Ergonomics
Vol. 54, No. 10, October 2011, 879–890

Age-related differences in warning symbol comprehension and training effectiveness: effects of familiarity, complexity, and comprehensibility

M.F. Lesch[a]*, W.J. Horrey[a], M.S. Wogalter[b] and W.R. Powell[a,c]

[a]*Center for Behavioral Sciences, Liberty Mutual Research Institute for Safety, 71 Frankland Road, Hopkinton, MA 01748, USA;* [b]*Department of Psychology, North Carolina State University, Raleigh, NC, USA;* [c]*Center for Disability Research, Liberty Mutual Research Institute for Safety, 71 Frankland Road, Hopkinton, MA 01748, USA*

(*Received 20 April 2011; final version received 28 June 2011*)

Age-related changes in selective attention, inhibitory efficiency, and the ability to form new associations suggest that older adults may have greater difficulty with more complex and less comprehensible symbols. We examined comprehension of symbols varying in terms of ratings of familiarity, complexity, and comprehensibility, by younger (aged 18–35) and older (aged 55–70) adults. It was found that older adults have greater difficulty than younger adults in comprehending warning symbols and that accident scenario training improves comprehension. Regression analyses indicated that familiarity and comprehensibility were important in determining performance on the pre-training comprehension test by both younger and older adults. However, training eliminated the effects of stimulus characteristics for younger adults, while older adults' comprehension continued to be significantly influenced by comprehensibility. We suggest that symbol design incorporates cues to knowledge to facilitate the linkage between new knowledge (i.e. the warning symbol) and relevant knowledge in long-term memory.

Statement of Relevance: Symbol characteristics play an important role in age-related differences in warning symbol comprehension. To optimise comprehension by older adults, symbols should have a clear relationship with a real-world referent. Alternatively, symbol design could incorporate cues to knowledge to facilitate the linkage between new knowledge and relevant knowledge in long-term memory.

Keywords: warnings; safety symbols; ageing; training

1. Introduction

Symbols play an important role in safety communications due to their potential ability to communicate to target populations of varying language backgrounds and reading skills. Additionally, symbols may be useful in addressing some of the perceptual and cognitive changes that occur as a result of ageing. For example, decreased visual acuity can impact the perception of print, and text comprehension may suffer due to declines in working memory (Baddeley 1986). However, research indicates that many warning symbols are poorly understood, and comprehension *may* be worse for older adults. Some studies have observed poorer comprehension of warning symbols by older adults (e.g. Easterby and Hakiel 1981, Collins and Lerner 1982, Zwaga and Boersema 1983, Morrell *et al*. 1990, Hancock *et al*. 1999, Lesch 2003), some no difference (e.g. Lesch 2005), and at least one study (Mayer and Laux 1989) demonstrated better comprehension for older adults for some symbols. This inconsistent pattern may reflect the use of different measures of comprehension, different types of symbols, as well as varying definitions of 'younger' versus 'older' (see Nichols *et al*. 2003). However, an accurate assessment of symbol comprehension is necessary to protect against critical confusions that result in increased risk and to minimise the costs associated with symbol development.

Just as the designers of products should attempt to design-out the hazards associated with those products, designers of warning symbols should attempt to design-out potentially hazardous misunderstandings. However, many hazardous situations are challenging to represent symbolically. For those symbols that are difficult to improve through re-design, training may be the only option. Therefore, it is important to design symbols so as to maximise both initial comprehension and subsequent learning. Here, we are particularly interested in the design of symbols that will be well understood by older populations. It is estimated that, by the year 2030, 25% of the US population will be over the age of 60 (Administration on Aging 2010). Since the age of the workforce is rapidly increasing and miscommunication of warning information can

*Corresponding author. Email: mary.lesch@libertymutual.com

ISSN 0014-0139 print/ISSN 1366-5847 online

DOI: 10.1080/00140139.2011.606924
http://www.informaworld.com

lead to injury or death, it is more important than ever to improve our understanding of age-related differences in warning symbol comprehension.

We suggest that warning symbol design should be sensitive to age-related cognitive changes such as those that occur in attention, language, and memory. The ability to distribute attention across different stimuli or tasks declines with age. A decline in selective attention (including the ability to selectively attend to higher priority signals at the expense of lower priority signals) has been demonstrated for both visual and auditory stimuli (e.g. Allen *et al*. 1994, McCalley *et al*. 1995, Alain *et al*. 1996). When attentional resources must be divided across multiple tasks, older adults tend to experience greater dual-task costs (e.g. Lindenberger *et al*. 2000, Hancock *et al*. 2003). Declines in selective attention may also mean that older adults will find it more difficult to focus on the most relevant portions of a warning or symbol.

Verbal abilities also shift over the lifespan. While older adults tend to have larger vocabularies than younger adults, research indicates that word-finding ability (e.g. Cohen 1979, Pratt *et al*. 1989) and verbal fluency (e.g. McCrae *et al*. 1987) tend to decrease with age. For example, when retelling stories, the verbal output of older adults is characterised by fewer words (and details) and an increased frequency of pauses relative to younger adults. It has been suggested that word-finding difficulty in older adults is due to reduced activation in the memory network (see, for example, Rissenberg and Glanzer 1987).

Theoretical accounts of age-related memory changes include general slowing (e.g. Salthouse 1985), reduced processing resources (e.g. Salthouse *et al*. 1989), and loss of inhibitory functions (e.g. Kane *et al*. 1994) among others (see Luo and Craik 2008). Older adults demonstrate a relative inability 'to effectively inhibit the processing of marginally relevant, irrelevant, and/or distracting stimuli and thoughts' (Kane *et al*. 1994, p. 103). This decline in 'inhibitory efficiency' also impacts memory by preventing 'the dampening of activation along irrelevant retrieval pathways' (Hasher *et al*. 1991, p. 168, see also Zacks *et al*. 1996). One relevant study by Jost *et al*. (2010) examined filtering efficiency of younger and older adults in a visual memory task. On each trial, participants were presented with an array of coloured rectangles of varying orientations. The task was to remember the orientations of only the red items and to ignore the blue and green items. Set size and the number of relevant and irrelevant stimuli were varied. Using neural measures (event-related potential effects), it was determined that older adults (ages 64–92) showed smaller filtering scores than the younger adults (ages 19–38) – that is, to be ignored stimuli remained more active for older adults. This finding is consistent with the inhibition-deficit theory. Such age-related memory changes suggest that older adults may have greater difficulty understanding more complex symbols since activation of irrelevant features is 'dampened' to a lesser extent than would be expected for younger adults.

Older adults also tend to show greater difficulty establishing and retrieving associative links between previously unrelated entities (see Luo and Craik 2008). Therefore, older workers may have greater difficulty in understanding, and learning, the meaning of warning symbols that lack a clear relationship with a real-world referent (i.e. abstract or arbitrary symbols). Some suggestive evidence comes from Lesch (2004) who found that age-related differences in warning symbol comprehension appeared to be related to the rated comprehensibility of the symbols (i.e. given the symbol's verbal label, how easy is it to understand this symbol in isolation?). The symbols that were initially well understood by the older adults, as well as those symbols that benefited from training, tended to be rated as highly comprehensible. Furthermore, while training eliminated effects of familiarity for both younger (18–35 years of age) and older adults (50–65 years of age), comprehensibility continued to correlate with performance by older adults such that symbols that received low ratings of comprehensibility were more poorly understood. It was suggested that comprehensibility reflects the extent to which the symbol is capable of triggering appropriate, context-specific information and that older adults' performance hinges to a greater extent on effective cues to knowledge. However, as the study was not specifically designed to assess comprehensibility effects, a limited range of comprehensibility values was included.

In the discussion thus far, we have documented age-related cognitive changes that have implications for warning symbol comprehension. The manner in which symbol comprehension is gauged also merits special consideration, as age-related differences in symbol comprehension can also be due to the method of measurement. Some guidelines (e.g. ANSI Z535.3 2007) suggest that warning symbol comprehension should be assessed using an open-ended test in which the respondent is asked to describe the symbol's meaning in their own words. However, since word-finding ability and verbal fluency tend to decrease with age, test measures that emphasise verbal output may underestimate comprehension for older adults.

Multiple choice tests, on the other hand, minimise verbal output by requiring the respondent to select the correct answer from a pool of distractors. However, due to age-related declines in selective attention and inhibitory efficiency, older adults may have greater difficulty rejecting incorrect responses. That is, based on their relationship with the target, plausible distractors will be somewhat active in memory. If inhibitory processes are working properly, those incorrect (but related) responses will, effectively, be 'shut off', and the correct response 'wins'. If inhibitory processes are not working properly, there will be a strong competition among possible responses resulting in a decreased likelihood that the correct response is selected.

In light of age-related changes in verbal abilities and inhibitory processing, we suggest the use of a semantic relatedness task to assess comprehension in older adults (see Lesch 2005). In this task, a symbol is presented with a verbal label and the participant must decide whether the verbal label corresponds to the meaning of the symbol. The semantic relatedness task, like the multiple choice test, has the advantage that verbal output is not required. However, unlike a multiple choice test, the correct and distractor labels are presented on separate trials so as to minimise any effect of inhibitory (in)efficiency. Wolff and Wogalter (1998) demonstrated that relative plausibility of distractors plays a major role in determining comprehension levels observed using multiple choice tests. With the semantic relatedness task, however, correct and distractor labels are presented on separate trials, thereby lessening the likelihood of respondents performing the task by comparing relative plausibility of response options.

To provide guidance in symbol design for older workers, we examined both younger and older workers' comprehension of symbols which varied in terms of complexity and comprehensibility – variables which were expected to be sensitive to age-related changes in cognitive processing. Familiarity was also included since its effects on comprehension are well documented (e.g. Hancock *et al.* 2004, see also Ng and Chan 2007). To examine the effect of these variables, participants' ratings of symbol characteristics were used to predict their comprehension performance. Comprehension was measured by accuracy of, and confidence in, judgements on the semantic relatedness task. We also examined the extent to which the accident scenario training improves comprehension of these different symbol types by younger and older workers. During training, symbols were presented paired with accident scenarios which further expanded on the nature of the hazard, required or prohibited actions, and consequences of failures to comply. Earlier studies have documented the effectiveness of accident scenario training (Lesch 2003, 2004, 2008a, 2008b). However, it was of further interest to determine the extent to which this training is effective in addressing comprehension difficulties associated with different symbol types. Prior to training, it was hypothesised that familiarity would be most important in determining comprehension of warning symbols. If a symbol is already familiar to a respondent, then its complexity and comprehensibility should have minimal effects. Comprehensibility and complexity would be expected to be more critical for relatively unfamiliar symbols. It was expected that training would serve to increase the familiarity of the symbols and that it might also aid in focusing attention on the most important aspects of the symbol. However, it is unclear whether training effectiveness would be impacted by comprehensibility of the symbols.

2. Method

2.1. *Participants*

One hundred and one study participants (50 females and 51 males) were recruited through local newspaper/online advertisements and received $40 for their participation. All participants were native English speakers with no uncorrected vision problems. Also, prospective participants who reported they took medication or had a health condition that might influence their performance in the study were excluded. Male and female participants were categorised into two age groups: a 'younger' group consisting of 50 individuals between the ages of 18 and 35 ($M = 26.3$, $SD = 5.0$) and an 'older' group consisting of 51 individuals 55–70 years of age ($M = 62.1$, $SD = 4.8$). We purposely use the terms 'younger' and 'older' to reflect a relative difference in age rather than an absolute classification as 'young' or 'old'. It should further be noted that our 'older' age group might be more appropriately described as 'middle-aged' (see Nichols *et al.* 2003). However, these age ranges were selected to both: (1) represent working age adults and (2) maximise the likelihood of observing age-related differences in warning symbol comprehension. The study procedures were approved by the Liberty Mutual Research Institute for Safety's Institutional Review Board.

2.2. *Apparatus*

Four personal computers (2.8 GHz processor) and 21-in. LCD monitors were used for data collection. The experimental program was developed and executed using E-Prime 2.0 (Psychology Software Tools, Inc.) software. The program measured all valid keyboard responses as well as reaction times (RTs) from stimulus presentation onset to participant keyboard response.

2.3. *Stimuli*

A pilot study was carried out in order to guide the selection of symbols for use in the main study. For this pilot work, 24 younger participants between the ages of 18 and 35 (Mean = 24.5, SD = 5.8) and 24 older participants between the ages of 55 and 70 (Mean = 60.2, SD = 4.0) rated 99 warning symbols on complexity ('How complex is this symbol?' Not at all complex–extremely complex), familiarity ('How familiar are you with this symbol?' Not at all familiar–extremely familiar), and comprehensibility ('How easy would it be to understand this symbol in isolation? That is, without knowing anything else about it, do you think the meaning would be obvious?' Not at all comprehensible–extremely comprehensible). None of these participants took part in the main study. For the comprehensibility ratings, the participants were given the meaning of the symbol in order to make their judgements. A manufacturer of safety labels/signs provided many of these symbols and additional symbols came from Dreyfuss (1984) and Modley (1976). All symbols were in current use and represented a variety of industries including (but not limited to) medical, chemical, construction, laboratory, and manual material handling. The symbols were presented in a 15 cm × 15 cm area and subtended approximately 15° of visual angle (see Figure 1 for sample symbols).

An attempt was made to select symbols in four categories: low complexity–low comprehensibility, low complexity–high comprehensibility, high complexity–low comprehensibility, and high complexity–high comprehensibility. However, it was found that complexity and comprehensibility tended to co-vary such that comprehensibility tended to decrease as complexity increased and vice versa. Therefore, symbols were not selected so as to vary characteristics independently but, instead, were selected to cover a range of familiarity, complexity, and comprehensibility.

Forty-six symbols were selected for use as the experimental stimuli which were tested for comprehension and received accident scenario training. On a scale from 1 to 5 (not at all–extremely), familiarity ranged from 1 to 4.7 (Mean = 2.3, SD = 1.0); complexity ranged from 1.2 to 4.6 (Mean = 2.9, SD = 0.9); and comprehensibility ranged from 1.4 to 4.6 (Mean = 2.8, SD = 1.1). Ten additional symbols served as filler items during the accident scenario portion of the study.

2.4. *Procedure*

Upon arrival, participants read and signed an informed consent form. Participants were seated in front of a computer and were told that the experiment was concerned with how people understand warning symbols. The study consisted of five main sections (administered in successive order): pre-training comprehension, training, demographic questionnaire, post-training comprehension, and symbol characteristic judgements.

2.4.1. *Pre-training comprehension*

Comprehension of the warning symbols was assessed via the semantic relatedness paradigm (Lesch 2003) in which the participant views a symbol paired with a verbal label and is asked to decide whether or not the verbal label conveys the meaning of the symbol.

Participants viewed each symbol twice (resulting in a total of 92 randomly presented trials) – once with a label that conveyed the meaning of the symbol (i.e. the correct label) and another time with a label that did not convey its meaning (i.e. the distractor label). The participant's task was to determine, as quickly as possible, whether the label conveyed the meaning of the symbol by pressing 'Yes' or 'No' on the keyboard. Immediately following their yes/no response, the participants reported their level of confidence in their decision from 1 (not at all confident) to 5 (certain). Distractor labels consisted of incorrect (but plausible) responses given by the participants in a pilot study (see Figure 2).

2.4.2. *Training*

In the training phase of the study, participants viewed each symbol with its verbal label (1500 ms), followed by an accident scenario which described an accident or 'close-call' related to the hazard indicated by the symbol (the display was terminated by the participant's button-press), followed by a second presentation of the symbol with its verbal label (i.e. its referent; 1500 ms). The accident scenarios further elaborated on the nature of the hazard depicted by the symbol, the recommended actions, as well as the possible consequence of failing to perform these actions (see Figure 3). The accident scenarios were derived from accident reports from a number of online sources (see references) including the U.S. Department of Labor Occupational Safety & Health Administration *Accident report fatal facts*, the National Institute for Occupation Safety and Health *Alerts*, the U.S. Department of Labor Mine Safety and Health Administration *Safety hazard alerts*, and the Centers for Disease Control's *Morbidity and mortality weekly report*. Earlier research (e.g. Lesch 2008a) indicated that the presentation of an accident scenario during training provides an additional benefit beyond provision of the verbal label alone. On 10 filler trials,

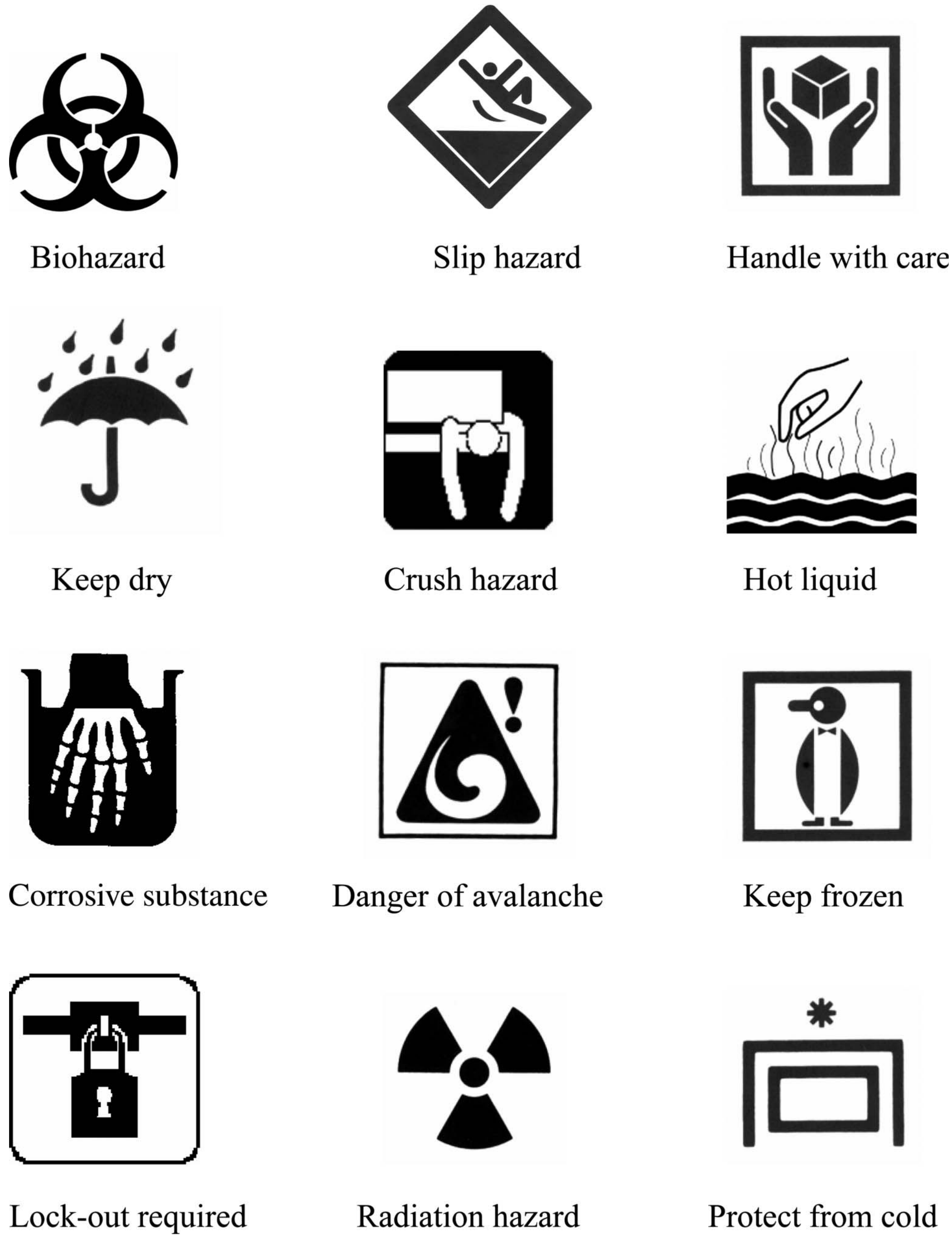


Figure 1. Sample of symbols used in the experiment.

the accident scenarios were followed by a question on the content of the accident scenario to encourage participants to actively read the associated text.

2.4.3. Demographic questionnaire

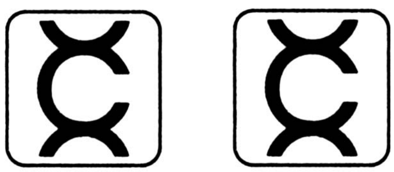
The demographic questionnaire asked participants their age, gender, level of education, and years in the workforce. They were also asked the following work-related accident questions: 'Have you ever worked in a factory, mill, construction site, etc. in which there were serious hazards to life and limb?', 'Have you ever had a work-related accident?', and 'Do you know of anyone who has had a work-related accident?'. The questionnaire was administered immediately following training and prior to the post-test comprehension assessment in order to provide a brief delay between training and post-test and to prevent rehearsal of the to-be-learned information.

2.4.4. *Post-training comprehension*

Participants were tested again on the meaning of the symbols using the same paradigm used in the pre-training comprehension test.

2.4.5. *Symbol characteristics*

Participants were asked to rate each symbol on three dimensions on a 1–5 scale: familiarity ('Before today, how often did you encounter this symbol?' Not at all–very frequently), complexity ('How complex is this symbol?' Not at all complex–extremely complex), and comprehensibility ('Pretend you are seeing this symbol for the first time and you know nothing about it. How likely do you think it is that you could guess that it means [symbol label]?' Not at all likely to guess –extremely likely to guess). Participants first rated all symbols (in random order) on familiarity, followed by complexity, and then comprehensibility.



(a) (b)

Figure 2. Symbol for *cancer-causing substance*. (a) Trial with the correct verbal label. (b) Trial with the distractor label. (a) Cancer causing substance. (b) Check for broken links in chain.

3. Results

3.1. *Accident scenario comprehension*

On average, participants obtained 85.5% (8.6 items out of 10) correct on the questions associated with the accident scenarios on the filler trials suggesting that they were actively reading the accident scenarios. There were no effects of Gender or Age on comprehension performance.

3.2. *Warning symbol comprehension*

To assess warning symbol comprehension, two dependent measures focused on the accuracy of responses: (1) *per cent correct* (where 'correct' was defined as correct acceptance of the target *and* correct rejection of the distractor) and (2) *composite confidence scores* which incorporated confidence ratings together with the comprehension responses. This involved merging the confidence ratings for target and distractor trials into a single measure of performance. Confidence ratings were first transformed to a scale from 1 to 10. For target trials (where the correct answer was 'yes'), 'no' responses

You work at a sewage treatment facility. When you arrived at work the other day, a rescue operation was in progress. Some of your co-workers had been working in a confined space: An empty 180,000 gallon tank. There was only one way in or out of this tank and ventilation was poor. The workers were overcome by toxic fumes from a glue they had been using. Unfortunately, they couldn't be saved in time. They had failed to take basic precautions while working in a confined space: They did not use air respirators, wear safety belts, or post a look-out outside the tank. Any of these precautions might have prevented this tragedy.

Figure 3. Symbol for *confined space* with its associated accident scenario.

were assigned values from 1 to 5 with 'certain-no' responses receiving a score of 1 and 'not at all certain-no' responses receiving a score of 5. 'Yes' responses received scores from 6 to 10 with 'not at all certain-yes' responses receiving a score of 6 and 'certain-yes' responses receiving a score of 10. Therefore, for trials in which the correct answer is 'yes', 'certain-no' responses get the least credit, a 'not at all certain-no' receives somewhat more credit, and 'certain-yes' responses receive the most credit. A similar procedure was used to transform the confidence ratings on the distractor trials (where the correct answer was 'no'). The two transformed scores associated with each symbol were averaged to produce the composite confidence score. Therefore, *comprehension*, as reflected by *composite confidence scores*, is highest when the correct answer is accepted with certainty and the distractor is rejected with certainty. Presumably, the ratings of confidence reflect the strength of the associated knowledge – that is, a *certain* 'yes' should reflect a stronger memory representation/activation than a *somewhat confident* 'yes'. RTs were also analysed.

3.2.1. Percent correct

An analysis of variance (ANOVA) was conducted on percent correct with the within-subject variable Test Session (pre-training and post-training) and the between-subjects variables Age (younger and older) and Gender (male and female). There was a main effect of Test Session, $F(1,97) = 710.23$, $p < 0.05$, partial $\eta^2 = 0.88$, such that participants obtained 73.6% correct following training compared with only 31.6% correct prior to training. There was also a main effect of Age with higher comprehension for the younger adults (56.4% correct) than for the older adults (48.7% correct), $F(1,97) = 9.08$, $p < 0.05$, partial $\eta^2 = 0.09$, but no effect of Gender, $F < 1$. Neither Age nor Gender interacted with Test Session (see Figure 4).

3.2.2. Composite confidence scores

An ANOVA of composite confidence scores indicated a main effect of Test, $F(1,97) = 752.85$, $p < 0.05$, partial $\eta^2 = 0.89$, with a higher mean post-training composite confidence score (Mean = 8.3) compared with the mean pre-training composite confidence score (Mean = 5.9). There was also a main effect of Age with higher comprehension for the younger adults (Mean = 7.5) than for the older adults (Mean = 7.1), $F(1,97) = 7.80$, $p < 0.05$, partial $\eta^2 = 0.07$. While there was no main effect of Gender, $F < 1$, there was a significant Test × Gender interaction, $F(1,97) = 5.18$, $p < 0.05$, partial $\eta^2 = 0.05$, such that females showed a higher rate of improvement from pre-training to post-training than did males, 2.7 vs. 2.4 for females and males, respectively. However, there was no Test × Age interaction ($p = 0.2$) indicating that the rate of improvement was similar for younger and older adults (see Figure 5).

3.2.3. Reaction times

An ANOVA of RTs to make correct decisions indicated a main effect of Test, $F(1,97) = 248.14$, $p < 0.05$, partial $\eta^2 = 0.72$, such that correct RTs were 2.2 s faster following training. There was also a main effect of Age, $F(1,97)$, 38.91, $p < 0.05$, partial $\eta^2 = 0.29$, with younger participants responding 1.5 s faster than the older participants. There was no main effect of Gender and neither Gender nor Age interacted with Test (see Figure 6).

3.2.4. Characteristic ratings

Of greater interest is whether symbol characteristics affected symbol comprehension. Participants rated the

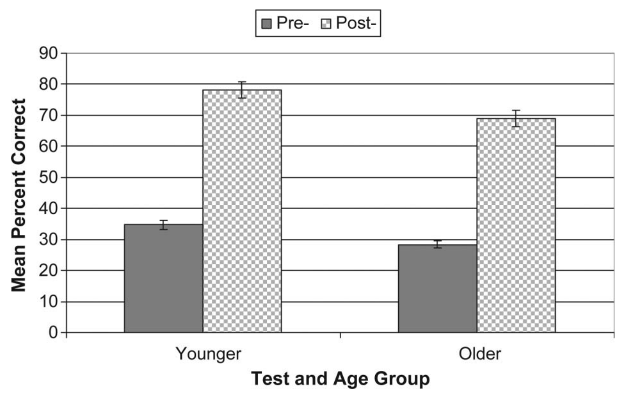


Figure 4. Mean per cent correct and SE as a function of test and age group.

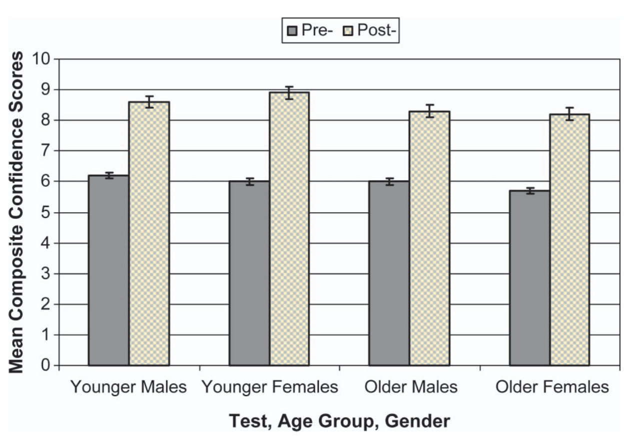


Figure 5. Mean composite confidence scores and SE as a function of test, age group, and gender.

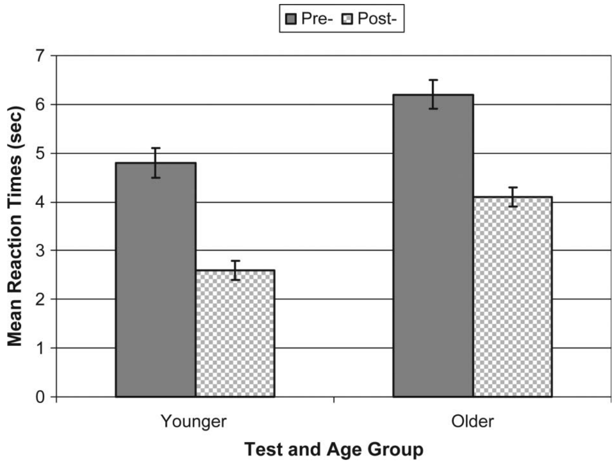


Figure 6. Mean RTs in seconds and SE as a function of test and age group.

Table 1. Mean characteristic ratings (SE) as a function of age group and gender.

	Characteristic		
	Familiarity	Complexity	Comprehensibility
Age–gender			
Younger males	2.5 (.1)	2.9 (.1)	3.1 (.1)
Younger females	2.1 (.1)	2.4 (.1)	3.1 (.1)
	2.3	2.7	3.1
Older males	2.3 (.1)	2.4 (.1)	3.0 (.1)
Older females	1.9 (.1)	2.3 (.1)	3.1 (.1)
	2.1	2.4	3.1

symbols on familiarity, complexity, and comprehensibility. It should be noted that the participants rated the symbol characteristics *after* comprehension testing and training of the symbols. This was done so that the experience of rating the symbols would not influence the initial comprehension measurement. To assess the reliability of the characteristic judgements, Pearson product moment correlations were calculated between the ratings collected during the pilot study and those ratings collected during the main study. The correlation coefficients ranged from .80 to .90 indicating 'good' to 'excellent' consistency in the measurement of characteristic judgements (see Cicchetti and Sparrow 1981, Anastasi and Urbina 1997).

The mean familiarity rating was 2.2 (SD = 0.6) indicating that participants were not very familiar with the symbols. A 2 (older vs. younger) × 2 (male vs. female) ANOVA indicated that males reported being more familiar (Mean = 2.4) with the warning symbols than did the females (Mean = 2.0), $F(1,97) = 15.45$, $p < 0.05$, partial $\eta^2 = 0.14$. However, there was no significant effect of Age (see Table 1).

The mean complexity rating was 2.5 (SD = 0.1) indicating that participants judged the symbols to be moderately complex. A 2 (older vs. younger) × 2 (male vs. female) ANOVA indicated that males rated the symbols as more complex (Mean = 2.6) than did the females (Mean = 2.3), $F(1,97) = 11.65$, $p < 0.05$, partial $\eta^2 = 0.11$. There was also a main effect of Age with younger adults rating the symbols as more complex (Mean = 2.6) than the older adults (Mean = 2.4), $F(1,97) = 10.05$, $p < 0.05$, partial $\eta^2 = 0.09$ (see Table 1).

The mean comprehensibility rating was 3.1 (SD = 0.05). A 2 (older vs. younger) × 2 (male vs. female) ANOVA indicated that mean comprehensibility rating did not vary as a function of Age or Gender, Fs < 1 (see Table 1).

A generalised estimating equation (GEE; Liang and Zeger 1986) regression approach was used in an attempt to disentangle the effects of the different stimulus characteristics on comprehension and training effects (as measured by composite confidence scores). This approach takes into account the dependency of observations within an individual by estimating a working correlation matrix from the residuals. Then, regression coefficients are estimated, correcting for the correlation. The Wald χ^2 is used to assess the significance of the results. This test statistic is conceptually analogous to the F or t statistic and is calculated by dividing a regression coefficient (B) by its standard error, then squaring the quotient (see Liang and Zeger for further information on the GEE approach).

3.3. *Younger adults*

Familiarity and comprehensibility significantly impacted pre-training composite confidence scores for younger adults, Wald $\chi^2 = 16.87$, $p < 0.01$, and Wald $\chi^2 = 49.34$, $p < 0.01$, for familiarity and comprehensibility, respectively. The parameter estimates indicated that, for each unit increase in familiarity, composite confidence scores increased by .56 units and, for each unit increase in comprehensibility, composite confidence scores increased by .64 units. However, there was no significant effect of complexity ($p = 0.08$). Finally, there was a significant interaction between familiarity and comprehensibility such that increasing familiarity resulted in a decreased effect of comprehensibility, Wald $\chi^2 = 11.94$, $p < 0.01$.

A different pattern emerged for the post-training composite confidence scores: None of the stimulus characteristics exerted a significant effect on post-training comprehension by the younger adults suggesting that training was effective in addressing comprehension difficulties associated with these variables.

3.4. *Older adults*

As was the case for younger adults, pre-test composite confidence scores for older adults were significantly impacted by familiarity and comprehensibility of the symbols, Wald $\chi^2 = 19.10$, $p < 0.01$, and Wald $\chi^2 = 26.40$, $p < 0.01$, for familiarity and comprehensibility, respectively. The parameter estimates indicated that, for each unit increase in familiarity, composite confidence scores increased by .21 units and, for each unit increase in comprehensibility, composite confidence scores increased by .42 units. While there was no main effect of complexity, $p = 0.5$, complexity moderated the effect of comprehensibility as indicated by a significant complexity by comprehensibility interaction, Wald $\chi^2 = 7.23$, $p < 0.01$. Increases in complexity resulted in a decreased effect of comprehensibility.

Following training, only comprehensibility significantly influenced composite confidence scores for older adults, Wald $\chi^2 = 12.74$, $p < 0.01$, suggesting that comprehensibility of the symbols is also critical for older adults' learning of warning symbols.

To summarise, as predicted, both familiarity and comprehensibility were important in determining performance on the pre-training comprehension test by both younger and older adults. Furthermore, increasing familiarity served to decrease comprehensibility effects, but only for younger adults. It was hypothesised that, for highly familiar symbols, other stimulus characteristics would be less influential. That is, if the association between the symbol and its meaning is well established for the respondent, *decoding* of the symbol becomes unnecessary. For older adults, on the other hand, the effect of comprehensibility was moderated by complexity rather than familiarity. One interpretation is that the meaning of complex symbols is less apparent (i.e. less comprehensible) for older adults due to a relative inability to focus on the most relevant aspects of the symbol. This interpretation is consistent with the research reviewed earlier indicating that the ability to selectively attend tends to decline with age.

Training served to eliminate the effects of stimulus characteristics for younger adults, while older adults' comprehension continued to be significantly (but less) influenced by comprehensibility of the symbols. Therefore, even the increased exposure to the symbols (i.e. increased familiarity due to training) did not fully counteract the effect of comprehensibility for older adults.

4. Discussion

Age-related effects on warning symbol comprehension have been inconsistently observed in the literature. We suggest here, and elsewhere, that a number of factors could contribute to this pattern of results including comprehension measure, type of symbols, as well as varying definitions of 'younger' versus 'older'. We suggested that age-related changes in cognitive processing have implications for both comprehension of warning symbols and the methods used to assess that comprehension. The semantic relatedness task was used to assess comprehension, since it minimises verbal output (verbal fluency tends to decline with age). We hypothesised that, due to age-related cognitive changes in attention, language, and memory, certain types of symbols would be especially difficult for older adults to comprehend – that is, older adults would have greater difficulty with more complex and less comprehensible symbols.

The current results replicated our earlier findings that older adults have greater difficulty than younger adults in comprehending warning symbols and that accident scenario training improves comprehension, confidence in responses, as well as speed of responding (Figures 4–6). We further found that both familiarity and comprehensibility of the symbols exerted a significant influence on pre-training comprehension for both younger and older adults. Contrary to our predictions, complexity did not significantly influence pre-training comprehension for either younger or older adults, although it did moderate the effect of comprehensibility for older adults. This finding is consistent with the suggestion that age-related decreases in inhibitory efficiency result in decreased comprehensibility. For younger adults, the effect of comprehensibility was moderated by familiarity – for highly familiar symbols for which the meaning is well established for the respondent, comprehensibility is less influential.

We also found that comprehensibility of the symbols was strongly related to comprehension after training, but only for the older adults. This finding is consistent with the research on age-related memory changes that indicates that older adults have greater difficulty in forming new associations (see Luo and Craik 2008) – less comprehensible symbols lack a clear relationship with their real-world referent and tend to be more abstract, or even arbitrary, in that relationship. Our results further suggest that comprehensibility plays an important role in learning by older adults – that is, despite training, comprehensibility continued to predict comprehension performance. For less comprehensible symbols, it is relatively ineffective to provide accident scenarios that further elaborate on the nature of the hazard, the required actions, and the consequences of failures to comply. Therefore, an important question is whether symbol design can somehow facilitate learning of less comprehensible symbols? An example of a symbol design that, if unfamiliar, may appear incomprehensible is the symbol for cancer-causing substance (see Figure 2). However, this symbol incorporates a verbal retrieval cue – the letter 'c' for 'cancer'. We are currently investigating the role that such cues to knowledge may have in addressing initial comprehension, and learning, of less comprehensible symbols by the older adults. We hypothesise that these cues will facilitate the linkage between new knowledge (i.e. the warning symbol) and relevant knowledge in long-term memory. In essence, the cue helps to make the new knowledge 'old', or already known, by providing a clearly recognisable and familiar piece of information that relates the new information to already known information.

A limitation of the current study is that we were unable to vary the symbol characteristics independently in order to more definitively determine their relative contributions. Therefore, some caution must be taken in interpreting the results. However, it was still possible to use regression methods to demonstrate that the relative importance of stimulus characteristics varied as a function of age and training. We are currently exploring the possibility that the amount of information (i.e. complexity) contained in the symbol may not be the critical factor, but, rather, whether or not that information is relevant to determining the meaning of the symbol. If this is the case, then it should be possible to further pull-apart the effects of complexity and comprehensibility by varying the relevance of additional information.

Another potential limitation is that the method of obtaining comprehensibility ratings (i.e. first providing the meaning and then asking participants how likely they would be to guess that meaning) might have introduced a 'hindsight' bias – that is, participants might have over-estimated the likelihood of guessing the symbol's meaning. Fischoff (1975) observed that participants tend to assign a higher likelihood of occurrence to outcomes that they have been told are true. However, it seems unlikely that the relative ratings of symbols would be impacted by this potential bias and, even if this bias was operating in the current study, it would not explain the differing patterns of comprehensibility effects for the younger and older participants.

Chan and Ng (2010, see also Ng and Chan 2011) recently reported the results of another study examining the effects of sign characteristics on training effectiveness but failed to observe effects of familiarity, concreteness, simplicity, meaningfulness, and semantic closeness. These results appear to be in contrast to the results reported in this study. However, Chan and Ng used simple paired-associate learning (symbol-verbal label), recall training, and recognition training. None of these training methods further expanded on the nature of the hazard communicated by the symbols as did the accident scenario training used here. Furthermore, Chan and Ng used a multiple choice test to assess comprehension, whereas we used a semantic relatedness task– as argued earlier the method used to assess comprehension can influence observed comprehension levels. Finally, while some of the characteristics studied by Chan and Ng bear some relationship to the characteristics studied here, there are likely to be differences. Our results suggest that familiarity, complexity, and comprehensibility are closely related and interact with one another to influence comprehension and that their influence on training effectiveness differs across younger and older adults. It will be important for future research to further disentangle the effects of symbol characteristics on comprehension and training so that we can more effectively design warning symbols as well as training programmes to improve their comprehension. However, based on the current results, we are able to make some tentative suggestions regarding symbol design and training: (1) symbols should be visually simple and representational – simple and direct is best; (2) as regards complexity, increased detail should be relevant and necessary in communicating the symbol's meaning; (3) if a symbol must be abstract, it should not be arbitrary (arbitrary symbols are those symbols that have little meaning in and of themselves); (4) however, if the use of abstract or arbitrary symbols cannot be avoided, inclusion of contextual or verbal cues in the design may facilitate initial symbol comprehension as well as increase training effectiveness. While we believe these recommendations are especially important in ensuring older adults' comprehension and memory for warning symbols, our results also indicate that they should benefit younger adults' comprehension as well.

Acknowledgements

The authors would like to thank Angela Garabet for her assistance during the pilot phase of the study; Marvin Dainoff, Yulan Liang, and Glenn Pransky for their comments on earlier versions of this manuscript; and Peg Rothwell for assistance in manuscript preparation.

References

Administration on Aging, Department of Health and Human Services, 2010. *Projected future growth of the older population*. Available from: http://www.aoa.gov/aoaroot/aging_statistics/future_growth/future_growth.aspx#age [Accessed 13 August 2011].

Alain, C., Ogawa, K.H., and Woods, D.L., 1996. Aging and the segregation of auditory stimulus sequences. *Journal of Gerontechnology. Series B, Psychological Sciences & Social Science*, 51, 91–93.

Allen, P.A., Weber, T.A., and Madden, D.J., 1994. Adult age differences in attention: filtering or selection? *Journal of Gerontechnology*, 49, 213–222.

American National Standards Institute (ANSI), 2007. *ANSI Z535.3-2007*. Rosslyn, VA: National Electrical Manufacturers Association.

Anastasi, A. and Urbina, S., 1997. *Psychological testing*. Michigan: Prentice Hall.

Baddeley, A.D., 1986. *Working memory*. London: Oxford University Press.

Centers for Disease Control and Prevention, n.d. *Morbidity and mortality weekly report*. Available from: http://www.cdc.gov/mmwr//mmwrsrch.htm [Accessed 13 August 2011].

Chan, A.H.S. and Ng, A.W.Y., 2010. Effects of sign characteristics and training methods on safety sign training effectiveness. *Ergonomics*, 53, 1325–1346.

Cicchetti, D.V. and Sparrow, S.S., 1981. Development of criteria for establishing the interrater reliability of specific items in a given inventory: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86, 127–137.

Cohen, G., 1979. Language comprehension in old age. *Cognitive Psychology*, 11, 412–429.

Collins, B.L. and Lerner, N.D., 1982. Assessment of fire-safety symbols. *Human Factors*, 24, 75–84.

Dreyfuss, H., 1984. *Symbol sourcebook: an authoritative guide to international graphic symbols*. New York: John Wiley & Sons.

Easterby, R.S. and Hakiel, S.R., 1981. Field testing of consumer safety signs: the comprehension of pictorially presented messages. *Applied Ergonomics*, 12, 143–152.

Fischoff, B., 1975. Hindsight ≠ foresight: the effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288–299.

Hancock, H.E., Rogers, W.A., and Fisk, A.D., 1999. Understanding age-related differences in the perception and comprehension of symbolic warning information. *In*: *Proceedings of the HFES 43rd annual meeting*. Santa Monica, CA: Human Factors and Ergonomics Society, 617–621.

Hancock, P.A., Lesch, M., and Simmons, L., 2003. The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention*, 35, 501–514.

Hancock, H.E., *et al*., 2004. Safety symbol comprehension: effects of symbol type, familiarity, and age. *Human Factors*, 46, 183–195.

Hasher, L., *et al*., 1991. Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17 (1), 163–169.

Jost, K., *et al*., 2010. Are old adults just like low working memory young adults? Filtering efficiency and age differences in visual working memory. *Cerebral Cortex*. first published online: 30 September 2010.

Kane, M.J., *et al*., 1994. Inhibitory attentional mechanisms and aging. *Psychology and Aging*, 9 (1), 103–112.

Lesch, M.F., 2003. Comprehension and memory for warning symbols: age-related differences and impact of training. *Journal of Safety Research*, 34, 495–505.

Lesch, M.F., 2004. Age-related effects on warning symbol comprehension. *In*: *Proceedings of the HFES 48th annual meeting*. Santa Monica, CA: Human Factors and Ergonomics Society, 233–237.

Lesch, M.F., 2005. A semantic relatedness paradigm for assessing comprehension of warning symbols. *In*: *Proceedings of the HFES 49th annual meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Lesch, M.F., 2008a. A comparison of two training methods for improving warning symbol comprehension. *Applied Ergonomics*, 39, 135–143.

Lesch, M.F., 2008b. Warning symbols as reminders of hazards: impact of training. *Accident Analysis & Prevention*, 40, 1005–1012.

Liang, K.Y. and Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13–22.

Lindenberger, U., Marsiske, M., and Baltes, P.B., 2000. Memorizing while walking: increase in dual-task costs from young adulthood to old age. *Psychology and Aging*, 15, 417–436.

Luo, L. and Craik, F.I.M., 2008. Aging and memory: a cognitive approach. *Canadian Journal of Psychiatry*, 53, 346–353.

Mayer, D.L. and Laux, L.F., 1989. Recognizability and effectiveness of warning symbols and pictorials. *In*: *Proceedings of the Human Factors Society 33rd Annual Meeting*. Santa Monica, CA: Human Factors Society, 984–988.

McCalley, L.T., Bouwhuis, D.G., and Juola, J.F., 1995. Age changes in the distribution of visual attention. *Journals of Gerontechnology. Series B, Psychological Sciences & Social Sciences*, 50, 316–331.

McCrae, R.R., Arenberg, D., and Costa, P.T., 1987. Declines in divergent thinking with age: cross-sectional, longitudinal, and cross-sequential analyses. *Psychology and Aging*, 2, 130–137.

Modley, R., 1976. *Handbook of pictorial symbols: 3,250 examples from international sources*. New York: Dover Publications.

Morrell, R.W., Park, D.C., and Poon, L.W., 1990. Effects of labeling techniques on memory and comprehension of prescription information in young and old adults. *Journal of Gerontology*, 45, 166–172.

National Institute for Occupational Safety and Health, n.d. *Alerts*. Available from: http://www.cdc.gov/niosh/alerts2.html.

Nichols, T.A., Rogers, W.A., and Fisk, A.D., 2003. Do you know how old your participants are? Recognizing the importance of participant age classifications. *Ergonomics in Design*, 11, 22–26.

Ng, A.W. and Chan, A.H., 2007. The guess ability of traffic signs: effects of prospective-user factors and design features. *Accident Analysis and Prevention*, 39, 1245–1257.

Ng, A.W. and Chan, A.H., 2011. Investigation of the effectiveness of traffic sign training in terms of training methods and sign characteristics. *Traffic Injury Prevention*, 12, 283–295.

Pratt, M.W., *et al*., 1989. Telling tales: aging, working memory, and the narrative cohesion of story retellings. *Developmental Psychology*, 25, 628–655.

Rissenberg, M. and Glanzer, M., 1987. Free recall and word finding ability in normal aging and senile dementia of the Alzheimer's type: the effect of item concreteness. *Journal of Gerontology*, 42, 318–422.

Salthouse, T.A., 1985. *A Theory of cognitive aging*. Amsterdam, The Netherlands: Elsevier Science.

Salthouse, T.A., *et al*., 1989. The effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 507–516.

U.S. Department of Labor, Mine Safety and Health Administration, n.d. *Safety hazard alerts*. Available from: http://www.msha.gov/ [Accessed 13 August 2011].

U.S. Department of Labor, Occupational Safety and Health Administration, n.d. *Accident report fatal facts*. Available from: http://www.osha-slc.gov/OshDoc/toc_FatalFacts.html [Accessed 13 August 2011].

Wolff, J.S. and Wogalter, M.S., 1998. Comprehension of pictorial symbols: effects of context and test method. *Human Factors*, 40, 173–186.

Zacks, R.T., Radvansky, G., and Hasher, L., 1996. Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 143–156.

Zwaga, H.J. and Boersema, T., 1983. Evaluation of a set of graphic symbols. *Applied Ergonomics*, 14, 43–45.