One kind of emergency in which auditory warnings are used is for fire. Karter (2014) reported there were 3,240 civilian deaths and 15,925 civilian injuries of 1,240,000 fires reported in 2013 in the U.S. Most fire warning systems typically use simple, loud auditory warnings. Data show that such alarms are useful. Ahrens (2014) reported that the home-fire death rate was at least two times lower in homes with smoke alarms compared to death rates in homes without or with an inoperable smoke alarm. These and other fire prevention systems are beneficial but have not completely prevented fire injury and damage.

One disadvantage of conventional fire warnings is the need to identify the sound as being a fire alarm. There are other simple, loud alarms used in emergencies such as the detection of carbon monoxide and leakage of explosive gas. Also, conventional fire alarms do not communicate other potentially important information such as evacuation instructions. During an emergency situation, people's attentional processing can be restricted or narrowed and could result in a failure to process all of the pertinent cues. Almost all adults know to avoid using elevators during building fires and could say so if surveyed about their knowledge on what to do in case of a fire in multi-story building. But in actual emergency situations, relevant knowledge in long-term memory may not be cued into awareness. Improper use of elevators in a fire emergency might still occur due to habitual use of elevators for egress. Speech fire warnings could be used to announce (and cue) the method of evacuation to avoid the elevators and use the stairs during a fire egress situation.

Fire evacuation instructions are usually visually posted on signs near elevators and stairways, but their content has received little research. An exception is a study by Taylor and Wogalter (2012). They manipulated the content of visually-presented fire warnings and asked participants to use a scale to rate the warnings on their acceptability. Their results showed that warnings with specific visual instructions, e.g., to use the stairs and not the elevator, produced higher ratings than warnings without that information. However, the warnings in that study were only presented in the visual (print) modality. No study has systematically manipulated spoken fire warnings. This was a purpose of the present research.

In the current study, participants were presented with both speech and non-speech fire warnings. The speech warnings were comprised of short phrases such as “Fire, Fire” or also included various egress immediacy phrases (e.g., exit now) and directives (e.g., use stairs). It was expected that the speech warnings would be rated higher than nonverbal sounds (conventional fire alarm and white noise) and speech warnings with egress immediacy information and/or egress directives would be rated higher than speech warnings without that information. The order or sequencing of the phrase content was also examined.

METHOD

Participants

The participants were 31 male (47.7%) and 34 female (52.3%) undergraduates from a psychology course at North Carolina State University with ages ranging from 18 to 25 years (M = 18.7, SD = 1.33). Three people had mild hearing impairments according to the scoring criteria of the Hearing Screen Inventory (Coren & Hakstian, 1992). Each participant received research credit for taking part in the study.

Materials

*Preparation of the voiced fire-warning stimuli***.** Recordings from six native English speakers (3 males and 3 females) were used as the spoken fire warnings in the main experiment. The speakers were recruited from the same pool used for the main experiment but none participated in both parts. Use of multiple speakers was intended to promote generalizability. Speakers participated individually in the recording sessions. The session began with voice training. The experimenter instructed speakers to enunciate the warnings in a way to capture the attention of building occupants. Speakers were instructed to say the warnings out loud quickly but also to distinctly enunciate all of the words. Speakers practiced vocalizing the warnings into a microphone until they maintained the sound level at a fairly consistent and approximate level of 80 dB (measured by a sound level meter) and could produce recordings without distortion (i.e., no clipping shown on recording software). Recording sessions were held in a small, quiet room with approximately 40 db of background noise. Audacity sound editing (open source) software was used to collect the recordings.

The warning statements were randomized for each speaker such that no speaker vocalized them in the same order. Speakers spoke all 15 fire-warning statements twice. The second recording was a backup recording. Each speaker received a card stack of the entire set of fire-warning statements with each card having a different printed statement according to a predetermined random order. The experimenter then re-arranged the cards into another pre-determined random order for the second recording session.

Main experiment. There were 90 spoken warning statements (15 statements from each of 6 speakers) and two non-speech sounds, i.e., 92 separate items. One non-speech sound was a short audio clipping of a fire alarm (3.16 s, ~80 dB) produced by a Simplex 4100U Addressable Fire Alarm system to represent a conventional buzzer-type fire alarm. The other sound was white noise (1.27 s, ~80 dB), which was generated using Audacity software. White noise served as a control condition. The warnings used in the experiment are shown in Table 1.

For the main experiment, five different random orders of the spoken warnings (and sounds) were produced. Each was presented in a playlist to 13 participants using VLC Media player. The playlists were organized such that 7 seconds of silence separated each warning to provide time for participants to write down their ratings.

Procedure

After participants signed a consent form, the experimenter read aloud a set of instructions to participants. Within these instructions was a fire-alarm scenario, in which participants were asked to imagine that while working inside an office building they heard a loud alarm sound. They were told to assume that fire alarms could take different forms and that they would be presented a series of auditory sounds or statements and imagine each being played over a loud speaker as the fire alarm. Participants donned a set of Sony MDR-XD100 stereo over-the-ear headphones to control ambient noise and sound level. Participants rated each warning on how acceptable it would be as a fire alarm. This was intended to measure beliefs on the overall suitability of the fire warnings. In the current study, participants were asked to think of acceptability as a combination of several measures such as appropriateness. The scale for acceptability is described in Taylor and Wogalter (2012). After rating the warnings, participants completed surveys for hearing loss and demographics.

RESULTS

The means for the spoken fire warnings were collapsed across speakers. Table 1 shows the mean acceptability ratings and standard deviations for the speech and non-speech warnings. A one-way repeated measures ANOVA was significant, $F(3.99, 255.30) = 37.65$, $MSE = 5.09$, $p < .001$, $\eta_p^2 = .37$. Post-hoc tests using a Bonferroni correction were conducted. The fire alarm (non-speech warning) ($M = 4.52$, $SD = 2.08$) had a significantly higher mean rating than the shortest spoken warning "Fire, Fire" ($M = 2.71$, $SD = 1.78$) and the white noise warning ($M = 1.66$, $SD = 2.41$).

Table 1.
Mean acceptability (and standard deviations) as a function of warning statement condition in descending order.

Warning condition	*M* (*SD*)
Fire, Fire, Exit Immediately, Do Not Use Elevator, Use Stairs	5.16 (1.33)
Fire, Fire, Exit Immediately, Use Stairs, Do Not Use Elevator	5.04 (1.40)
Fire, Fire, Exit Now, Do Not Use Elevator, Use Stairs	4.85 (1.41)
Fire, Fire, Exit Now, Use Stairs, Do Not Use Elevator	4.70 (1.41)
Fire, Fire, Use Stairs, Do Not Use Elevator	4.59 (1.22)
<Non-speech fire alarm sound>	4.52 (2.08)
Fire, Fire, Exit Immediately, Use Stairs	4.49 (1.22)
Fire, Fire, Exit Now, Use Stairs	4.41 (1.26)
Fire, Fire, Exit Now, Do Not Use Elevator	4.33 (1.17)
Fire, Fire, Exit Immediately, Do Not Use Elevator	4.33 (1.13)
Fire, Fire, Do Not Use Elevator, Exit Now	4.29 (1.23)
Fire, Fire, Use Stairs	4.03 (1.39)
Fire, Fire, Exit Now	4.01 (1.42)
Fire, Fire, Exit Immediately	3.85 (1.14)
Fire, Fire, Do Not Use Elevator	3.74 (1.28)
Fire, Fire	2.71 (1.78)
<White noise sound>	1.66 (2.41)

The non-speech fire alarm also received a higher mean rating than some of the spoken fire warnings but the differences were not significant. White noise was rated significantly lower than all other warnings, falling between “not at all acceptable” and “somewhat acceptable” according to the scale anchors. White noise also had the highest standard deviation of all conditions. Other than white noise, all other warnings’ ratings were rated at least “somewhat acceptable” according to the scale anchors. As can be seen in this table, the longer statements containing egress immediacy and directives were rated highest.

Several additional analyses were conducted on subsets of statements that formed factorial designs. These analyses were intended to examine the effects of content components and whether parts or factors interact. A 3 (Egress Immediacy: <none>, “Exit Now,” “Exit Immediately") X 2 (Includes Stairs Directive: no, yes) X 2 (Includes Elevator Directive: no, yes) factorial repeated measures ANOVA was conducted. Table 2 shows the means and standard deviations as a function of the above-mentioned warning-content factors.

There was a significant main effect of Egress Immediacy, $F(1.56, 99.99) = 34.59$, $MSE = 1.34$, $p < .001$, $\eta_p^2 = .35$. Post-hoc tests using a Bonferroni correction showed that warnings with the immediacy phrases "Exit Now" ($M = 4.38$) or "Exit Immediately" ($M = 4.45$) were rated higher than warnings that lacked an immediacy phrase ($M = 3.77$). No significant difference was found between the two immediacy phrases (“now” versus “immediately”).

Table 2.
Mean Acceptability Rating (and standard deviations) as a Function of Egress Immediacy, Stairs, and Elevator Phrases.

Egress Immediacy		Includes Stairs Directive: No	Yes	Overall Mean
None	No Elevator Directive	2.71 (1.78)	4.03 (1.39)	3.37
	Elevator Directive	3.74 (1.28)	4.59 (1.22)	4.16
Exit Now	No Elevator Directive	4.01 (1.42)	4.41 (1.26)	4.21
	Elevator Directive	4.31 (1.09)	4.77 (1.36)	4.54
Exit Immediately	No Elevator Directive	3.85 (1.14)	4.49 (1.22)	4.17
	Elevator Directive	4.33 (1.13)	5.10 (1.27)	4.72
Overall Mean		3.83	4.57	

There was significant main effect of the Stairs Directive, $F(1, 64) = 73.65$, $MSE = 1.45$, $p < .001$, $\eta_p^2 = .54$. Post-hoc tests using a Bonferroni correction showed that warnings with the "Use Stairs" ($M = 4.57$) directive were rated significantly higher than those without the stairs directive ($M = 3.83$). There was significant main effect of the Elevator Directive, $F(1, 64) = 25.12$, $MSE = 2.42$, $p < .001$, $\eta_p^2 = .28$. Post-hoc tests using a Bonferroni correction showed warnings with the "Do Not Use Elevator" ($M = 4.47$) directive were rated higher than those with no elevator directive ($M = 3.92$).

There were also several significant interactions. One was an interaction of egress immediacy and stairs directive, $F(2, 128) = 21.03$, $MSE = .34$, $p < .001$, $\eta_p^2 = .25$. Figure 1 shows a bar graph of this interaction. When a stairs directive was present, warnings with “Exit Immediately” produced significantly higher ratings than warnings with “Exit Now” and warnings with no immediacy phrase, the latter of which did not differ. However, when the stairs directive was absent, there was no difference between the two immediacy phrases, but “Exit Now” was significantly higher than warnings with no immediacy phrase. Also notable in this graph is the low mean rating when both the stairs and immediacy phrases were absent.

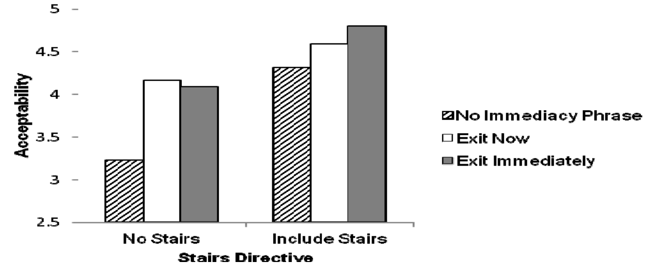


Figure 1. Egress Immediacy X Stairs Directive Interaction. Each bar represents a level of the egress immediacy factor.

There was a significant interaction effect between egress immediacy and elevator directive, $F(2, 128) = 11.69$, $MSE = .30$, $p < .001$, $\eta_p^2 = .15$. Figure 2 shows the bar graph of this interaction. The pattern of results is similar to the interaction described above, but in this case, the elevator directive is substituted for the stairs directive. The pattern of these results appeared to show that the warnings with the elevator directive, "Exit Immediately," were rated significantly higher than warnings with "Exit Now," which in turn was significantly higher than warnings with no immediacy phrase. However, for warnings that lacked the elevator directive, there was no difference between "Exit Now" and "Exit Immediately," but both were higher than warnings with no immediacy phrase.

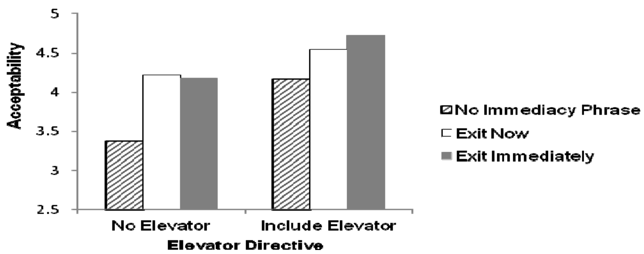
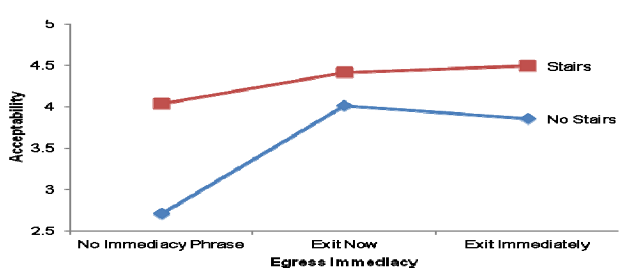


Figure 2. Egress Immediacy X Elevator Directive Interaction. Each bar represents a level of the egress immediacy factor.

There was also a significant 3-factor interaction, $F(1.74, 111.56) = 4.13$, $MSE = .49$, $p = .02$, $\eta_p^2 = .06$. The means associated with this interaction are shown in the two parts of Figure 3 (a & b). Figure 3a shows the means for warnings with no elevator statement and figure 3b shows means for warnings with the elevator phrase present.

(a) NO ELEVATOR PHRASE



(b) ELEVATOR PHRASE INCLUDED

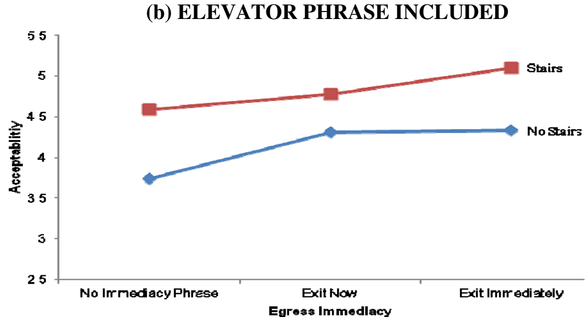


Figure 3 (a and b). Mean acceptability ratings in a significant three-factor interaction shown as a function of two graphs. The top graph (a) shows the egress immediacy and stairs statement when the elevator phrase is absent; the bottom graph (b) shows the same two factors when the elevator phrase is included (present).

The overall pattern of means is the same as already described but it appears that the interaction is being driven by one particularly low mean. It is the condition with the most limited content in the entire set of spoken warnings: "Fire, Fire." The mean this warning is shown as the bottom left point in Figure 3a. Lastly, statements were manipulated with respect to changing the order of component phrases with all content held constant. Comparisons among warnings with the same content, but differently-ordered phrasing yielded no significant differences.

DISCUSSION

In this study, people rated a set of systematically-manipulated voiced fire evacuation warnings and two non-speech sounds on their acceptability as fire alarms. In general, all warnings with the exception of the white noise were rated as being at least somewhat acceptable. The fire alarm sound garnered significantly higher ratings than the shortest and simplest speech warning, "Fire, Fire." This might be due to past experience of participants having heard the nonverbal fire alarm before on the university campus (or elsewhere) during fire drills, and as a result, associated the sound with an actually-used fire emergency alarm. The fire alarm sound probably fits peoples' pre-existing expectations as to how a typical fire warning would sound, and thus resulting in its relatively high mean rating.

The white noise sound was given significantly lower ratings than all other conditions. White noise is frequently used to mask background noise. It is a sound that is not semantically associated with fire emergencies. It also had the highest standard deviation, which suggests some confusion about its appropriateness in this application. It should be noted that the white noise played for a slightly shorter duration than some of the other warnings. An alternative explanation is that this short burst of sound might have startled some participants leading them to give lower acceptability ratings.

As expected, spoken warning content had an effect on ratings. In general, the results appear to support the ideas of brevity and completeness being important in speech warnings. The shortest warning, "Fire, Fire" had the lowest rating of all spoken fire warnings. It did not provide any information other than to identify itself as a fire warning.

Several components of the warning content were manipulated. One was egress immediacy. The results showed that warnings with an egress immediacy phrase were rated more acceptable than those without an immediacy phrase. The inclusion of these phrases probably benefitted ratings because they give direct commands not to delay evacuation. In actual emergency situations, these immediacy phrases could be useful in quick and effective evacuations.

Although the findings show that immediacy information is important to include, there was no difference between the two phrases, "Exit Now" and "Exit Immediately." Both were rated higher than no immediacy phrase. The words *now* and *immediately* are synonyms. This may be a reason why the ratings did not differ. However, "Exit Immediately" is longer and takes more time to say. Although the main effect showed no difference between the two immediacy phrases, there were interaction results suggesting the longer immediacy phrase might be better in certain circumstances (when a stairs or elevator directive is used).

Warnings that included egress directives received higher ratings than those that did not include directives. The instructions to use stairs and/or avoid using the elevator could serve as reminders to people during a fire emergency. Their explicit presentation could increase the likelihood that this information or knowledge will be brought to awareness during an emergency evacuation and potentially "break" the use of automatic behavior of taking an elevator and to use the stairs instead. Of the two egress directives, the stairs directive had a greater effect than the elevator directive and having both in a warning produced the highest ratings. This finding does not support brevity as an exclusive criterion for warnings. On the surface, one could argue that having one or the other directive would suffice. Knowing to avoid elevators would imply to use stairs. However, participants believed that both components were beneficial, possibly because it leaves no ambiguity on what to do and what not to do.

The 3-factor interaction results confirmed what was found in the one-way ANOVA: "Fire, Fire" was rated significantly less acceptable compared to all other spoken fire warnings. The other spoken warnings contained explicit evacuation instructions and/or an immediacy phrase. Again, this result supports the notion that acceptable spoken fire alarms should have a certain amount of completeness in expression as opposed to extreme brevity. However, we did not test extremely long and extraneous content in this study.

As noted previously, speech warnings have several advantages over visual warnings and may be more appropriate for certain applications than visual print warnings (Conzola & Wogalter, 1999). The presence of warnings in both modalities is usually better than either modality alone (Barlow & Wogalter, 1993; Wogalter, Shaver, & Kalsher, 2014).

Relative to complex (coded) non-speech sounds, speech given in a known language generally requires less training—a benefit for warnings in non-occupational settings and for children and illiterates. However, speech will provide little assistance to persons who do not understand the language. Thus, environments that are likely to include persons who do not understand the primary language (e.g., international airports, multi-cultural communities) will need supplemental presentation. Similar to the benefits of multi-modal warnings, mixed auditory warnings (voice and electronic sound) is one kind of supplement, but also text, pictorial symbols, and presentation in more than one language are others. Future research could evaluate the mixing of modality and medium in warnings to varied populations. Additional research on intonation and voicing (cf. Barzegar & Wogalter, 1998) of speech-based fire warnings and their effects on different groups would benefit knowledge in this area.

There are other ways in which research on fire warnings could be extended. The warnings could be rated on dimensions other than acceptability. Acceptability has been used in other research and provides an overall impression or judgment about warnings. However, warning researchers have also used measures such as judgments of urgency and intended carefulness to evaluate auditory warnings. Future studies could include these measures (and others) to determine if a similar pattern of results is found. Deepened knowledge on the relationships among measures is needed.

One limitation of this study was that warnings were presented in a noise-absent environment. During a real fire emergency, the environment may be noisy from people hurrying to evacuate, other speech and other alarms sounding, as well as objects falling and possibly sounds from explosions. If there is a significant overlap in warning and noise characteristics or a high noise-to-warning ratio, then warning information could be masked and unintelligible (see Haas & Edworthy, 2006). Baldwin (2011) manipulated the signal-to-noise (S/N) ratio of spoken collision avoidance warnings. She found that perceived urgency, annoyance, and warning effectiveness increased and response time decreased as the S/N ratio increased. Future studies on spoken fire warnings should consider using S/N ratio as a factor.

There is also a need to examine warnings involved in systems that detect multiple hazards such as smoke, carbon monoxide detectors and explosive gas. Most current systems have alarm sounds that sound similar, i.e., not discriminable. Better warnings would identify the specific danger and give specific instructions relevant to the danger.

REFERENCES

Ahrens, M. (2014). Smoke alarms in U.S. home fires. *National Fire Protection Association*. Retrieved from http://www.nfpa.org

Baldwin, C. L. (2011). Verbal collision avoidance messages during simulated driving: Perceived urgency, alerting effectiveness and annoyance. *Ergonomics, 54*, 328-337. doi: 10.1080/00140139.2011.558634

Barlow, T., & Wogalter, M. S. (1993). Alcoholic beverage warnings in magazine and television advertisements. *Journal of Consumer Research, 20,* 147-156.

Barzegar, R. S., & Wogalter, M. S. (1998). Intended carefulness for voiced warning signal words. *Proceedings of the Human Factors and Ergonomics Society*, *42*, 1068-1072.Conzola, C. V., & Wogalter, M. S. (1999). Using voice and print directives and warnings to supplement product manual instructions. *International Journal of Industrial Ergonomics*, *23*, 549-556.

Coren, S., & Hakstian, A. (1992). The development and cross-validation of a self-report inventory to assess pure-tone threshold hearing sensitivity. *Journal of Speech & Hearing Research*, *35*, 921-928.

Haas, E. & Edworthy, J. (2006). An introduction to auditory warnings and alarms. In M. S. Wogalter (Ed.), *Handbook of Warnings.* (pp. 189-198). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Karter, M. J., Jr. (2014). Fire loss in the United States during 2013. *National Fire Protection Association*. Retrieved from http://www.nfpa.org

Laughery, K. R. (2006). Safety communications: Warnings, *Applied Ergonomics*, 37 (4), 467-478.

Taylor, J. R. I., & Wogalter, M. S. (2012). Acceptability of evacuation instruction fire warnings. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*, 1753-1757.

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

Wogalter, M. S., DeJoy, D. M., & Laughery, K. R. (1999). Organizing framework: A consolidated communication human information processing (C-HIP) model. In M. S. Wogalter, D. M. DeJoy, & K. R. Laughery (Eds.). *Warnings and Risk Communication*. (pp. 15-24). London: Taylor and Francis.

Wogalter, M. S., Shaver, E. F., & Kalsher, M. J. (2014). Effect of presentation modality in direct-to-consumer (DTC) prescription drug television advertisements. *Applied Ergonomics, 45*, 1330-1336.