Applied Ergonomics 45 (2014) 1285–1290

Contents lists available at SciVerse ScienceDirect

Applied Ergonomics

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

Facilitating pictorial comprehension with color highlighting

Brannan R. McDougald*, Michael S. Wogalter

North Carolina State University, Psychology Department, Human Factors and Ergonomics Program, Raleigh, NC 27695-7650, USA

ARTICLE INFO

Article history:
Received 20 July 2012
Accepted 8 May 2013

Keywords:
Warnings
Highlighting
Color

ABSTRACT

Pictorials can aid in communicating warning information, but viewers may not always correctly comprehend them. Two experiments focused on whether the use of relevant highlighting could benefit pictorial comprehension. A set of warning-related pictorials were manipulated according to three-color highlighting conditions: highlighting areas more relevant to correct comprehension, highlighting areas less relevant to comprehension, and no highlighting. Participants were asked to describe the purpose and meaning of each pictorial presented to them. The findings from both experiments indicate that comprehension of warning pictorials is higher for the relevant highlighting condition than the other two conditions. The highlighting of less relevant areas reduced comprehension compared to no highlighting. Use of appropriately placed highlighting could benefit the design of a complex symbol by pointing out pertinent areas to aid in determining its intended conceptual meaning.

1. Introduction

Understanding safety information concerns, products, and equipment is important for their proper use (Wogalter, 2006). Communicating safety information has become increasingly challenging as people speaking different languages intermix through travel and trade. Illiteracy can increase this challenge. To overcome these and related difficulties, symbols offer a potential way to address the language barrier as a sort of "universal" language. However, a common finding in research on symbol comprehension is that the intended communication may not be understood properly (e.g., Hancock et al., 1999). Worse yet, some symbols in certain contexts can confuse the viewer by conveying the wrong information (e.g., Zwaga and Boersema, 1983). Nevertheless effective symbols seem to offer benefits. In addition, symbols appear to be useful in attracting attention, which is important to processing safety communications (Bzostek and Wogalter, 1999; Laughery et al., 1993).

Because of difficulties in understanding symbols, domestic and international standards organizations have produced methods and criteria to assess comprehension adequacy of symbols. The American National Standard Institute's (ANSI) Criteria for Safety Symbols requires 85 percent correct in a comprehension test with a sample of 50 people reporting the intended concept. Criteria include that no more than five percent of the sample may experience critical confusion (an opposite or very wrong response; ANSI Z535.3, 2011). The International Organization for Standardization (ISO) also has a set of guidelines for symbol comprehension (ISO 9186, 2007). ISO and ANSI are in the process of harmonizing their symbol guidelines.

Designing symbols that meet ANSI criteria is reportedly difficult (e.g., Davies et al., 1998; Hancock et al., 2004; Zwaga and Boersema, 1983). To aid designers in creating symbols that meet these criteria, the human factors literature offers strategies and methods to increase comprehension (e.g., Collins and Lerner, 1982; Easterby and Hakiel, 1981; Hancock et al., 2004; Wogalter et al., 2006). For example, one strategy to influence symbol comprehension is to enhance legibility — the visual clarity of a symbol (Wogalter et al., 2002). Traditionally, symbol simplicity is preferred to enhance legibility (Wogalter et al., 2006) resulting in the use of bold lines in lieu of fine lines or details.

Some guidelines (e.g., FMC, 1985; Westinghouse Electric Corporation, 1981) and standards (e.g., ANSI Z535.3, 2011; ISO 3864-1, 2003) suggest the use of simple symbols for warnings. However, these suggestions may be inappropriate in situations necessitating specific and complex communication of information. When symbols lose increasingly more information due to simplification, the symbol may lack the necessary information for individuals to interpret its intended meaning. Fig. 1 provides an example of a symbol that appears to lack enough information to attain a high level of comprehension. So sometimes introducing increased information or detail may benefit symbols, which is contrary to conventional design strategies emphasizing minimal detail. Detail may be necessary for some symbols to meet high levels of comprehension. The additional information may aid the

* Corresponding author. Tel.: +1 919 306 2964; fax: +1 919 515 1716.
E-mail address: brmcdoug@ncsu.edu (B.R. McDougald).

0003-6870/$ – see front matter
http://dx.doi.org/10.1016/j.apergo.2013.05.008

Fig. 1. Symbol of a dam.

viewer in comprehending situational considerations, avoidance strategies, and consequences related to the symbol. Fig. 1 would likely be benefitted by at least some additional information within the symbol, such as more specifics about the dam and surrounding environment or a different perspective of the structure. One of the main reasons that guidelines on warnings mandate limited detail is that lesser important information (e.g., detail) may capture and hold attention to the detriment of attention to relevant details and the determination of the symbol's intended meaning (Wogalter et al., 2006). Symbols designed with minimal complexity are differentiated from pictorials, which are symbols designed with greater amounts of detail and information. Pictorials might benefit from prominent aspects or characteristics that direct the viewer's attention to the most relevant information.

Salience is a stimulus-driven or bottom–up process whereby physical characteristics tend to "pop-out" from the context and seemingly to stand out effortlessly (Fecteau and Munoz, 2006; Yantis, 2000). The addition of color is one method to increase salience (e.g., Itti et al., 1998; Peters et al., 2005). Colored backgrounds are commonly used in warning pictorials to attract viewer's attention from environmental stimuli to the warning (ISO 3864-1, 2003). Highlighting a limited area of the pictorial with color could be used to increase the salience of the most relevant details of the pictorial. Focusing attention through highlighting could reduce the potential adverse effects of introducing greater pictorial complexity.

The use of highlighting to direct viewer's attention has been demonstrated in previous research in other areas. For example, Wickens et al. (2004) investigated the utility of varying intensity of highlighting in the performance of map reading tasks by increasing the salience and discriminability of stimuli. Martin et al. (1987) found that highlighting can increase performance of visual search with minimal performance cost to ignoring incorrectly highlighted stimuli. Wu and Yuan (2003) demonstrated the superiority in reading times of traditional, color highlighting in comparison to no highlighting, and other, non-traditional forms of highlighting (e.g., flashing text). In a review of color coding research, Christ (1975) concluded that color facilitates identification and searching of objects, particularly when the color is known to be uniquely associated with the target.

Potential harmful effects of highlighting have also been observed (e.g., Fisher and Tan, 1989; Tamborelloii and Byrne, 2007). In visual search tasks, when highlighting is placed over the target stimuli, search time decreases. However, when highlighting is placed over a distracter, search time increases. Whether highlighting is placed over the target or a distracter, it is drawing the viewer's attention toward that area. Additionally, inappropriate highlighting has been shown to interfere with text comprehension when placed on less relevant or irrelevant information (e.g., Gier et al., 2009). The harmful effects of color were also identified by Christ's (1975) review. For example, when color is added to distracters, the accuracy of identifying features of targets without color decreases.

Two experiments described in this report examine whether color highlighting can improve pictorial comprehension performance. Two highlighting conditions are examined: relevant and irrelevant. In the relevant highlighting condition, a portion highly pertinent or relevant to the intended meaning of the pictorial is overlaid with the color yellow. It was expected that highlighting would enhance pictorial comprehension compared to the same pictorials with no highlighting. This enhancement could be explained by the highlighting focusing viewers' attention to pertinent aspects of the pictorial compared to the absence of relevant highlighting. In the irrelevant condition, a portion minimally pertinent or relevant to the intended meaning of the pictorial is overlaid with the color yellow. It was expected that this highlighting would diminish pictorial comprehension compared to the same pictorials with no highlighting. This diminishment in performance could be explained by the highlighting focusing viewer's attention away from pertinent aspects of the pictorial compared to the absence of irrelevant highlighting. Irrelevant highlighting may misguide viewers' attention resulting in lower comprehension. The second experiment differed from the first by sampling from a different population and addressing methodological issues.

2. Experiment 1

Pictorial comprehension is investigated comparing relevant and irrelevant highlighting to no highlighting.

2.1. Method

2.1.1. Participants

Eighty-four North Carolina State University undergraduate students (mean age = 18.7 years, SD = 1.2) participated as part of a course requirement for an introductory psychology course. The sample was comprised of 35 males and 49 females.

2.1.2. Experimental design and stimuli

Participants completed an online questionnaire with 13 pictorials. Each pictorial was accompanied with a short statement briefly describing the context where it might be located. Three of the pictorials were specious and intended to help disguise the purpose of the study. Responses to the specious pictorials were not analyzed. The other 10 pictorials were experimentally manipulated to form three conditions. The three experimental conditions were no highlighting, less relevant (or irrelevant) highlighting and relevant highlighting. To form these conditions, a set of "base" pictorials was first produced. All of these had no highlighting and served as stimuli in the no highlighting condition. From the base non-highlighted pictorials the other two experimental conditions were produced. To form the relevant highlighting set of pictorials yellow was added that encircled and covered the most pertinent portion(s) of the pictorial in order to determine its meaning. This area was generally the focal point of the interaction between the human figure and the instructed action. The less relevant (irrelevant) highlighting added color to the base pictorials that did not cover the focal point of the interaction between the human figure and the instructed action. This area included any distinct object within the pictorial that did not overlap with the relevant highlighting portion. For examples, see Figs. 3 and 4 shown later in this article. Highlighting was always accomplished using a highly saturated yellow hue. (Note: the pictorials shown in the figures as illustrative examples use gray highlighting in lieu of yellow.)

The specific conceptual meanings (referents) of the ten manipulated pictorials were: (1) hold on with your hand to ladder while

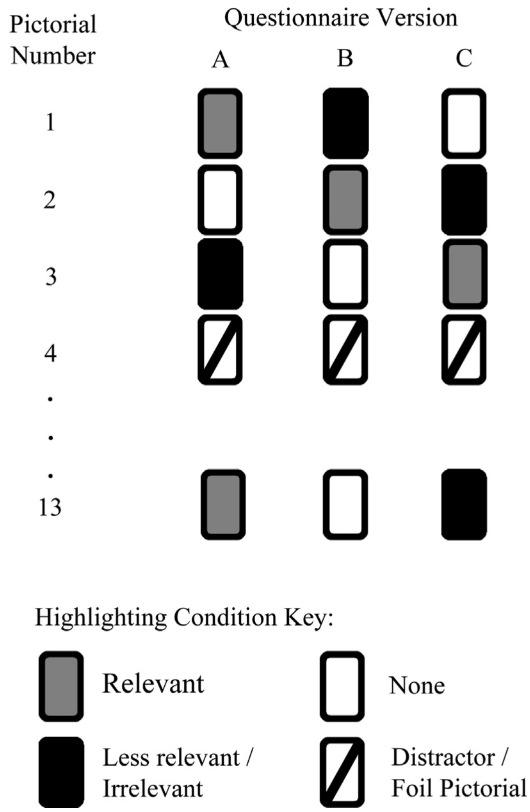

Fig. 2. Three versions of the pictorials. Note: each participant completed the task with one version.

keeping both feet on it, (2) use safety clamps on barbell when lifting weights, (3) do not stand under ladder while holding it, (4) keep wrists straight and lifted while typing, (5) keep medication out of reach of children, (6) check mirrors while driving, (7) hold onto something while standing in a bus, (8) look both ways before crossing a road, (9) do not overload electrical plugs, and (10) turn pan handles inward while on stove. Thus from the 10 base pictorials, 30 experimental pictorials were produced. Each base pictorial had a representative in each of the three highlighting conditions.

The study was a within-subjects design. In an attempt to control for pictorial specific effects, all base pictorials were presented in each of the highlighting conditions across three different versions of the task. Each base pictorial was presented only once in each of the versions. Each version included pictorials in all three highlighting conditions. See Fig. 2 for a graphical representation of how pictorials were presented to participants. Because there were 10 pictorials presented in three conditions, there was an unequal distribution of pictorials in each condition. Each participant saw three pictorials in two conditions and four pictorials in the other

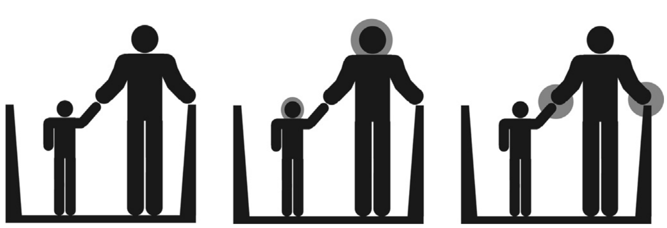

Fig. 3. Pictorial referent is "hold on to side and hold child's hand while on moving walkway". Yellow highlighting is shown in gray. Order of conditions from left to right: no highlighting, less relevant highlighting, relevant highlighting.

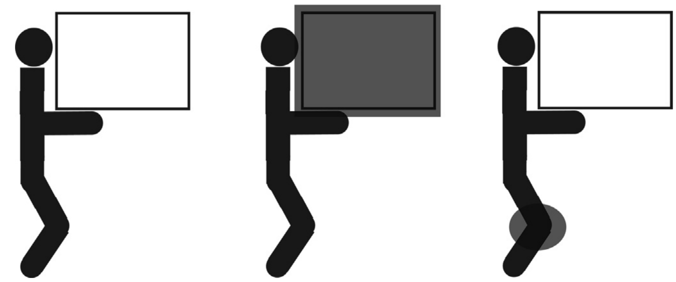

Fig. 4. Pictorial referent is "bend at the knee". Yellow highlighting is shown in gray. Order of conditions left to right: no highlighting, less relevant highlighting, relevant highlighting.

condition. However, pictorials and conditions were fully counterbalanced with equal numbers of participants so that there was no bias due to the unequal numbers of pictorials in the three conditions that any given participant observed. A box was provided adjacent to each pictorial and contextual description for participants to type in their responses describing their understanding of each pictorial's meaning.

A demographic questionnaire asked participants about their age, gender, education level attained and ethnicity.

2.1.3. Procedure

Participants signed up to take part in the study via a web link that navigated them to an online questionnaire. At that point participants were assigned to one of three versions of the pictorials using block random assignment. All participants saw a sequence of 13 pictorials. Ten were experimentally manipulated pictorials in each of the three highlighting conditions. Three were specious pictorials intended to be fillers that were not manipulated or scored. Participants read a short statement describing context in which the pictorial might be viewed and viewed the pictorial. Participants were asked to type into a dialog box adjacent to the pictorial on what they believed the pictorial was attempting to communicate. Participants were encouraged to give as detailed and specific descriptions that they could. After completing each response, they continued to the next pictorial until they completed the entire set. Upon completing their responses to the pictorials, participants completed a short demographic questionnaire.

Two independent raters evaluated participants' responses. Raters compared the conceptual meanings to participants' responses to determine correctness: correct (assigned a score of 1) or incorrect (assigned a score of 0). Responses needed to reasonably reflect the intended meaning of the referent. It did not have to match a verbatim description. For example, the pictorial with the meaning "hold on with your hand to a ladder while keeping both feet on it" could be correctly answered by a number of conceptually similar responses such as "maintain three points of contact with the ladder at all times." General responses such as "use ladder properly" were not counted as correct responses.

Table 1
Mean frequency and proportion of correct responses as a function of highlighting condition (Experiment 1).

Highlighting condition	Frequency	Proportion
	M (SD)	*M* (SD)
No	1.07 (0.91)	0.32 (0.26)
Less relevant (Irrelevant)	0.47 (0.67)	0.15 (0.22)
Relevant	2.07 (1.10)	0.61 (0.30)

$N = 84$.

2.2. Results

For each participant, the number of correct interpretations of pictorials was calculated for each of the three conditions. Between the two independent raters, inter-rater reliability was 0.92, and Cohen's Kappa, a measure of nominal scale response agreement between two raters by taking into account chance agreement between raters (Cohen, 1960) was 0.84. Descriptive statistics for number of correct responses per participant are shown in Table 1 under "frequency". The total number of correct responses in relation to the total number of possible responses is provided under "proportion" in Table 1. An ANOVA on the proportion data indicated the use of highlighting (no highlighting, less relevant highlighting, relevant highlighting) showed a significant effect of the conditions, $F(1.73, 142.17) = 68.29$, MSE = 0.92, $p < 0.001$, $\eta_p^2 = 0.45$. Comparisons among means using the Bonferroni correction showed that comprehension was higher for relevant highlighting than no highlighting or less relevant highlighting. Also, comprehension was lower for less relevant highlighting than for no highlighting (all $ps < 0.02$).

2.3. Discussion

The data show that relevant color highlighting assisted comprehension of the pictorials compared to no highlighting. Higher pictorial comprehension for relevant highlighting is probably due to viewers' focusing visual attention on the most relevant information. Additionally, less relevant or irrelevant highlighting reduced comprehension of the pictorials probably because this highlighted information captured visual focus on the wrong information leading to confusion. If relevant portions of the pictorial are highlighted then there are comprehension benefits. However, if the highlighted area is less or not relevant, then it may focus attention away from pertinent information and lead interpretation astray. Thus, highlighting appears to be either beneficial or detrimental depending on the relevance of the areas highlighted to its intended meaning.

There were two methodological difficulties; neither difficulty was major but worth pointing out. The first concerned the number of pictorials manipulated experimentally. There was an imbalance in the number of pictorials in each condition each participant viewed (e.g., three in two conditions and four in one condition). The unequal number of pictorials in each condition required that a more complex experimental design be used in which extra care was necessary to have all three variations of the pictorials (relevant, irrelevant, and no highlighting) to be seen by an equal number of participants by rotating the imbalanced sets across conditions and participant groups. Second, the participants were all undergraduate students. Although they were young adults majoring in a broad range of subjects, they are a narrow demographic compared to the broader population expected to interpret pictorials. A second experiment was conducted to address these concerns.

3. Experiment 2

Experiment 2 replicates Experiment 1 with three exceptions. One is that in Experiment 2 every participant viewed the same number of pictorials in each highlighting condition. This improvement was aimed to address the imbalance mentioned above in Experiment 1. Experiment 2's experimental design was simplified compared to Experiment 1 by increasing the number of pictorials so there were an equal number of pictorials in each of the three conditions. Second, some of the pictorials were redrawn or modified slightly based on comments from a group of persons with a human factors/ergonomics background asked to give feedback. Third, Experiment 2 included a more diverse group of sampled participants than the undergraduate participants from Experiment 1.

3.1. Method

3.1.1. Participants

A total of 207 persons (71 males and 136 females) participated. Participants' reported ages ranged from 18 to 81 with a mean of 34.1 years (SD = 12.0). Participants represented a variety of occupations and a broad range of education levels from earning a high school or equivalent degree to completing graduate or professional school. Participants were recruited through Mechanical Turk (MTurk; Amazon.com, 2010). MTurk is an online marketplace that utilizes crowdsourcing to accomplish tasks by participants representing a wider population sample than undergraduate students (e.g., Paolacci et al., 2010). Behrend et al. (2011) found MTurk participants to be older, more ethnically diverse, and to possess more experience working than undergraduate students. All participants were reportedly U.S. residents and compensated at a level slightly above the median MTurk wage at $1.38 per hour (Horton and Chilton, 2010).

3.1.2. Experimental design and stimuli

The set of ten pictorials from Experiment 1 served as the starting point for the development of the pictorials in Experiment 2. Two additional pictorials were created to balance the number of pictorials per condition. Six Human Factors subject matter experts (SMEs) were provided with printouts of the pictorials from Experiment 1 along with the two additional pictorials accompanied by a short description of where the warning could be posted. The SMEs gave feedback pertaining to design recommendations for the purpose of maximizing comprehension by viewers. After initial feedback was collected, it was followed by a short discussion of each pictorial to clarify its intended meaning and any additional suggestions were solicited. Revisions to three of the pictorials were made based on this feedback.

This process resulted in 12 "base" pictorials each pertaining to a unique intended meaning that functioned in the "no highlighting" condition. The referents of the 12 manipulated base pictorials were: (1) hold on with your hand while keeping both feet on the ladder, (2) use safety clamps on barbell when lifting weights, (3) do not support ladder from underneath, (4) look both ways before crossing, (5) keep out of reach of children, (6) hold on while standing on bus, (7) do not use weights on exercise ball, (8) turn handles inward, not outward on a stove, (9) do not overload outlets/plug in too many electric devices, (10) hold rail and child's hand while on moving walkway, (11) bend at the knees, and (12) secure the spreaders on ladder. See Figs. 3 and 4 for examples of pictorials in the three conditions.

From the pictorials without highlighting, the other two experimental conditions were produced. The principles used to highlight the relevant areas and less relevant or irrelevant areas of the pictorials were the same as those described in the Method section of Experiment 1. A total of 36 pictorial-condition combinations were created. According to a 3 × 3 Latin Square, the three sets of pictorials were combined with three groups of participants. Each set of pictorials contained one of each of the 12 base pictorials. The sets differed in which pictorials were in which highlighting condition. Each of the three sets included 12 unique pictorial-condition combinations and four pictorials for each condition. Thus, each group of participants viewed 12 pictorials with four pictorials in each of the three highlighting conditions. Participants were randomly assigned to three groups that saw different sets of pictorials in the highlighting condition. No participant saw the same pictorial in more than one condition, and all pictorials appeared an equal number of times across conditions and participants. The same rating procedure from Experiment 1 was utilized.

3.2. Results

Inter-rater reliability was 0.90, and Cohen's Kappa (Cohen, 1960) was 0.79. Descriptive statistics for mean frequency of correct responses per participant are shown in the middle column of Table 2. The proportion correct relative to the total number of responses is provided in the right-most column of the table. An ANOVA conducted on the proportions correct showed a significant effect of the conditions, $F(2, 412) = 192.85$, $MSE = 0.07$, $p < 0.001$, $\eta_p^2 = 0.48$. Comparisons among means corrected for by the Bonferroni method showed that comprehension was higher for relevant highlighting than no highlighting or less relevant highlighting. Also, comprehension was lower for less relevant highlighting than for no highlighting (all $ps < 0.02$). An ANOVA and follow-up comparisons conducted on the frequency data instead of proportional data as described above produced similar results.

3.3. Discussion

The data support the notion that highlighting relevant portions of pictorials assisted viewers' comprehension. Salience of particularly important aspects in pictorials apparently serves to focus viewers' attention aiding in the comprehension of the intended meaning. The no highlighting condition, lacking focal attention cues, resulted in lower comprehension performance than relevant highlighting. Comprehension of pictorials was reduced when less relevant or irrelevant information was highlighted. Noteworthy is that performance in this condition was lower than the no highlighting condition. Thus, irrelevant highlighting was harmful as shown by its negative effect on comprehension.

4. General discussion

Both experiments showed that highlighting relevant portions of pictorials can assist in their comprehension. Although simple symbols are generally desirable, sometimes greater detail is necessary to distinguish and elaborate content of a pictorial. However, additional detail may include a cost of decreased comprehension. Highlighting may direct or focus the viewer's attention to the most relevant parts of the pictorials leading to better comprehension. Highlighting could, in effect, de-clutter or perceptually simplify the pictorial. Additionally, highlighting may enable the communication of small differences in conceptual meaning to be conveyed that distinguish between similar concepts. Adding relevant highlighting may do this by reducing scanning time associated with the examination of a detailed and complex graphic by helping to direct attention to relevant detail.

Additionally, irrelevant highlighting was demonstrated in both experiments to be harmful to comprehension of intended meanings. Reduced performance for pictorials with less relevant or irrelevant information highlighting suggests that designers should exercise caution regarding what information is highlighted. Methods to determine how best to identify regions of a pictorial to be highlighted could be demonstrated in future research.

In the two experiments, different samples of the population were used. Experiment 1 used university students, and Experiment 2 used a more diverse sample of participants from the Internet. Fig. 5 shows a comparison of the proportion of correct responses between the two experiments. The pattern of findings across conditions in both studies is similar.

An application of where relevant highlighting may be useful is for pictorials that fail to acquire adequate levels of comprehension, such as less than 85 percent correct or have more than 5 percent critical confusion errors (ANSI Z535.3, 2011). These "failed" pictorials might benefit from greater detail and relevant highlighting. Some failed pictorials may only require relevant highlighting and thus save in the redesign costs of the pictorial. Highlighting relevant details can be a "tool" in a designers' toolbox to reduce development time while also raising the comprehension level.

Further research on the effectiveness of pictorial highlighting is warranted. One area that might profit from additional research pertains to older adults having greater difficulties than younger adults in understanding pictorials (e.g., Hancock et al., 2001; Zwaga, & Boersema, 1983). These difficulties are associated with the age-related cognitive and perceptual declines (e.g., Lesch et al., 2011; Rogers and Fisk, 2001; Salthouse, 1996), such as the inhibition of irrelevant information (e.g., Hasher et al., 1991; Hasher and Zacks, 1988; Madden et al., 1996), increases in visual search times (e.g., Plude and Hoyer, 1986), and decreases in visual acuity (e.g., Gittings and Fozard, 1986; Watanabe et al., 1994). These decrements suggest that older adults experience greater difficulty in tasks related to the comprehension of more complex pictorials. Older adults may demonstrate substantial improvement in complex pictorial comprehension with relevant highlighting when inhibition of irrelevant information and visual search are the sources of the difficulties. Additionally, Lesch et al. (2011) emphasize that comprehensibility, the ease with which a pictorial isolated from its typical environment can be understood by a viewer, should not suffer at the expense of decreasing complexity. Highlighting will likely not address decrements in visual acuity, so the increased detail that highlighting affords may compound the issue for older adult's pictorial comprehension. Regarding improperly placed highlighting, it might be expected that older adults would be more adversely affected than younger adults.

Several pictorials in the present study were relatively complex. "Hold on with your hand while keeping both feet on the ladder" and "hold child's hand and rail on moving walkway" are examples demonstrating multiple instructions. The ladder pictorial included instructions about (1) your hand and (2) your feet. The child on the

Table 2
Mean frequency and proportion of correct responses as a function of highlighting condition (Experiment 2).

Highlighting condition	Frequency	Proportion
	M (SD)	*M* (SD)
No	1.51 (1.19)	0.38 (0.30)
Less relevant (Irrelevant)	0.70 (0.78)	0.17 (0.20)
Relevant	2.65 (1.14)	0.66 (0.28)

$N = 207$.

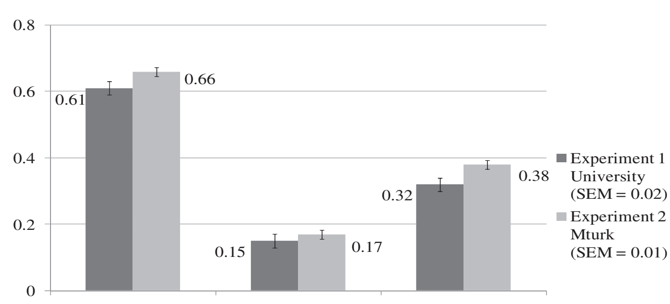

Fig. 5. Comparison between university and MTurk participants' mean proportion of correct responses as a function of highlighting condition. (SEM = Standard Error of the Mean).

walkway pictorial included instructions about (1) holding the child's hand and (2) holding the rail. In the no highlighting condition, participants seemed more likely to miss one of the instructions in a multi-instruction pictorial, such as the two previous examples, in comparison to the highlighting condition. Future research could investigate the validity of these particular findings.

Research suggests that people can perform more quickly in visual search tasks with colored icons that indicate an area of important warning text (Bzostek and Wogalter, 1999; Huang et al., 2008). Highlighting relevant regions appears to hasten processing by directing the viewer to the relevant information. Verifying this difference in processing speed is a question for future research. Among the kinds of additional questions that could be addressed include: (a) how many distinct colors may be used, (b) how many areas can be highlighted, (c) what other highlighting methods besides color could be beneficial, and (d) do special considerations need to be made for color-blind individuals?

Acknowledgments

The authors thank Ann Dixon for her help in evaluating responses. We would also like to thank the reviewers for their helpful comments. A portion of this research was given at the 55th Annual Human Factors and Ergonomics Society meeting in Las Vegas, Nevada, USA (McDougald and Wogalter, 2011).

References

Amazon.com, 2010. Mechanical Turk. Retrieved from: http://www.mturk.com/.
American National Standards Institute (ANSI), 2011. Criteria for Safety Symbols. ANSI Z535.3-2011. National Electrical Manufacturers Association, Rosslyn, VA.
Behrend, T.S., Sharek, D.J., Meade, A.W., Wiebe, E.N., 2011. The viability of crowdsourcing for survey research. Behav. Res. Meth.. Retrieved from: http://www.springerlink.com/content/hkx2248321308124/.
Bzostek, J.A., Wogalter, M.S., 1999. Measuring visual search time for a product warning label as a function of icon, color, column and vertical placement. In: Proceedings of the Human Factors and Ergonomics Society 43nd Annual Meeting. Human Factors and Ergonomics Society, Houston, Texas, pp. 888–892.
Christ, R.E., 1975. Review and analysis of color coding research for visual displays. Hum. Factors 17 (6), 542–570.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46.
Collins, B.L., Lerner, N.D., 1982. Assessment of fire-safety symbols. Hum. Factors 24, 75–84.
Davies, S., Haines, H., Norris, B., Wilson, J.R., 1998. Safety pictograms: are they getting the message across? Appl. Ergon. 29, 15–23.
Easterby, R., Hakiel, S., 1981. Field testing of consumer safety signs: the comprehension of pictorially presented messages. Appl. Ergon. 12, 143–152.
Fecteau, J.H., Munoz, D.P., 2006. Salience, relevance, and firing: a priority map for target selection. Trends Cogn. Sci. 10, 382–390.
Fisher, D.L., Tan, K.C., 1989. Visual displays: the highlighting paradox. Hum. Factors 31, 17–30.
FMC, 1985. Product Safety Sign and Label System. FMC Corporation, Santa Clara, California.
Gier, V.S., Kreiner, D.S., Natz-Gonzalez, A., 2009. Harmful effects of preexisting inappropriate highlighting on reading comprehension and metacognitive accuracy. J. Gen. Psychol. 136, 287–300.
Gittings, N.S., Fozard, J.L., 1986. Age changes in visual acuity. Exp. Gerontol. 21, 423–434.
Hancock, H.E., Rogers, W.A., Fisk, A.D., 1999. Understanding age-related differences in the perception and comprehension of symbolic warning information. In: Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting. Human Factors Society, Santa Monica, CA, pp. 617–621.
Hancock, H., Rogers, W., Fisk, A., 2001. An evaluation of warning habits and beliefs across the adult life span. Hum. Factors 43, 343–354.
Hancock, H., Rogers, W., Schroeder, D., Fisk, A., 2004. Safety symbol comprehension: effects of symbol type, familiarity, and age. Hum. Factors 46, 183–195.
Hasher, L., Stoltzfus, E.R., Zacks, R.T., Rypma, B., 1991. Age and inhibition. J. Exp. Psychol. Learn. Mem. Cogn. 17, 163–169.
Hasher, L., Zacks, R.T., 1988. Working memory, comprehension, and aging: a review and a new view. In: Bower, G.H. (Ed.), 1988. The Psychology of Learning and Motivation, vol. 2. Academic Press, San Diego, CA, pp. 193–225.
Horton, J.J., Chilton, L.B., 2010. The labor economics of paid crowdsourcing. In: Proceedings of the 11th ACM Conference on Electronic Commerce. ACM, New York, NY, USA, pp. 209–218.
Huang, K., Chen, C., Chiang, S., 2008. Icon flickering, flicker rate, and color combinations of an icon's symbol/background in visual search performance. Percept. Mot. Skills 106, 117–127.
International Standards Organization (ISO), 2003. Graphical Symbols — Safety Colours and Safety Signs — Part 1: Design Principles for Safety Signs in Workplaces and Public Areas. ISO 3864-1. Geneva, Switzerland.
International Standards Organization (ISO), 2007. Graphical Symbols – Test Methods for Judged Comprehensibility and for Comprehension. ISO 9186. Geneva, Switzerland.
Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. 20, 1254–1259.
Laughery, K.R., Young, S.L., Vaubel, K.P., Brelsford, J.W., 1993. The noticeability of warnings on alcoholic beverage containers. J. Publ. Pol. Market. 12, 38–56.
Lesch, M.F., Horrey, W.J., Wogalter, M.S., Powell, W.R., 2011. Age-related differences in warning symbol comprehension and training effectiveness: effects of familiarity, complexity, and comprehensibility. Ergonomics 54, 879–890.
Madden, D.J., Pierce, T.W., Allen, P.A., 1996. Adult age differences in the use of distractor homogeneity during visual search. Psychol. Aging 11, 454–474.
Martin, D.W., McDonald, D.R., Patton, C.R., 1987. The benefit/cost of VDT highlighting. In: Salvendy, G., Sauter, S.L., Hurrell Jr., J.J. (Eds.), Social, Ergonomic and Stress Aspects of Work with Computers. Elsevier Science Publishers, Amsterdam, pp. 89–96.
McDougald, B.R., Wogalter, M.S., 2011. Increased comprehension of warning pictorials with color highlighting. In: Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting. Human Factors Society, Santa Monica, CA, pp. 1769–1772.
Paolacci, G., Chandler, J., Ipeirotis, P.G., 2010. Running experiments on amazon mechanical turk. Judgm. Decis. Mak. 5, 411–419.
Peters, R.J., Iyer, A., Itti, L., Koch, C., 2005. Components of bottom-up gaze allocation in natural images. Vis. Res. 45, 2397–2416.
Plude, D.J., Hoyer, W.J., 1986. Age and the selectivity of visual information processing. Psychol. Aging 1, 4–10.
Rogers, W.A., Fisk, A.D., 2001. Understanding the role of attention in cognitive aging research. In: Birren, J.E., Schaie, K.W. (Eds.), Handbook of the Psychology of Aging, fifth ed. Academic Press, San Diego, CA, pp. 267–287.
Salthouse, T.A., 1996. The processing-speed theory of adult age differences in cognition. Psychol. Rev. 103, 403–428.
Tamborelloii, F., Byrne, M., 2007. Adaptive but non-optimal visual search behavior with highlighted displays. Cogn. Syst. Res. 8, 182–191.
Watanabe, R.K., Gilbreath, M.K., Sakamoto, C.C., 1994. The ability of the geriatric population to read labels on over-the-counter medication containers. J. Am. Optom. Assoc. 65, 32–37.
Westinghouse Electric Corporation, 1981. Product Safety Label Handbook. Westinghouse Printing Division, Trafford, Pa.
Wickens, C., Alexander, A., Ambinder, M., Martens, M., 2004. The role of highlighting in visual search through maps. Spat. Vis. 17, 373–388.
Wogalter, M.S., 2006. Purposes and scope of warnings. In: Wogalter, M.S. (Ed.), Handbook of Warnings. Lawrence Erlbaum Associates Publishers, Mahwah, NJ US, pp. 3–9.
Wogalter, M.S., Conzola, V.C., Smith-Jackson, T.L., 2002. Research-based guidelines for warning design and evaluation. Appl. Ergon. 33, 219–230.
Wogalter, M.S., Silver, N., Leonard, S., Zaikina, H., 2006. Warning symbols. In: Wogalter, M.S. (Ed.), Handbook of Warnings. Lawrence Erlbaum Associates Publishers, Mahwah, NJ US, pp. 159–176.
Wu, J.-H., Yuan, Y., 2003. Improving searching and reading performance: the effect of highlighting and text color coding. Inf. Manage. 40, 617–637.
Yantis, S., 2000. Goal-directed and stimulus-driven determinants of attentional control. In: Monsell, S., Driver, J. (Eds.), Control of Cognitive Processes: Attention and Performance XVIII. The MIT Press, Cambridge, MA, pp. 73–103.
Zwaga, H., Boersema, T., 1983. Evaluation of a set of graphic symbols. Appl. Ergon. 14, 43–54.