Applied Ergonomics 1994, **25** (1) 53–57

The effect of alternative product-label design on warning compliance

Michael S. Wogalter* and Stephen L. Young†

*Ergonomics Program, North Carolina State University, Department of Psychology, Poe Hall, Raleigh, NC 27695–7801, USA
†Human Factors Program, Rice University, Department of Psychology, Sewall Hall, Houston, TX 77251, USA

Many potentially hazardous products are packaged in small containers. Because of the limited amount of space available on these containers for warnings and other information, manufacturers often reduce the size and amount of printed material on the labels. This frequently impairs the message's legibility, noticeability and comprehensibility. Recently, several alternative label designs have been investigated using preference ratings, but whether the designs facilitate safer behaviour has not been determined. In the present experiment, two alternative designs (tag and wings) were compared with a conventional (control) design for their effect on behavioural compliance with a warning on a very small container of glue. Participants performed a parts-assembly task using the glue without being informed of the study's real purpose. Whether participants wore protective gloves as directed by the warning was measured. Results showed that the tag design produced significantly greater compliance than the other two designs. Measures of noticing, reading and recall of the warning mirrored the compliance results. While participants generally preferred the control label, they most preferred the tag warning. Overall, the results suggest that alternative designs like the tag can enhance warning communication and compliance in cases where surface area is limited.

Keywords: Warning, labels, consumer products, instructions, compliance

Introduction

Consumers are often unaware of the hazards associated with many of the products that they use. One way that manufacturers try to communicate information about the hazards to users is through on-product instructions and warnings. However, limited label space on products packaged in small containers often forces manufacturers to decrease the amount and/or the size of the information presented. This reduces its noticeability, legibility and, sometimes, comprehensibility. In an attempt to remedy the problems associated with limited space, manufacturers sometimes provide additional information on the external packaging, and in product inserts and owners' manuals. However, consumers may discard or misplace these materials after initial product use, thereby making this information less accessible when the product is used at a later time or by other persons.

Recently, Barlow and Wogalter (1991) began to address the problem of limited surface area by examining consumers' preferences for six alternative product-label designs for a glue product contained in a very small bottle (8.9 ml). Each of the designs provided a different method of increasing the surface area of the label relative to a conventional (control) label design. Two of the alternative designs (tag and wings) plus the control are shown in Figure 1. Elderly participants (mean = 76 years) and college students rated the label designs on attractiveness, ease of use, willingness to purchase, ease of reading the label in general, and likelihood of noticing and reading the warning. The results showed that the wings design was preferred by the elderly participants on most of the factors tested. However, the college students preferred the tag design with regard to noticing and reading the warning, the wings design for ease of reading the label in general, and the control design for attractiveness, ease of use, and purchase intentions. More recent follow-up research (Wogalter *et al*, 1993) demonstrates that differences between these two subject populations, with regard to the tag and wings designs on the warning-related dimensions, are attributable to the larger print

 0003-6870/94/01 0053–05 © 1994 Butterworth-Heinemann Ltd

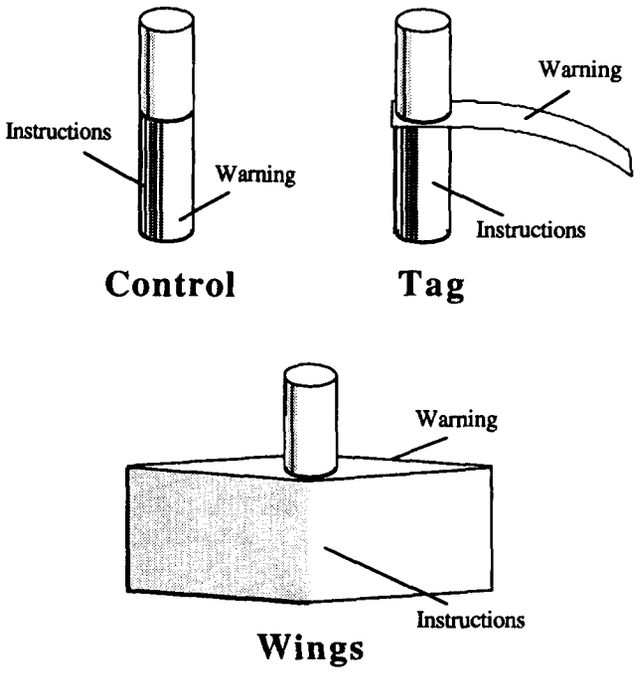

Figure 1 The control and two alternative label designs

afforded by these labels compared with the control design.

The two rating studies suggest that alternative label designs may enhance the communication of warnings and instructions. However, people's judgements about the label designs may not reflect any increase in the amount of information actually transmitted by the warnings or produce any increase in behavioural compliance. In other words, people's preference for a particular label design may not be related to its ability to produce safe behaviours on the part of the user.

The purpose of the present study is to examine the effectiveness of a warning (in terms of behavioural compliance) on two alternative labels (tag and wings) relative to a control design. Participants interacted with one of three container designs for a glue product while performing a parts-assembly task. The instructions and task were set up so that the true focus of the study was never explicitly stated or readily apparent (that is, under incidental exposure conditions). This manipulation helped to ensure that warning compliance behaviour could be assessed under fairly realistic product-use conditions.

Method

Participants

Forty-four undergraduate students from Rice University participated for credit in an introductory psychology course. They were randomly distributed across the three label conditions such that each group contained approximately the same number of participants. The control and tag conditions each had 15 participants, and the wings condition had 14 participants.

Materials

Product information and warnings were presented on realistic-appearing, but fictional, glue containers. The product was held in identical 8.9 ml (0.3 fl oz) glass-cylinder bottles with brush-applicator caps. The bottle had a circumference of 5.0 cm and a total height of 6.4 cm (3.7 cm and 2.7 cm for the glass and cap, respectively). The three label designs were control, tag and wings (see Figure 1). The printed label on the control bottle occupied all the available space of the bottle's glass section. The tag label was 2.0 cm wide and extended 6.2 cm beyond the edge of the cap. The wings label was 3.7 cm tall and extended 2.9 cm on either side of the bottle (5.8 cm width on each label side).

For the control design, the non-warning material was printed in 5 point Times Roman and the warning was printed in 5.5 point Times Roman. The two alternative designs had identical font and size characteristics for all the information except the warnings. Because greater surface area was available, the print size of both alternative-label warnings was enlarged to 9 point Times Roman Bold.

The added surface area of the wings was fabricated using foam-core board. The tags were made with stiff paper labels. Labels were laser-printed and all surfaces were covered with clear plastic laminate. A representation of the information included on the labels is shown in Figure 2.

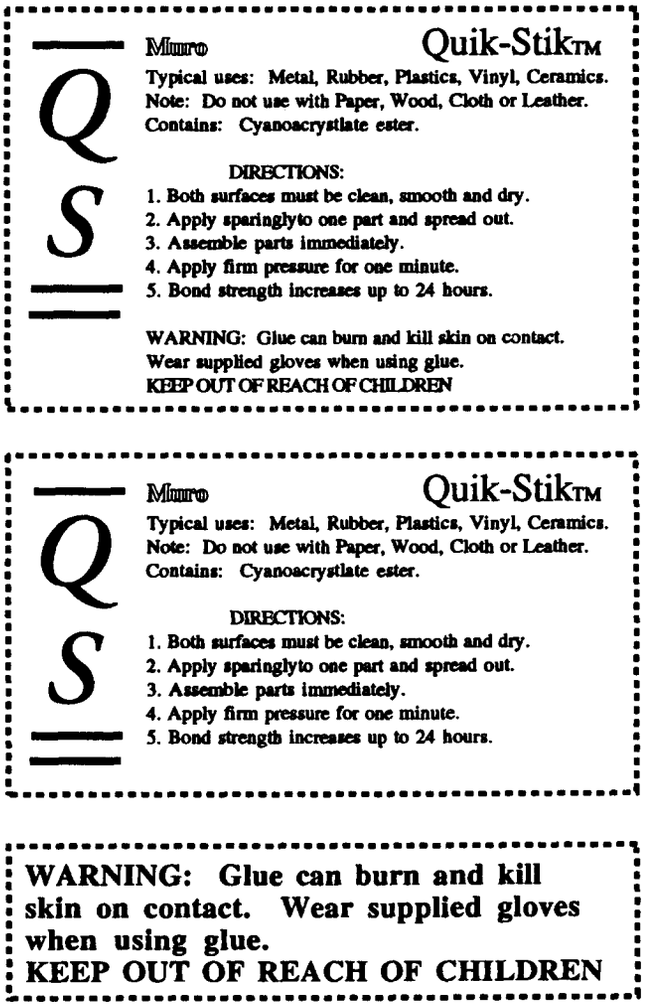

Figure 2 The printed labels: top, the control label; bottom, the tag and wings label (with the separate warning section)

Rubber cement served as the 'glue' enclosed in the bottles. Rubber cement was chosen because: (1) it is an adhesive; (2) it does not actually harm skin, thereby reducing the ethical difficulties associated with allowing volunteers to use a truly hazardous substance; (3) it has a slight (but non-toxic) odour; and (4) it is 'stringy' and could be seen as likely to come into contact with skin during normal use (especially when applied with a brush).

Procedure

A Testors Cessna 0–2A/B Skymaster model airplane was used. All the parts for the model, as well as the box, some sanding paper, a model knife, the glue and a pair of latex gloves, were laid out on a table in front of the participants. The impression given to participants was that all of these materials were supplied with the model. Participants were given the following instructions:

> In this study you are going to put together part of a model airplane. Models like the one in front of you are always accompanied by instructions which have multiple stages. Within each stage, there are usually several parts to be assembled, but the instructions sometimes do not state the order in which each part should be added. We are interested in how people determine what parts are assembled within a stage and in what order these parts are applied. Thus, we want you to examine the first stage of this model and assemble a portion of it, but not all of it. We want you to put together as many parts of the first stage as you feel are necessary to have completed a sufficient part of that stage. It is up to you to determine what is sufficient. Remember, there are no right or wrong solutions. We are merely interested in the way people decide what to do. In order not to influence your decisions, I will be at the other side of room and will not be able to answer any questions once you start.

The experiment was conducted in a 5 × 11 m room with several one-way mirrors. After presenting the instructions to participants, the experimenter sat at a table facing away from the participant at the other end of the room. The appearance of the experimenter's being disinterested in the task (for example, by moving to the opposite end of the room) was intended to enhance the belief that the experimenter was only concerned with the final results of their work. Participants' actions were observed by a second experimenter through one of the mirrors. Although each mirror was covered with an opaque shade, the observing experimenter could see through the side of one shade located farthest from the participants. Participants' behaviour of wearing or not wearing the gloves was recorded.

After completing the model-construction task, participants were asked several questions concerning the glue bottle:

1 whether they noticed the instructions;
2 whether they read the instructions;
3 to recall (or guess) what the instructions said;
4 whether they noticed the warning;
5 whether they read the warning;
6 to recall (or guess) what the warning said;
7 if they had any experience constructing models; and
8 approximately how many models they had previously constructed.

All of the items were scored as dichotomous variables (0 = 'no' and 1 = 'yes') except for the last item, which could take on any non-negative integer. Recall of the instructions and the warning were scored using a lenient criterion. The responses were counted as correct if the participant recalled the core or basic concepts (for example, the need to wear gloves because of potential injury), rather than the exact wording of the material presented.

Lastly, participants were asked their preference for the three label designs. They were shown all three designs and asked to rank the containers according to (a) their general preference for the labels (excluding the warning information) and (b) their preference for the labels with regard to how well each displayed the warning information. Participants were allowed to assign two or more of the containers the same rank, if they believed they were equivalent on one or both factors. After these questions, participants were fully debriefed as to the nature and purpose of the study.

Results

Compliance frequencies for the three bottle-label designs are shown in the top row of Table 1. The overall chi-square analysis was significant; χ^2 (2, N = 44) = 14.05, $p < 0.001$. Paired comparisons using Fisher's exact probability tests with one degree of freedom showed that participants complied significantly more often with the tag warning than with either the wings ($p < 0.03$) or the control warning ($p < 0.001$). The wings and control designs did not differ ($p > 0.05$).

Table 1 Dependent measures as a function of label design

	Label design		
Dependent measure	**Control (n=15)**	**Wings (n=14)**	**Tag (n=15)**
Behavioural compliance	2	5	12
Label instructions			
Notice	6	4	4
Read	4	4	3
Recall	4	4	2
Warning message			
Notice	4	7	15
Read	2	5	13
Recall	2	5	13
Preference rankings[1]			
General	1.4	2.5	2.1
Warning	2.9	1.9	1.2

[1] Lower scores on the two rank order measures indicate greater preference. The ranks ranged from 1 (best) to 3 (worst)

Table 1 also shows the results from the post-task questions. Chi-square analyses failed to show any reliable differences for noticing, reading and recalling the (non-warning) instructions ($p > 0.05$). However, the analyses of those items involving the warning showed significant effects: for noticing the information in the warning, χ^2 (2, $N = 44$) = 17.39, $p < 0.001$; for reading it, χ^2 (2, $N = 44$) = 17.05, $p < 0.001$; and for recalling it, χ^2 (2, $N = 44$) = 17.05, $p < 0.01$. Paired comparisons showed that more participants in the tag condition noticed, read and recalled the warning than participants in the wings and control conditions. The wings and control conditions did not differ for any of these measures ($p > 0.05$).

As the pattern in the table suggests, most of the participants who reported noticing the warning donned the protective gloves (phi coefficient, $\Phi = 0.73$, $p < 0.0001$), and virtually all of the participants who read and recalled the warning wore the gloves (both Φs = 0.96, $p < 0.0001$). Substantial positive relationships were also seen among the warning-related questions: noticing and reading ($\Phi = 0.76, p < 0.001$), noticing and recall ($\Phi = 0.76, p < 0.001$), and reading and recall of the warning ($\Phi = 1.0, p < 0.001$). No significant relationship was found for participants' reported previous experience constructing models.

Mean ranks for general-label and warning preference are shown at the bottom of Table 1. The Friedman within-subjects multicondition test for ranks showed a significant effect of container design for general-label, χ^2 (2, $N = 43$) = 27.35, $p < 0.0001$, and warning preference, χ^2 (2, $N = 44$) = 66.86, $p < 0.0001$. Paired comparisons using Wilcoxon's signed-rank test (corrected for ties) showed that all differences were significant ($p < 0.05$). Participants preferred the control for general label design, followed by the tag, and last by the wing. However, with regard to the display of the warning, participants preferred the tag, followed by the wing, and last by the control. These preferences were independent of the label condition in which the participants took part in the main experiment. Specifically, chi-square tests demonstrated that participants' preference rankings for the different bottle designs were not biased by exposure to any one design during the model-building task ($p > 0.05$).

Discussion

The results demonstrate that participants complied more often with the tag warning than with the wings or the control warnings. The facilitated compliance by the tag label can be attributed largely to the tag's presenting the warning in a more noticeable way than the other two label designs. The tag warning was facing upwards, which allowed the participants to see the warning at a more direct angle. In the other two designs, it was necessary to pick the bottle up from the resting position on the table to see the warning straight on. Also, the warning placement of the tag method was more salient because the message was continuously visible as the participants interacted with the product during the model-construction task. In the other two designs, the warning could be covered by the participant's hand while holding the product container.

Another indication that warning noticeability was an important factor for compliance was revealed by the post-task questions. The reports of noticing and reading the warning, and its correct recall, virtually mirrored the compliance results. In other words, if participants noticed the warning, they were very likely to read, recall and comply with it. Thus a major facilitating factor for compliance in this case was the tag warning's ability to attract or even demand attention, which the wings and control designs did less well. Indeed, 100% of the participants in the tag condition noticed the warning, and 80% complied.

This pattern of high conditional probabilities between noticing, reading, recalling and compliance differs from other research showing relatively large decreases from initial (noticing) to later (compliance) stages of processing (see DeJoy, 1989, for a review). However, the pattern exhibited in the current experiment supports other findings that show a strong positive association between noticing or reading a warning and compliance (Wogalter *et al*, 1985; 1987). In any event, the noticeability factor is probably not the sole reason for the relatively high compliance rate. Because the gloves were readily available and required little effort to don, the costs (in terms of time and effort) associated with compliance were low (Wogalter *et al*, 1989). If the protective equipment had been more difficult to obtain and use, then it is likely that there would have been a reduction in compliance, in spite of the warning being noticed and read. Thus, while noticing the warning is necessary, this factor alone is not sufficient for compliance. Other factors [such as social influence (Wogalter *et al*, 1989), costs of compliance and non-compliance] can mediate the processes leading from noticing the warning to behavioural compliance.

These findings also confirm earlier subjective preference research showing that college student populations favour the way the warning is displayed on the tag compared with the other designs (Barlow and Wogalter, 1991; Wogalter *et al*, 1993). One discrepancy from the earlier work is that students preferred the wings design for presenting the (non-warning) instructions in the previous research, whereas students preferred the tag design in the present study. However, the disparity is not large, because the earlier work showed that the tag design followed closely behind the wings design on the general label measures. Further confirmation of the earlier research was shown by the preference ranks for warning information. While the college students generally preferred the control design over all others, they preferred the tag design for presenting the warning information.

The reasons for students' general preference for the control label are not particularly clear, but may be due to familiarity and perceived price. The control label is a conventional design, which the students are more accustomed to, constituting a basis for liking this design. In addition, many students are on limited budgets and they might have had the belief that the other designs would be more expensive, producing a dislike for them. Furthermore, it is likely that the students could read the very small print on the control label, and so the advantage of having greater surface

area for bigger print in the alternative designs might not have been a highly important feature for this population. Some support for the expense explanation is provided by earlier work (Barlow and Wogalter, 1991) showing that the students were more willing to purchase the control than the other designs. However, further investigation is necessary to provide more substantive evidence on why they had this preference.

Other comments on the study's limitations are worth noting. First, only one segment of the general population was tested. In earlier research, the wings label was favourably judged by elderly participants. While the present study did not test the elderly, the rather strong confirmation of the students' preferences by the current behavioural study suggests that subsequent research involving behavioural measurement of the elderly might support the earlier preference research as well. Another basis for predicting the wings to be better than the tag for the elderly population is suggested in comments made by several elderly participants in earlier research. They commented that the wings design provides a better grip, which would make the container easier to hold and to remove the top compared with the other designs. A behavioural study evaluating this aspect of the containers is needed to test its validity. However, subsequent research might need to employ a different task (other than the model-construction activity used in this research), because the elderly participants may find the task to be less relevant and engaging than the students did.

Second, only one kind of product and task was tested. Different factors might be involved in compliance with warnings on other products and its associated activities: for example, the proper consumption of medicine as directed by pharmaceutical-container labels. Again, the relevance of the product and task to a given population may be an important consideration.

Third, the noticing and reading measures are self-report measures. These indices are usually considered to be more subjective (less objective) than other kinds of measures that could have been collected, such as the recording of eye movements. With self-reports, it is sometimes not clear whether participants have responded in a bona fide manner. However, given that noticing and reading reports correspond to the recall and compliance scores, it is probably the case that most participants answered truthfully on the self-report items.

Fourth, the label designs tested in this and earlier research are not the only ones that could have been developed and studied. The designs were prototypes (mock-ups); other designs could be developed and tested. For example, the tag and the wings labels could be combined to form several kinds of hybrid design. Future designs should consider not only the effectiveness in transmitting warning information, but also the attractiveness of the container in general. Particularly good designs would be ones that present and convey warning information effectively, while at the same time motivating the consumer to purchase the product (or at least not to avoid the product because of its appearance).

In summary, this research demonstrates that one of the alternative designs, the tag label, is a viable and effective method of enhancing the communication of on-product warnings. In all of the warning-related measures (compliance behaviour and container-design preferences) the tag was favoured over the other two designs. While the wings label was not significantly different from the control, almost all the measures showed a trend towards its being a better design than the conventional label. By increasing the available surface area and by occupying a good location on the bottle, alternative designs can increase the quantity and quality of space provided for important hazard information. The choice of label design will not only be determined by the importance of conveying the warning information, but also by consideration of other factors, such as the demographics of the expected consumer population, the needs of the population, the product itself, and the anticipated tasks.

Alternative designs such as the tag provide a solution for the problem of limited space on consumer product labels. They can reduce the need for product manufacturers to (a) exclude potentially important information because of space considerations, (b) make the print so small that it is difficult to read, or (c) move warning information to less accessible locations such as the packaging materials, inserts and manuals. It is expected that the provision of warning information in a more salient and more accessible manner will increase the likelihood that consumers will use the product appropriately and take appropriate precautionary measures. The tag design is a very inexpensive and practical method to increase the likelihood that the warning will be noticed, read and complied with.

Acknowledgements

Portions of this research were reported in a poster session at the Human Factors Society 36th Annual Meeting, October 1992, Atlanta, Georgia.

References

Barlow, T and Wogalter, M S 1991 'Increasing the surface area on small product containers to facilitate communication of label information and warnings' in *Proc Interface 91* Human Factors Society, pp 88–93

DeJoy, D M 1989 'Consumer product warnings: review and analysis of effectiveness research' in *Proc Human Factors Soc 33rd Ann Mtg*, pp 936–940

Wogalter, M S, Fontenelle, G A and Laughery, K R 1985 'Behavioral effectiveness of warnings' in *Proc Human Factors Soc 29th Ann Mtg*, pp 679–683

Wogalter, M S, Godfrey, S S, Fontenelle, G A, Desaulniers, D R, Rothstein, P and Laughery, K R 1987 'Effectiveness of warnings' *Hum Factors* **29** (5), 599–612

Wogalter, M S, Allison, S T and McKenna, N A 1989 'The effects of cost and social influence on warning compliance' *Hum Factors* **31** (2), 133–140

Wogalter, M S, Forbes, R M, Van't Slot, L J and Barlow, T 1993 'Facilitating communication of label information and warnings by increasing the surface area and print size on small product containers' in *Proc Interface 93* Human Factors Society, pp 181–186