Applied Ergonomics 80 (2019) 57–66

Contents lists available at ScienceDirect

Applied Ergonomics

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

Specific egress directives enhance print and speech fire warnings

Jesseca R.I. Taylor*, Michael S. Wogalter

North Carolina State University, Psychology Department, Human Factors and Applied Cognition, 640 Poe Hall, Raleigh, NC, 27695-7650, USA

A B S T R A C T

Despite its importance, research examining the effectiveness of different egress warnings in fire emergencies has been sparse. This research examines language-based warnings presented visually in Experiment 1 and spoken format in Experiment 2 on their rated suitability as fire egress warnings. Two types of phrases were manipulated: egress immediacy and egress directives. Phrase ordering was also varied. Both experiments demonstrated a minimal fire warning without egress information is perceived less acceptable than warnings with egress information. The warnings rated as most acceptable contained egress directives indicating how to evacuate safely and that it should be done quickly. Furthermore, analyses on warning length revealed longer, specific warnings were rated higher than shorter warnings. However, data in Experiment 2 suggested the longest statements were rated lower than ones that were somewhat shorter. Results are discussed in terms of application to fire emergency warnings and general warning issues in emergencies.

1. Introduction

When hazards cannot be practically designed out or guarded against, warnings are typically used as a means of controlling hazards to limit potential injury. One hazard that is difficult to fully prevent is fires. In the U.S., fire departments respond to over one million fires every year with roughly half of those fire emergencies resulting in thousands of injuries and deaths (Karter, 2014). Anecdotal and research-based evidence show fire-warning systems are partly successful in expediting fire egress, resulting in less injury and death. For instance, 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 smoke alarms or with inoperable ones.

Tall, multi-story buildings have increased fire dangers than shorter buildings with respect to egress due to more limited escape routes. Virtually all elevators (lifts) in multistory buildings have signs that warn potential users to use the stairs instead of the elevator during a fire emergency. These signs are typically static displays, combining both words and one or more pictographic images. Advances in technology have provided other potential ways to provide warnings. Fire warnings could be displayed differently, such as using increasingly less expensive flat (and curved) displays, LED information displays, and e-paper with the added potential benefit of being able to dynamically present warnings depending on the current or anticipated type and level of danger (e.g., Wogalter and Mayhorn, 2006). Despite the pervasiveness of static elevator signage and the availability of newer technology to deliver information potentially with greater prominence and utility, the specific wording of fire egress signs has received very little attention in the warning literature (e.g., Smith and Wogalter, 2007; Thomas and Bruck, 2010).

Since the mid-1980s there has been remarkable growth of warning research, most of which has been published in the human factors and ergonomics literature (e.g., see Wogalter, 2006). From this research, guidelines can be extracted for warning designers to use. Two goals said to benefit warnings are brevity and completeness. Warning text should be short but also complete. These two guidelines can be in opposition when applied to warning designs (e.g., Laughery and Page-Smith, 2006). A warning used in a fire emergency should be concise so that people can quickly apprehend the hazard and determine what to do to avoid the hazard. However, the message should also be reasonably complete to give necessary information for a safe evacuation. Clearly it would be inappropriate to provide an unnecessarily lengthy message for an emergency situation, but also it would be unsuitable to leave out critical information that people at risk may need.

Although most published research tends to focus on the design of visually presented warnings (e.g., Wogalter et al., 1997), warnings are also frequently presented as auditory messages across many situations and environments (e.g., Laughery et al., 1994). One advantage of auditory warnings over visual warnings is that they are omnidirectional. That is, receivers do not have to orient their head toward the source to hear a warning, unlike a visual warning in which the eyes must be aimed in the general area of the warning to receive it. In situations that involve complex visual tasks with high workload, receivers might not notice a visual warning that they might otherwise see if the visual environment was less complex and the workload is lighter. In such cases, auditory warnings may be beneficial (cf. Kahneman, 1973; Wickens, 1984).

Auditory warnings can be categorized in several ways. One way is to

* Corresponding author.
E-mail addresses: jessecaitaylor@gmail.com (J.R.I. Taylor), wogalterm@gmail.com (M.S. Wogalter).

https://doi.org/10.1016/j.apergo.2019.04.009
Received 8 June 2018; Received in revised form 15 April 2019; Accepted 16 April 2019
Available online 24 May 2019
0003-6870/

distinguish non-speech versus speech sounds (Haas and Edworthy, 2006; Wogalter and Young, 1991). Further, non-speech sounds can be classified as either simple or complex. Most auditory warnings are simple non-speech sounds, e.g., beeps, buzzers, alarms, bells, horns, etc. They are often used as indicators, i.e., signal the presence of something, but they usually lack a clear indication of their meaning. More complex non-speech sounds have been used to indicate different specific meanings. Their waveforms can be multifaceted (varying over time in frequency, amplitude, and timbre), and thus could potentially convey more information than a simple indicator-type sound. A substantial amount of research on non-speech alarms exists, spanning topic areas of alarm recognition (e.g., Edworthy et al., 2013; Edworthy et al., 2017), subjective impressions of alarms (e.g., Guillaume et al., 2003), and performance (e.g., Suied et al., 2008). Although there are numerous complex auditory signals available for use as warnings, the problem with using them is they require training to learn their meanings and retraining for memory maintenance. In some occupational settings, such as cockpits and intensive care units, operators may need to recognize and properly interpret many alarms.

Typical fire warning systems involve simple loud sounds. One disadvantage of this type of warning is that receivers must know, associate and identify it as a fire alarm. Potentially it could be an alarm for something other than fire, e.g., an auditory warning for carbon monoxide or explosive gas. Also, conventional fire alarms do not inherently communicate other potentially useful emergency information such as specific egress instructions that could be important in potentially stressful emergency situations.

Most adults probably know that they and other people should use elevators when evacuating a multi-story building in a fire emergency. However, during a stressful emergency, people may not do what they may “know.” In emergency situations, there is an aroused state that restricts or narrows people's attentional processing, which could result in the failure to process all pertinent cues (Wickens, 1996). Another reason is relevant knowledge in long-term memory (i.e., not to use the elevator) may not be cued into awareness at the relevant time, i.e., during an emergency event. Improper use of an elevator can occur due to habitual use during typical, non-emergency egress.

Advancements in technology have enabled other methods of warning presentation that could be used. Electronic displays could be used instead of typical static signs to use stairs in a fire emergency. Such displays could be changed dynamically to reduce habituation effects, which could mitigate inappropriate use of elevators at critical times. Also, inexpensive methods are available to deliver digitized speech warnings. So instead of typical fire alarms, speech (voice) warnings could help to identify and cure into awareness the purpose of the alarm and the proper method of egress to use the stairs during a fire egress situation.

Like visual warnings, speech warnings should be concise. A long-duration warning takes time to present from the beginning to the end, which may be too long to be useful in emergency situations. Also, a lengthy message could reduce comprehension due to working memory limitations or postpone necessary emergency responses (Wogalter et al., 1999). It should be brief having only essential, priority information to allow safe egress for persons at risk. There has been relatively little research on speech warnings. Most of it has focused on differences conveyed by spoken signal words (e.g., Danger, Caution) (e.g., Barzegar and Wogalter, 1998; Edworthy et al., 2003; Hellier et al., 2002) with a few exceptions (e.g., Barzegar and Wogalter, 2000). Most recent research has concerned factors involved in the delivery of vehicle collision information and navigation systems (e.g., Baldwin, 2011; Bella and Silvestri, 2017; Nees et al., 2016; Wan et al., 2016). However, this research does not systematically examine specific manipulations of content or the issue of length or duration of presentation. Evaluation of visual print warnings has mostly been conducted on single word warnings (e.g., Hellier and Edworthy, 2006; Wogalter and Silver, 1995) with some exceptions (e.g., Wogalter and Barlow, 1990; Wogalter et al., 1987).

The present research involves two experiments examining a set of multi-word fire warnings in which the wording is manipulated. The statements were rated by participants on their perceived suitability or acceptability as fire evacuation warnings. Acceptability was chosen to measure people's overall judgment or belief about the warning stimuli (e.g., Baldwin and Lewis, 2014; Wogalter et al., 2014).

Experiment 1 involved the manipulation of a set of visually-presented (printed text) statements according to phrase content, order, and length. Experiment 2 was similar to Experiment 1, except the warnings were given through voice presentation.

2. Experiment 1

Thirteen visually-presented statements differing in egress immediacy and egress directives were examined on rated acceptability as a fire evacuation warning.

2.1. Method

2.1.1. Participants

One hundred five individuals (58 males, 47 females) from the Raleigh, North Carolina USA metropolitan area participated. The overall mean age of the sample was 28.8 years ($SD = 14.9$). Eighty-three (79%) reported English as their native language. Seventy-two (68.6%) participants were full-time students from psychology courses at North Carolina State University; they received course credit in exchange for participation. The student participants' self-reported ethnic backgrounds were 57 Caucasians, 2 African Americans, 7 Asians, 4 Hispanic/Latinos, 1 Pacific Islander, and 1 Other. Fifty-nine (81.9%) full-time students reported English as their native language. The remaining 33 participants were recruited at a local flea market, and received t-shirts for participation. Their self-reported ethnicities were 22 Caucasians, 2 African Americans, 3 Asians, and 6 Hispanic/Latinos. Twenty-four (72.7%) were native English speakers.

2.1.2. Materials and procedure

Participants completed a questionnaire consisting of a consent form, demographics questions (e.g., gender, ethnicity), and questions on various safety-related topics. The instructions informed participants that the study sought to identify ways to alert people in fire emergencies and get them to respond appropriately. The instructions gave a scenario in which participants were asked to imagine each fire warning was presented during a fire emergency occurring in a multi-story building and to rate each based on their acceptability as a fire warning. In making their ratings, participants were asked to think of acceptability as a combination of several dimensions such as perceived hazardousness, appropriateness, and how careful one would be to follow the statement. Specifically, participants rated each warning using a 9-point rating scale with the following textual anchors: *0 = Not all acceptable*, *2 = Somewhat acceptable*, *4 = Acceptable*, *6 = Very acceptable*, and *8 = Extremely acceptable*. Ratings were written in blanks that corresponded with the printed warnings.

The 13 warnings that were rated are shown in Table 1. All warnings started with the key signal term *Fire*, which was given twice as Fire, Fire. Some warnings included an *egress* immediacy phrase that intended to encourage building occupants to evacuate quickly and when the phrase was present it was either *Exit Now* or *Exit Immediately*. These two statements have the same meaning, i.e., they are synonymous, but differ in letter length. Additionally, some warnings included instructions on how to evacuate; these are egress directive phrases, and when present, had either *Use Stairs* or *Do Not Use Elevator* or had both directives. Two orders of statements were used; one was randomized and the other was the reverse order.

Table 1
Mean acceptability ratings and standard deviations of 13 fire warnings in descending order.

Warning	Mean (*SD*)
Fire, Fire, Exit Now, Use Stairs	5.56 (1.72)
Fire, Fire, Exit Immediately, Use Stairs	5.44 (2.00)
Fire, Fire, Exit Immediately, Do Not Use Elevator	5.27 (2.30)
Fire, Fire, Exit Now, Do Not Use Elevator, Use Stairs	5.10 (2.44)
Fire, Fire, Exit Now, Use Stairs, Do Not Use Elevator	4.98 (2.47)
Fire, Fire, Exit Now, Do Not Use Elevator	4.84 (2.04)
Fire, Fire, Use Stairs, Do Not Use Elevator	4.80 (1.98)
Fire, Fire, Do Not Use Elevator, Exit Now	4.54 (2.02)
Fire, Fire, Use Stairs	4.43 (2.08)
Fire, Fire, Do Not Use Elevator	4.32 (1.84)
Fire, Fire, Exit Immediately	4.21 (2.00)
Fire, Fire, Exit Now	3.99 (2.13)
Fire, Fire	2.39 (2.42)

2.2. Results

Table 1 shows the two highest rated warnings were *Fire, Fire*, *Exit Now*, *Use Stairs* and *Fire, Fire*, *Exit Immediately*, *Use Stairs*. The lowest was *Fire, Fire*. The highest versus the lowest rated warnings differed by 3 points on the rating scale, ranging from "very acceptable" to "somewhat acceptable."

2.2.1. One-way ANOVA

Table 1 shows the mean warning acceptability ratings ordered from highest to lowest. A one-way repeated measures ANOVA showed a significant effect of statement content, $F(4.81, 500.55) = 21.99$, $MSE = 8.01$, $p < .001$, $\eta_p^2 = 0.18$; corrected degrees of freedom were used based on the Greenhouse-Geisser ε of 0.40). Post hoc comparisons were conducted using a Bonferroni correction. *Fire, Fire, Exit Now, Use Stairs*, was rated significantly more acceptable than warnings ranging from *Fire, Fire, Use Stairs, Do Not Use Elevator* to *Fire, Fire* in Table 1. *Fire, Fire, Exit Immediately, Use Stairs* was rated higher than warnings ranging from *Fire, Fire, Do Not Use Elevator, Exit Now* to *Fire, Fire*. *Fire, Fire, Exit Immediately, Do Not Use Elevator* was more acceptable than *Fire, Fire, Do Not Use Elevator, Exit Now* and warnings ranging from *Fire, Fire, Do Not Use Elevator* to *Fire, Fire*. Lastly, *Fire, Fire* was rated significantly less acceptable than all others. Overall, the comparisons show warning statements with an immediacy phrase and a directive phrase were rated as more acceptable than those that included one phrase. The comparisons did not reveal differences among the warnings that had more information (i.e., 3 phrases).

2.2.2. Phrase order

Another analysis involved phrase order. Examined were the ratings for warnings that were identical in content but had a different order of phrases, for example, *Fire, Fire*, *Do Not Use Elevator*, *Exit Now* versus *Fire, Fire*, *Exit Now*, *Do Not Use Elevator*. There were no significant differences in any of the analyses examining phrase ordering.

2.2.3. Two-way factorial ANOVA

Additional analyses were conducted to examine whether there were effects of warning statement content using factorial designs. When significant effects were found, post hoc comparisons were performed using Bonferroni corrections. One analysis used the ratings from 9 warnings from the original set of 13 were used to form a factorial analysis in a 3 (Immediacy: none, *Exit Now*, *Exit Immediately*) X 3 (Directive: none, *Use Stairs*, *Do Not Use Elevator*) repeated measures design. The 9 warning statement conditions used in the analysis are listed in Table 2 (Exp. 1). The ANOVA revealed a significant immediacy phrase main effect, $F(1.64, 170.82) = 43.83$, $MSE = 4.06$, $p < .001$, $\eta_p^2 = 0.30$; using corrected degrees of freedom (Huynh-Feldt estimate with ε = 0.82). Table 3 (Exp. 1) shows the main effect means for immediacy phrase along the bottom row.

Warning statements with an immediacy phrase (either *Exit Now* or *Exit Immediately*) were rated as significantly more acceptable than warning statements without an immediacy phrase. There was no significant difference between the two immediacy phrases.

There was also a significant directive phrase main effect, $F(1.80,186.87) = 40.01$, $MSE = 6.35$, $p < .001$, $\eta_p^2 = 0.28$; degrees of freedom were corrected based on Huynh-Feldt ε of 0.90. The means are shown in the last column of Table 3 (Exp. 1). Warning statements with a directive (*Use Stairs* or *Do Not Use Elevator*) had significantly higher ratings than ratings for warning statements with no directive. There was no significant difference between the two directive phrases.

There was also a significant interaction of Immediacy X Directive phrase, $F(3.13, 325.52) = 6.24$, $MSE = 2.20$, $p < .001$, $\eta_p^2 = 0.06$; degrees of freedom were corrected based on Huynh-Feldt ε of 0.78. Fig. 1 (Exp. 1) shows the interaction graphically. Overall, the pattern seen in the figure largely reflects the main effects previously described. The presence of either phrase contributed to acceptability relative to their absence, i.e., having either immediacy or a directive phrase increased rated acceptability relative to statements without those phrases. Having both phrases was better than one. A simple effects analysis indicated that for statements that included the component "*Fire, Fire, Exit Now*" the addition of *"Use Stairs"* was rated significantly higher than when the phrase "*Do Not Use Elevator*" was included.

2.2.4. Correlation

The next analysis evaluated the relationship between warning length—as measured with the number of letters—and acceptability ratings. Ratings for each of the 13 statements were collapsed across participants. A non-parametric test was used due to the data being negatively skewed (skewness = −1.65, $SE = 0.62$; kurtosis = 3.92, $SE = 1.19$). A correlation test revealed a positive relationship between warning length and ratings, $r_s\ (11) = 0.72$, $p = .006$; longer fire warnings were rated more acceptable than those with fewer letters. Table 4 (Exp. 1) suggests a curvilinear relationship between warning length and acceptability because the moderate-length (7 syllables) warning was rated highest compared to warnings with more or fewer syllables. A hierarchical regression was performed, but the quadratic factor that would indicate a curvilinear relationship was not significant.

2.3. Discussion

The results showed that certain visually-presented fire warnings are considered more acceptable than others. The shortest warning, *Fire, Fire*, was rated as the least suitable compared to the others. This makes sense because it does not give information beyond signaling a fire emergency. The highest rated statements tended to have the longest statements and have the most information.

Additional analyses of the warning statements revealed several other findings. Warnings with immediacy phrases were perceived as more acceptable than without the immediacy phrases. However, the two immediacy phrases used in this study did not produce significantly different ratings, probably because *Now* and *Immediately* are synonyms. This particular finding does not support the idea that a shorter word would be preferred over a longer word with the same meaning.

Warning statements with egress directives were more acceptable than those without those directives. Although the general trend showed no difference between the presence of the directive to use the stairs compared to avoiding the elevator, some data did not follow this trend. There was an interaction effect in which "*Fire, Fire, Exit Now*, *Use Stairs"* produced significantly higher ratings than "*Fire, Fire, Exit Now*, *Do Not Use Elevator."* One potential reason for this is result that not all multistory buildings have elevators, and thus it would not make sense in those cases to warn against using an elevator. The phrase "*Use Stairs"* might have been considered a better choice because it is applicable to any multistory building. Nevertheless, it should be noted the scenario

Table 2
(Exp. 1) Nine warning statements organized as a function of immediacy and directive phrase factors.

Directive Phrase	*Immediacy Phrase* *None*	*Exit Now*	*Exit Immediately*
None	Fire, Fire	Fire, Fire, Exit Now	Fire, Fire, Exit Immediately
Use Stairs	Fire, Fire, Use Stairs	Fire, Fire, Exit Now, Use Stairs	Fire, Fire, Exit Immediately, Use Stairs
Do Not Use Elevator	Fire, Fire, Do Not Use Elevator	Fire, Fire, Exit Now, Do Not Use Elevator	Fire, Fire, Exit Immediately, Do Not Use Elevator

Note. Warnings included in analysis were ordered with immediacy phrase, directive phrase. Warnings in other orders were not included.

described the building to have an elevator. Another potential reason is that participants may have considered the directive to use the stairs to be redundant and shorter than the directive not to use the elevator.

This experiment examined effects of visually-presented (print) text warnings. The findings would be useful in selecting terminology for static signage and dynamic display monitors near one or more elevators of a multistory building, and useful for certain vulnerable groups such as deaf persons. However, most warnings for fire emergencies are auditory alarms. Accordingly, the purpose of Experiment 2 was to evaluate whether the results would be similar when fire warning statements are presented auditorily as speech (voice). In certain circumstances, the auditory modality has several advantages over the visual modality. One is that sound is omnidirectional; individuals do not have look in the direction of the source of an auditory warning as is necessary for visual warnings. Warnings given in the auditory modality would be useful for blind occupants and in a visually degraded environment that could occur in a fire, reducing legibility (such as air filled with smoke) (Collins and Lerner, 1983).

3. Experiment 2

In this experiment, a similar set of egress fire warnings used in Experiment 1 was presented auditorily through speech (voice) warnings. In addition to the set of statements used in Experiment 1, the set was supplemented with: (a) a few lengthier warnings, and (b) two non-speech sounds (conventional fire alarm and white noise). Like Experiment 1, warning statements with egress information (immediacy and/or egress instructions) were expected to be rated as more suitable as fire warnings than those without that information. Alternatively, the pattern of results using spoken statements might yield a different pattern of results than was found in Experiment 1 using print warnings because a different modality is being used.

In Experiment 1, there appeared to be a linear increase in acceptability with length. However, this may be true only to a point, whereupon lengthier warnings become less acceptable. If so, this could be especially relevant to auditory warnings because longer warnings need more time to be presented (holding everything else constant), and could delay cognitive processing and actions needed to comply to the warnings (e.g., Rayner et al., 2010). To test this notion, two longer warning were added to the stimulus set using the same components used in the other statements in the set. One alternative is that the results might show the same trend as Experiment 1 with the longest warnings being rated more acceptable. Alternatively, the lengthier warning statements might be rated higher up to a point but even longer (the longest) warnings might be rated less acceptable relative to shorter ones. With

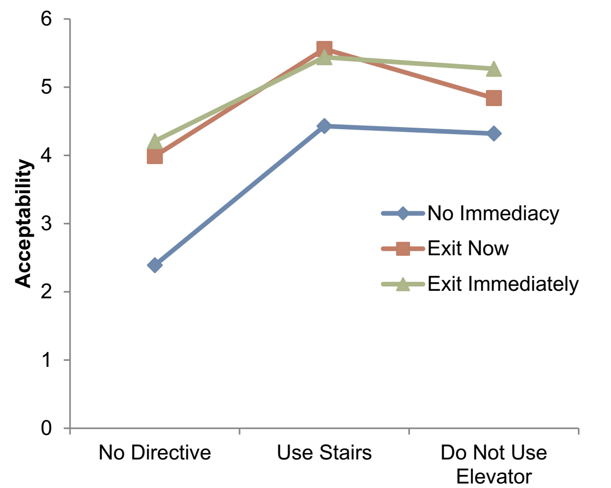

Fig. 1. (Exp. 1) Mean acceptability as a function of Immediacy X Directive Phrase Interaction.

Table 4
(Exp. 1) Mean acceptability ratings as a function of syllable count. Warning statements are in descending order of syllable count.

Syllable Count	Mean (*SD*)
16 syllables	5.27 (2.30)
Fire, Fire, Exit Immediately, Do Not Use Elevator	
14 syllables	5.04 (2.36)
Fire, Fire, Exit Now, Do Not Use Elevator, Use Stairs	
Fire, Fire, Exit Now, Use Stairs, Do Not Use Elevator	
12 syllables	4.73 (1.66)
Fire, Fire, Exit Now, Do Not Use Elevator	
Fire, Fire, Do Not Use Elevator, Exit Now	
Fire, Fire, Use Stairs, Do Not Use Elevator	
11 syllables	4.88 (1.58)
Fire, Fire, Exit Immediately, Use Stairs	
Fire, Fire, Do Not Use Elevator	
9 syllables	4.21 (2.00)
Fire, Fire, Exit Immediately	
7 syllables	5.56 (1.72)
Fire, Fire, Exit Now, Use Stairs	
5 syllables	3.99 (2.13)
Fire, Fire, Exit Now	
4 syllables	4.43 (2.08)
Fire, Fire, Use Stairs	
2 syllables	2.39 (2.42)
Fire, Fire	

Table 3
(Exp. 1) Mean acceptability ratings of warning statement as a function of immediacy and directive phrase factors. Standard deviations are in parentheses.

Directive Phrase	Immediacy Phrase None	Exit Now	Exit Immediately	Overall Mean
None	2.39 (2.42)	3.99 (2.13)	4.21 (2.00)	3.53 (1.77)
Use Stairs	4.43 (2.08)	5.56 (1.72)	5.44 (2.00)	5.14 (1.47)
Do Not Use Elevator	4.32 (1.84)	4.84 (2.04)	5.27 (2.30)	4.81 (1.74)
Overall Mean	*3.71 (1.50)*	*4.80 (1.35)*	*4.97 (1.64)*	

the latter, length (number of syllables and time duration) was expected to have a curvilinear-shaped relationship with acceptability ratings.

3.1. Method

3.1.1. Participants

Sixty-five students (31 males and 34 females) taking introductory psychology courses at North Carolina State University participated. They received course credit for their participation. The overall mean age of the sample was 18.7 years ($SD = 1.33$). Participants reported their ethnicities as follows: 42 Caucasians, 4 Africans, 7 African Americans, 8 Asians, 1 East Indian, 1 Hispanic/Latino, and 2 Mixed Race. The sample comprised of 84.6% native English speakers. Three people had mild hearing impairments according to the scoring criteria of the Hearing Screen Inventory (Coren and Hakstian, 1992).

3.1.2. Experimental design and stimuli

Recordings of six native English speakers (3 males and 3 females) were used for 15 fire warning statements. Thirteen of these 15 warnings are given in Table 1 (Experiment 1), and the two longer warnings added to the set for this experiment were: "*Fire, Fire*, *Exit Immediately*, *Use Stairs*, *Do Not Use Elevator"* and "*Fire, Fire*, *Exit Immediately*, *Do Not Use Elevator*, *Use Stairs*." For these two statements, only the order of the directive phrases was different. To produce the stimuli, speakers were recruited from the same pool of university students used in the subsequent acceptability ratings part of the experiment but none participated in both parts. Use of multiple speakers for generating the stimuli was intended to promote generalizability as opposed to using only one or two speakers, which is an unfortunate common practice in some prior speech warning research (see Barzegar and Wogalter, 1998, for a discussion about this issue). Speakers participated individually in the recording sessions. Speakers were instructed to speak loudly into a microphone to enunciate the warnings to capture attention and convey the messages clearly to building occupants. Speakers practiced vocalizing the warnings into a microphone until the sound level was kept at a fairly consistent and approximate level of 80 dBA (measured by a sound level meter), and could produce recordings without distortion. Recording sessions were held in a small, quiet room with approximately 40 dBA of background noise. Audacity sound editing (open source) software was used to capture the recordings. The order of the warnings was randomized so that no speakers vocalized them in the same order. Speakers spoke all 15 fire warnings twice with the second recording serving as a backup recording.

The final set included 90 speech warnings (15 statements from each of 6 speakers). There were also two non-speech sounds in the set of stimuli. One non-speech sound was a short audio clipping of a fire alarm (3.16 s, ~80 dBA) produced by a Simplex 4100U Addressable Fire Alarm system and represented a conventional buzzer-type fire alarm. The other sound was white noise (1.27 s, ~80 dBA) generated using Audacity software, and was considered a control condition. Thus, participants heard 92 warnings (15 statements X 6 speakers = 90 and two non-speech sounds). Five playlists were generated based on different random orders of the warnings. Thirteen participants were assigned to listen one of the five playlists. The playlists included 7 s of silence separating each warning to provide time for participants to write down their rating on a numbered response sheet.

3.1.3. Procedure

After participants signed a consent form, the task instructions and scenario were presented orally by the experimenter and on paper. The scenario asked participants to imagine they worked at a large company and a very loud fire alarm had been triggered. No one nearby knew whether it was a real fire or a fire drill. There was confusion with a few people waiting for an elevator instead of using the stairs. After that introduction, participants were informed that the study was examining different warnings to determine what would be most appropriate for alerting people during a real fire emergency. Furthermore, the instructions indicated a number of spoken statements and sounds, such as an alarm and white noise, would be presented. They were asked to imagine statements and sounds would be repeated several times so as to encourage people to leave the building safely and quickly. After each statement or sound was presented, participants were asked to give a rating on how acceptable it would be as a fire evacuation warning, using the same rating scale used in Experiment 1. They were told acceptability might include a variety of concepts, such as perceived hazardousness, appropriateness, and how careful one would be to follow the statement or sound.

Participants donned Sony MDR-XD100 over-the-ear headphones, which was used to control presentation and reception of the auditory warnings, including aspects associated with sound level and ambient noise. After the trials concluded, participants completed surveys for hearing loss and demographics. No one was excluded from participating based on these measures.

3.2. Results

Table 5 shows the means and standard deviations for the auditory stimuli used in this experiment. The two highest-rated warnings were also the longest warnings: *Fire, Fire*, *Exit Immediately*, *Do Not Use Elevator*, *Use Stairs* and *Fire, Fire*, *Exit Immediately*, *Use Stairs*, *Do Not Use Elevator*. The lowest-rated spoken warning was *Fire, Fire*, and lowest rated was white noise. Other than white noise, all other warnings' ratings were rated at least "somewhat acceptable" based on the anchors of the rating scale.

3.2.1. One-way ANOVAs

A one-way repeated measures ANOVA involving all 17 conditions (15 speech fire warnings and two non-speech sounds) was significant, $F(3.99, 255.30) = 37.65$, $MSE = 5.09$, $p < .001$, $\eta_p^2 = 0.37$. Mauchly's test indicated the assumption of sphericity had been violated, $\chi2(135) = 948.50$, $p < .001$; therefore, degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.25$). Post hoc comparisons were conducted using a Bonferroni correction. *Fire, Fire, Exit Immediately, Do Not Use Elevator, Use Stairs* was rated significantly more acceptable than *Fire, Fire, Exit Now, Use Stairs, Do Not Use Elevator, Fire, Fire, Use Stairs, Do Not Use Elevator*, and warnings ranging from *Fire, Fire, Exit Immediately, Use Stairs* to white noise in Table 5. *Fire, Fire, Exit Immediately, Use Stairs, Do Not Use Elevator* was rated higher than warnings ranging from *Fire, Fire, Exit Now, Use Stairs* to white noise. *Fire, Fire, Exit Now, Do Not Use Elevator, Use Stairs* was higher than warnings ranging from *Fire, Fire, Exit Now, Do Not Elevator*

Table 5
Mean acceptability (and standard deviations) of 17 warnings in descending order.

Warning	Mean (*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)

to white noise. *Fire, Fire, Exit Now, Use Stairs, Do Not Use Elevator* was higher than warnings ranging from *Fire, Fire, Do Not Use Elevator* to white noise. *Fire, Fire, Use Stairs, Do Not Use Elevator* was rated higher than warnings ranging from *Fire, Fire, Use Stairs* to white noise. The conventional fire alarm was rated more acceptable than *Fire, Fire* and white noise. *Fire, Fire, Exit Immediately, Use Stairs* was rated higher than *Fire, Fire, Use Stairs* and warnings ranging from *Fire, Fire, Exit Immediately* to white noise. *Fire, Fire, Exit Now, Use Stairs* was rated higher than warnings ranging from *Fire, Fire, Do Not Use Elevator* to white noise. *Fire, Fire, Exit Immediately, Do Not Use Elevator* was rated higher than warnings ranging from *Fire, Fire, Do Not Use Elevator* to white noise. *Fire, Fire, Do Not Use Elevator, Exit Now* was rated higher than *Fire, Fire, Do Not Use Elevator*, *Fire, Fire*, and white noise. *Fire, Fire* was rated significantly lower than the other spoken warnings. Lastly, white noise was rated significantly lower than all other conditions. In general, the warnings with more phrases were considered more acceptable than those with fewer phrases. The conventional fire alarm was perceived more acceptable compared to *Fire, Fire*.

3.2.1.1. Factorial ANOVAs. Factorial analyses were used to examine the effects of content phrases and whether parts interact. A 3 (Immediacy phrase: none, *Exit Now*, *Exit Immediately*) X 2 (Stairs directive phrase: no, yes) X 2 (Elevator directive phrase: no, yes) factorial repeated measures ANOVA was conducted. When significant effects were found, post hoc comparisons were performed using Bonferroni corrections. There was a significant Immediacy phrase main effect, $F(1.56, 102.04) = 34.59$, $MSE = 1.31$, $p < .001$, $\eta_p^2 = 0.35$. Mauchly's test indicated that the assumption of sphericity had been violated, $\chi 2\ (2) = 20.71$, $p < .001$; therefore, degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity ($\varepsilon = 0.80$). Warning statements with either of the immediacy phrases, *Exit Now* ($M = 4.38$, $SD = 1.03$) or *Exit Immediately* ($M = 4.45$, $SD = 0.97$), were rated significantly higher than warning statements that lacked an immediacy phrase ($M = 3.77$, $SD = 1.24$). No significant difference was found between *Exit Now* and *Exit Immediately*.

There was a significant Stairs directive phrase main effect, $F(1,64) = 73.65$, $MSE = 1.45$, $p < .001$, $\eta_p^2 = 0.54$. The bottom row of Table 6 shows the means and standard deviations. Warning statements with *Use Stairs* were rated significantly higher than those without *Use Stairs*. There was significant Elevator directive phrase main effect, $F(1,64) = 25.12$, $MSE = 2.42$, $p < .001$, $\eta_p^2 = 0.28$. Warning statements with *Do Not Use Elevator* ($M = 4.47$, $SD = 1.01$) were rated significantly higher than those without the elevator directive phrase ($M = 3.92$, $SD = 1.17$).

There were also several significant interactions. One was an Egress Immediacy X Stairs Phrase interaction, $F(2,128) = 21.03$, $MSE = 0.34$, $p < .001$, $\eta_p^2 = 0.25$. Fig. 2 shows a line graph of this interaction. Mean acceptability dropped significantly when the warning statements lacked both an immediacy and stairs directive phrases.

There was a significant Egress Immediacy X Elevator Phrase interaction, $F(2, 128) = 11.69$, $MSE = 0.30$, $p < .001$, $\eta_p^2 = 0.15$. Fig. 3 displays the means associated with this interaction. When the warning

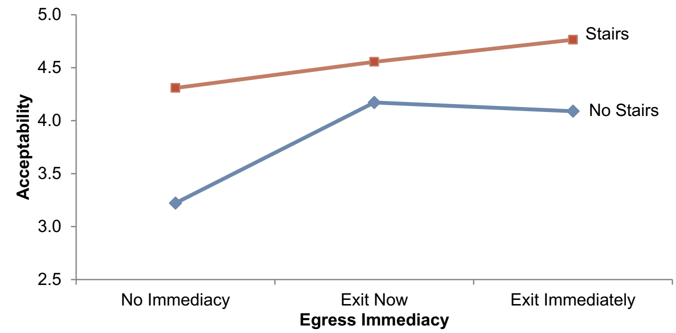

Fig. 2. Egress immediacy X stairs phrase interaction.

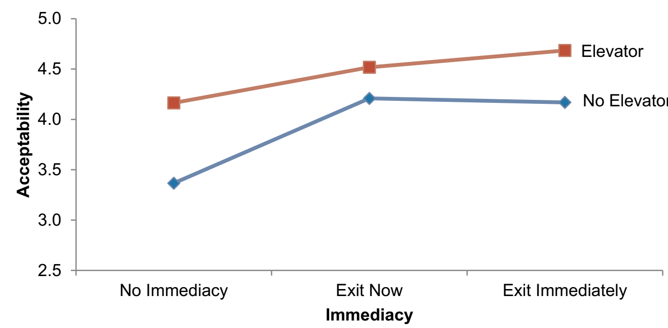

Fig. 3. Egress immediacy X elevator phrase interaction.

statements had no immediacy and no elevator directive phrases mean acceptability dropped significantly. Thus, both of the two-factor interactions suggest warning statements lacking certain information are considered less acceptable than warnings with more information.

The ANOVA also yielded a significant 3-factor interaction, $F(1.74, 111.56) = 4.13$, $MSE = 0.49$, $p = .02$, $\eta_p^2 = 0.06$. Mauchly's test indicated that the assumption of sphericity had been violated for the 3-way interaction effect, $\chi 2\ (2) = 10.05$, $p = .007$; therefore, the degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity ($\varepsilon = 0.80$). The means associated with this interaction are shown in two panels of Fig. 4 (a & b). The effect of the absence of information on acceptability suggested in the earlier-described two-factor interactions is further highlighted in the three-way interaction. Fig. 4a (top graph) shows one particularly low mean in the bottom left, the mean for *Fire, Fire*, which is the warning statement with the least amount information.

3.2.2. Hierarchical regression

Means were computed by collapsing ratings across participants for all warnings statements heard by participants (90 or 15 statements X 6 speakers). Each warning's recorded duration (in s) was converted to a difference score centered by subtracting the raw duration from the mean. An initial analysis revealed number of syllables and duration were strongly, positively correlated ($r = 0.83$). The overlap may exist because it generally takes more time to enunciate statements with more

Table 6
Means Organized by Egress Immediacy, Stairs, and Elevator Phrases. Standard deviations are in parentheses.

Egress Immediacy		Includes Stairs Directive No	Yes	*Overall Mean*
None	No Elevator Directive	2.71 (1.78)	4.03 (1.39)	3.37 (1.50)
	Elevator Directive	3.74 (1.28)	4.59 (1.22)	4.16 (1.12)
Exit Now	No Elevator Directive	4.01 (1.42)	4.41 (1.26)	4.21 (1.18)
	Elevator Directive	4.31 (1.09)	4.77 (1.36)	4.54 (1.14)
Exit Immediately	No Elevator Directive	3.85 (1.14)	4.49 (1.22)	4.17 (1.09)
	Elevator Directive	4.33 (1.13)	5.10 (1.27)	4.72 (1.08)
Overall Mean		3.83 (1.07)	4.57 (1.05)	

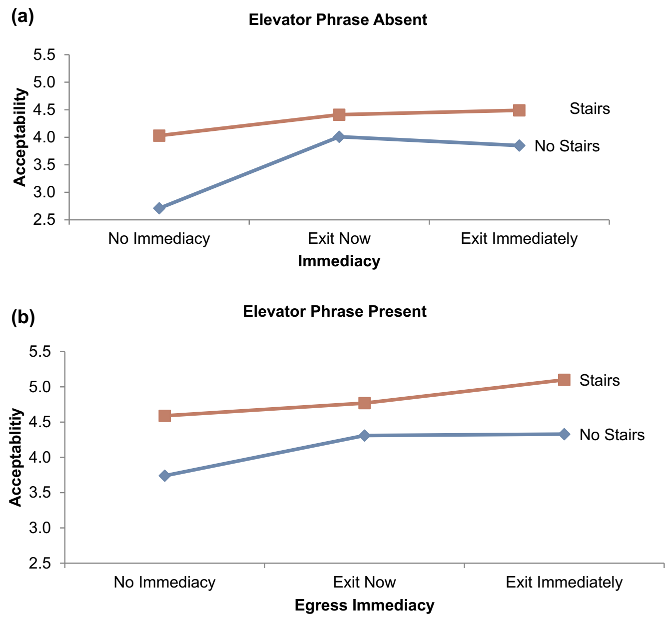

Fig. 4. (a & b). Mean acceptability in a three-factor interaction shown as a function of presence versus absence of elevator phrase (and thus split into two graphs). The top graph (a) shows the interaction of immediacy and stairs phrases when the elevator phrase is absent; the bottom graph (b) shows the same two factors when the elevator phrase is present.

syllables. The syllable variable was excluded in later analyses because it was not a significant predictor over and above the duration variable. A scatterplot of the standardized residuals against standardized predicted mean values showed homoscedasticity of variance was met. No standardized residuals were greater than 2.58. A hierarchical regression was used to test for a curvilinear relationship between speech fire warning duration and ratings. In order to test a non-linear relationship in a regression, an interaction term of a predictor variable with itself (i.e., cross product or squared term) is included in the model (Keith, 2015). The squared term is used to address whether the outcome variable depends on the level of the predictor variable. The predictor variable and its squared term are represented with centered scores; centering reduces the correlation between the predictor variable and squared predictor variable (Keith, 2015).

The first block of predictors in the regression included centered warning duration. As can be seen in Table 7, the first block accounted for 50% of the variance in acceptability ($F(1, 89) = 89.39$, $p < .001$). Duration was a positive, significant predictor. Squared, centered duration was added in the second block ($F(1, 89) = 50.07$, $p < .001$), and the amount of variance significantly increased to 54%. The squared term was a significant, negative predictor of acceptability, indicating the regression line contained a negative, convex curve. As warning duration increases past a certain point, acceptability ratings decrease. Fig. 5 (Exp. 2) shows the quadratic curve fitting line is a better fit of the

Table 7
Hierarchical regression with duration (predictor) and acceptability ratings (outcome).

	R^2	Adjusted R^2	*B*	*SE (B)*	Β
Block 1	.50	.50			
Constant				0.053	4.30
Duration Ctr.			.71*	0.053	.50
Block 2	.54	.52			
Constant				0.07	4.41
Duration Ctr.			.71*	0.052	.50
Duration Ctr.^{2}			-.18**	0.043	-.104

Note: $^*p < .001$; $^{**}p < .05$.

data compared to the linear fitting line. The curved fitting line is better fit for the lower duration warnings with lower ratings. Also, note at the rightmost part of the figure, there are a few points that show a downward trend; those points lay closer to the curved line as opposed to the linear one.

3.3. Discussion

As expected, speech warning content affected acceptability ratings. Some of the results in this experiment using speech warnings were similar to those of Experiment 1 using visual print (text) warnings. For example, *Fire, Fire* had the lowest rating of all statements. The results also showed that statements with egress immediacy and directives increased ratings. More about the similarity and differences in the findings between the two experiments are described in the General Discussion section.

The factorial ANOVA revealed the Use Stairs phrase had a greater positive effect on acceptability than the Elevator phrase. When both directive phrases were present in a warning, ratings were higher compared to ratings when one phrase was present. This result implies the use of both phrases was beneficial, possibly because it leaves no ambiguity on what to do and what not to do during egress. It reduces the potential for some persons not fully considering the concepts during an emergency situation when attention allocation is more limited. Statements that include both the stairs and elevator directives may provide a more forceful and explicit expression on what exactly should be done in a fire emergency.

Although warnings with multiple phrases were perceived more acceptable, the relationship between warning duration and acceptability ratings reveals that longer warnings were better than shorter warnings, except for the very longest warnings in which there was a small but drop in the acceptability ratings. This result seems to contradict the factorial ANOVA, but there were differences in the data used. The ANOVA used means of each statement, averaged over multiple speakers. These statements were not aggregated across speakers in the regression. Specific (and lengthier) wording was better up to a point. Another aspect of duration is the actual time taken to enunciate the information. Warnings need to be long enough to provide critical information but presented quickly so that receivers can respond quickly. The non-linear relationship revealed an increase in perceived acceptability up until a certain point in which warnings that took longest to be enunciated were rated less acceptable. This non-linear (curvilinear) trend is supported in earlier research. For example, in a study reported by Simpson and Marchionda-Frost (1984), pilots listened to voiced threat messages that varied in speech rate (duration decreasing with speech rate increasing) as well as other characteristics. The pilots showed preference for the messages voiced in a moderate speech rate as opposed to those messages in slow and fast rates. Very quick and very slow rates tend to reduce intelligibility. Because time constraints are likely in emergencies, shorter durations are preferred because lengthy warnings could delay important emergency responses. However, a longer warning can have the benefit of increased intelligibility (i.e., ability to identify all of the words) because more content (or context) can be heard particularly when some is obscured or masked due to noise being present (Sanders and McCormick, 1993). However, longer speech warnings may be less effective due to constraints of attentional capacity and memory (Wogalter et al., 1999).

In addition to comparing content in speech warnings, this research compared the speech warnings with non-speech auditory warnings. The results showed that non-speech fire-alarm sound was rated more acceptable compared to *Fire, Fire*. This was somewhat surprising because *Fire, Fire* identifies the hazard, which is more than what a traditional fire alarm explicitly indicates. A loud alarm, which may or may not be repetitive, could indicate many different abnormal situations, which may or may not be a fire emergency. Undoubtedly the participants have heard conventional fire alarms before (on the university campus and

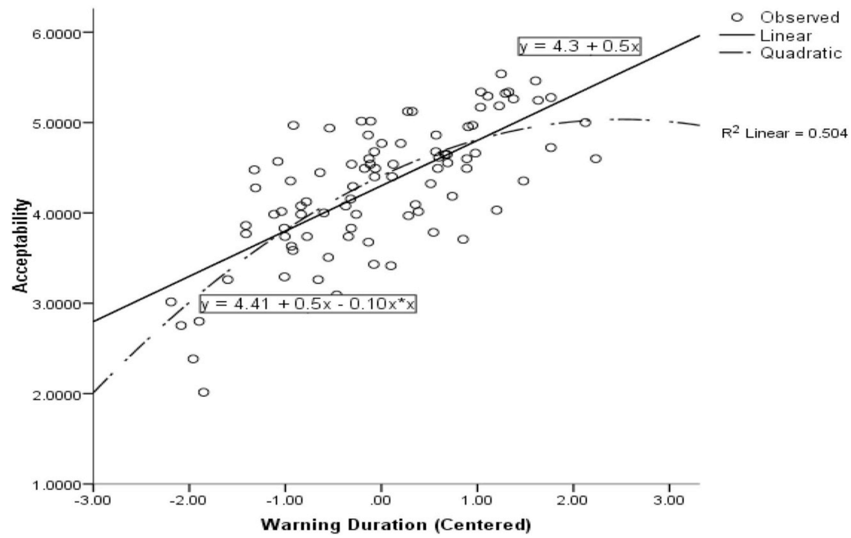

Fig. 5. (Exp. 2) Scatter plot of duration and acceptability ratings with linear and curved fitting lines.

elsewhere in their experience, such as during fire drills), and as a consequence, cued an association with our fire alarm condition. The fire alarm sound included in this study likely fit peoples' pre-existing expectations as to how a typical fire warning sounds, resulting in its moderately high mean rating.

White noise was included as a control condition because it is not semantically associated with fire emergencies. White noise received the lowest ratings of all conditions. Its relatively high standard deviation suggests some confusion about its appropriateness in this application. It was a short, harsh burst of sound, which might have startled some participants, which might have been influential in participants' ratings of this condition. Further research is needed to determine a more definitive conclusion regarding this condition.

During a real fire emergency, the aural background is likely to be noisy including noise of various types along with utterances by people hurrying to evacuate. If noise overlaps in characteristics with the egress warning or there is a low signal (warning) to noise (S/N) ratio, then the warning information could be masked and unintelligible (see, e.g., Haas and Edworthy, 2006). Baldwin (2011) showed increased perceived urgency and decreased response time as S/N ratio increased for collision avoidance speech warnings. This suggests that S/N ratio be considered in future spoken fire warning research.

4. General discussion

Both experiments revealed a similar pattern of ratings among the same warnings used. The mean ratings of the visually-presented warning statements (Experiment 1) and the same subset of auditorily-presented warnings (Experiment 2) were positively correlated, r_s (11) = 0.80, p = .001. The main difference in the findings involved the relationship between warning length and rated acceptability. Experiment 1 with visual print stimuli used the number of letters for length, whereas Experiment 2 with auditory speech stimuli used duration of utterance. In the first experiment there was a linear positive correlation with acceptability ratings, whereas in the second experiment there was a main linear component but also a curvilinear component where the longest warnings were rated less acceptable than warnings somewhat shorter. In Experiment 1, the highest-rated warning was one moderate in length as shown in Table 4. It is possible that the curvilinear relationship was non-significant in Experiment 1 because there was not enough power to get a significant finding. Experiment 1 had 13 ratings included in the correlational analysis, but Experiment 2 had 90 ratings in the regression analysis. It should be noted that the longest warnings in Experiment 2 were not present in Experiment 1's set, which might account for the difference in the overall trends. Further research using longer statements is needed to follow up on these warning length trends. In general, these findings suggest fire-warning statements should contain a certain amount of specific information so as to be complete in expression. Extreme brevity should be avoided, while the same is probably true for extreme lengthiness.

The present research offers empirical data that could serve to guide the selection of fire emergency warnings. However, there are other factors that would likely play a role and would need to be considered in designing real-world fire warnings. Explicit instructions to use stairs and/or avoid using the elevator could serve as reminders to people in conditions of stress of a real emergency. Explicit reminders would serve to increase the likelihood that associated knowledge will be brought into awareness at the appropriate time during an emergency egress, and potentially "break" the use of automatic, habitual behavior of taking an elevator to egress, and potentiating the use the stairs instead.

In a fire emergency, speech warnings have certain advantages over visual warnings and may be more appropriate in this context than visual warnings (Conzola and Wogalter, 1999; Wogalter and Young, 1991). However, other research indicates the opposite, i.e., visual warnings can be better than auditory warnings. For example, in television commercials with warnings in print streams versus voiceovers, visual presentation has been found to lead to better comprehension and memory than auditory/speech presentation (Barlow and Wogalter, 1993; Wogalter et al., 2014a). Although not tested in the present experiments, warning research generally shows that combining both modalities, and thus providing redundancy, is usually better than either modality alone (e.g., Barlow and Wogalter, 1993; Wogalter et al., 2014b; Wogalter and Young, 1991).

Giving specific information by way of text or voice provides more acceptable fire warnings than warnings that are nonspecific or absent relevant information (and making them potentially ambiguous). In both experiments, the briefest warning, *Fire, Fire,* was rated the lowest of the text and speech warnings. Although that warning is specific on the potential hazard, it does not provide any additional details. Of course, any single kind of warning is not going to be adequate for everyone. Auditory warnings have limited utility for deaf persons, and English language warnings have low effectiveness for nonusers of English. Thus, warnings given in multiple, supplemental ways, such as including pictorial symbols (e.g., Young and Wogalter, 1990), are likely to provide a wider reach. Future research could evaluate acceptability by varying the modality and medium in warnings to different populations and environments. Also further research on intonation and voicing (cf. Barzegar and Wogalter, 1998; Hollander and Wogalter, 2000) of speech-based fire warnings and their effects on different demographic groups, e.g., ethnicity, gender, native language (Smith-Jackson et al., 2013) would be beneficial in addressing needs and preferences of more vulnerable groups.

There are other ways in which research on fire warnings could be extended. The warnings could be rated on dimensions other than acceptability. While acceptability was an overall single judgment comprised of a number of considerations (perceived hazard, appropriateness, and anticipated carefulness), a future research focus could be to determine separate contributions of different dimensions important for warning effectiveness. Future studies could assess judgments of urgency and intended carefulness measures (and others) to determine whether a similar pattern of results is found. Greater knowledge on the relationships among measures would provide guidance for the design of effective warnings.

Additionally, subsequent testing of speech fire warnings could be done in simulated (augmented or virtual) environments. Virtual environments have been used to measure behavioral compliance to warnings (Duarte et al., 2010) including personalized work-related safety warnings (Duarte et al., 2012) and fire egress signs (Duarte et al., 2014). The benefit of using simulated environments is that the warnings can be presented in realistic-appearing conditions while not actually exposing participants to actual harm.

This study used a relatively novel approach to assessing warning design. It involved the manipulation of a relatively large set of statements comprising potential warnings. Although several earlier studies have looked at wording in warnings, the vast majority has looked at the differential effects of individual signal words presented in visual print (e.g., Chapanis, 1994; Hellier and Edworthy, 2006) or speech form (e.g., Arrabito, 2009; Barzegar and Wogalter, 1998; Edworthy et al., 2003; Hellier et al., 2002; Hollander and Wogalter, 2000; Ljungberg and Parmentier, 2012). In the real-world settings, most visual and auditory warnings are more complex than single words. The present study included warning statement of greater complexity than individual signal words. Future research should continue to examine systematically manipulated complex warnings to better understand what kinds of component wording benefit and hinder warning effectiveness measures.

The present study involved warnings for a scenario involving a fire in multistory building, however, future research could confirm and extend this work by examining other scenarios such as those involving other kinds of emergency situations and dangerous environments and other kinds of hazards including industrial and consumer products requiring different responses. Future research could also examine how to present warnings in integrated systems with multiple kinds of sensors and warnings (e.g., combined detectors of smoke, carbon monoxide, and explosive gas) and which would differentially prioritize and provide messages depending on variables such as areas of detection, building design, functioning systems, and the seriousness of the hazard. For example, for a given situation it may be possible to customize the egress information, e.g., instruct which staircase is most safe.

Acknowledgements

Portions of this research were presented at the Human Factors and Ergonomics Society Annual Meetings (Taylor and Wogalter, 2012, 2015). The authors are grateful for the help of Dr. Soyun Kim in helping operationalize the study.

References

Ahrens, M., 2014. Smoke alarms in U.S. home fires. National Fire Protection Association Retrieved from:. http://www.nfpa.org.
Arrabito, G.R., 2009. Effects of talker sex and voice style of verbal cockpit warnings on performance. Hum. Factors 51, 3–20.
Baldwin, C.L., 2011. Verbal collision avoidance messages during simulated driving: perceived urgency, alerting effectiveness and annoyance. Ergonomics 54, 328–337.
Baldwin, C.L., Lewis, B.A., 2014. Perceived urgency mapping across modalities within a driving context. Appl. Ergon. 45, 1270–1277.
Barlow, T., Wogalter, M.S., 1993. Alcoholic beverage warnings in magazine and television advertisements. J. Consum. Res. 20, 147–156.
Barzegar, R.S., Wogalter, M.S., 1998. Intended carefulness for voiced warning signal words. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 42, 1068–1072.
Barzegar, R.S., Wogalter, M.S., 2000. Intended carefulness ratings for voice warning statements. Proc. Int. Ergon. Assoc. Hum. Factors Ergon. Soc. Congr. 44 (3), 686–689.
Bella, F., Silvestri, M., 2017. Effects of directional auditory and visual warnings at intersections on reaction times and speed reduction times. Transport. Res. F Traffic Psychol. Behav. 51, 88–102.
Chapanis, A., 1994. Hazards associated with three signal words and four colours on warning signs. Ergonomics 37, 265–275.
Collins, B.L., Lerner, N.D., 1983. Evaluation of Exit Symbol Visibility. National Bureau of Standards (NTIS), Springfield VA.
Conzola, V.C., Wogalter, M.S., 1999. Using voice and print directives and warnings to supplement product manual instructions. Int. J. Ind. Ergon. 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. J. Speech Hear. Res. 35, 921–928.
Duarte, E., Rebelo, M.S., Wogalter, M.S., 2010. Virtual reality (VR) and its potential for evaluating warning compliance. Hum. Factors Ergon. Manuf. Serv. Ind. 20, 526–537.
Duarte, E., Rebelo, F., Teles, J., Wogalter, M.S., 2012. A personalized speech warning facilitates compliance in an immersive virtual environment. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 56, 2045–2049.
Duarte, E., Rebelo, F., Teles, J., Wogalter, M.S., 2014. Behavioral compliance for dynamic versus static signs in an immersive virtual environment. Appl. Ergon. 45, 1367–1375.
Edworthy, J., Hellier, E., Rivers, J., 2003. The use of male or female voices in warnings systems: a question of acoustics. Noise Health 6, 39–50.
Edworthy, J., Meredith, C., Hellier, E., Rose, D., 2013. Learning medical alarms whilst performing other tasks. Ergonomics 56 (9), 1400–1417.
Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., Pye, E., 2017. The recognizability and localizability of auditory alarms: setting global medical device standards. Hum. Factors 59 (7), 1108–1127.
Guillaume, A., Pellieux, L., Chastres, V., Drake, C., 2003. Judging the urgency of nonvocal auditory warning signals: perceptual and cognitive processes. J. Exp. Psychol. Appl. 9 (3), 196–212.
Haas, E., Edworthy, J., 2006. An introduction to auditory warnings and alarms. In: Wogalter, M.S. (Ed.), Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 189–198.
Hellier, E., Edworthy, J., 2006. Signal words. In: Wogalter, M.S. (Ed.), Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 407–417.
Hellier, E., Edworthy, J., Weedon, B., Walters, K., Adams, A., 2002. The perceived urgency of speech warnings: semantics versus acoustics. Hum. Factors 44, 1–17.
Hollander, T.D., Wogalter, M.S., 2000. Connoted hazard of voiced warning signal words: an examination of auditory components. In: Proceedings of the Human Factors and Ergonomics Society, vol. 44. pp. 702–705.
Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Englewood Cliffs, NJ.
Karter Jr., M.J., 2014. Fire Loss in the United States during 2013. National Fire Protection Association Retrieved from:. http://www.nfpa.org.
Keith, T.Z., 2015. Multiple Regression and beyond. Routledge, New York.
Laughery, K.R., Page-Smith, K.R., 2006. Explicit information in warnings. In: Wogalter, M.S. (Ed.), Handbook of Warnings. Lawrence Erlbaum Associates, Inc, Mahwah, NJ, pp. 419–428.
Laughery, K.R., Wogalter, M.S., Young, S.L. (Eds.), 1994. Human Factors Perspectives on Warnings: Selections from Human Factors and Ergonomics Society Annual Meetings 1980 - 1993. Human Factors and Ergonomics Society, Santa Monica, CA.
Ljungberg, J.K., Parmentier, F., 2012. The impact of intonation and valence on objective and subjective attention capture by auditory alarms. Hum. Factors 54 (5), 826–837.
Nees, M.A., Helbein, B., Porter, A., 2016. Speech auditory alerts promote memory for alerted events in a video-simulated self-driving car ride. Hum. Factors 58 (3), 416–426.
Rayner, K., Slattery, T.J., Bélanger, N.N., 2010. Eye movements, the perceptual span, and reading speed. Psychon. Bull. Rev. 17 (6), 834–839.
Sanders, M.S., McCormick, E.J., 1993. Human Factors in Engineering and Design, seventh ed. McGraw-Hill, New York.
Simpson, C.A., Marchionda-Frost, K., 1984. Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots. Hum. Factors 26, 509–517.
Smith, J.J., Wogalter, M.S., 2007. Perceived effectiveness of warning messages for use as vocal warnings in residential fire alarms. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 51, 1398–1402.
Smith-Jackson, T.L., Resnick, M.L., Johnson, K.T., 2013. Cultural Ergonomics: Theory, Methods, and Applications. CRC Press, Boca Raton, FL.
Suied, C., Susini, P., McAdams, S., 2008. Evaluating warning sound urgency with reaction times. J. Exp. Psychol. Appl. 14 (3), 201–212.
Taylor, J.R.I., Wogalter, M.S., 2015. Specific evacuation instructions enhance spoken fire warnings. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 59, 1486–1490.
Thomas, I., Bruck, D., 2010. Awakening of sleeping people: a decade of research. Fire Technol. 46 (3), 743–761.
Wan, J., Wu, C., Zhang, Y., 2016. Effects of lead time of verbal collision warning messages on driving behavior in connected vehicle settings. J. Saf. Res. 58, 89–98.
Wickens, C.D., 1984. Processing resources in attention. In: Parasuraman, R., Davies, R. (Eds.), Varieties of Attention. Academic Press, New York, pp. 63–101.
Wickens, C.D., 1996. Designing for stress. In: Driskell, James E., Salas, Eduardo (Eds.), Stress and Human Performance. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 279–296.

Wogalter, M.S., 2006. Handbook of Warnings. CRC Press, Boca Raton, FL Mahwah, NJ: Lawrence Erlbaum Associates.
Wogalter, M.S., Barlow, T., 1990. Injury likelihood and severity in warnings. Proc. Hum. Factors. Soc. 34, 580–583.
Wogalter, M.S., Begley, P.B., Scancorelli, L.F., Brelsford, J.W., 1997. Effectiveness of elevator service signs: measurement of perceived understandability, willingness to comply, and behaviour. Appl. Ergon. 28, 181–187.
Wogalter, M.S., DeJoy, D.M., Laughery, K.R., 1999. Organizing framework: a consolidated communication human information processing (C-HIP) model. In: Wogalter, M.S., DeJoy, D.M., Laughery, K.R. (Eds.), Warnings and Risk Communication. Taylor and Francis, London, pp. 15–24.
Wogalter, M.S., Godfrey, S.S., Fontenelle, G.A., Desaulniers, D.R., Rothstein, P., Laughery, K.R., 1987. Effectiveness of warnings. Hum. Factors 29, 599–612.
Wogalter, M.S., Lim, R., Nyeste, P., 2014a. On the hazard of quiet vehicles to pedestrians and drivers. Appl. Ergon. 45, 1306–1312.
Wogalter, M.S., Mayhorn, C.B., 2006. The future of risk communication: technology-based warning systems (Chap. 63). In: Wogalter, M.S. (Ed.), Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 783–793.
Wogalter, M.S., Shaver, E.F., Kalsher, M.J., 2014b. Effect of presentation modality in direct-to-consumer (DTC) prescription drug television advertisements. Appl. Ergon. 45, 1330–1336.
Wogalter, M.S., Silver, N.C., 1995. Warning signal words: connoted strength and understandability by children, elders, and non-native English speakers. Ergonomics 38, 2188–2206.
Wogalter, M.S., Young, S.L., 1991. Behavioural compliance to voice and print warnings. Ergonomics 34, 79–89.
Young, S.L., Wogalter, M.S., 1990. Comprehension and memory of instruction manual warnings: conspicuous print and pictorial icons. Hum. Factors 32, 637–649.