PRODUCT LABEL LIST FORMAT: EFFECTS OF ITEM ARRANGEMENT AND COMPLETENESS ON COMPARISON TIME AND ACCURACY

Michael S. Wogalter
Department of Psychology
North Carolina State University
Raleigh, NC 27695-7801

Michael J. Kalsher
Department of Psychology
Rensselaer Polytechnic Institute
Troy, NY 12180-3590

ABSTRACT

This research examined the influence of two factors on the ease of gaining information from item lists: the order and completeness of the items displayed. Food nutrition labels served as the vehicle to test the manipulated lists on comparison time and accuracy performance measures. Four booklets, each containing 12 pairs of nutrient labels, were constructed in which the listed items were: (1) either arranged in a standard order or in a random order, and (2) either had a complete set of nutrients (including nutrients not present in the product) or a partial set of nutrients (excluding nutrients with zero or near-zero amounts in the product). Thirty-two participants were instructed to assume that their physician has told them to increase their intake of three specific nutrients and decrease their intake of three other nutrients. One label of each pair contained a higher level of one nutrient that should be increased or a lower level of one nutrient that should be decreased. Participants were to determine which of the two labels would be better given the prescribed diet. Time and accuracy measures were collected. Participants made significantly faster judgments for nutrients arranged in a standard order than for nutrients arranged in a random order. For all conditions, the error rate was low. An interaction indicated that labels with a complete set of nutrients in a standard order produced fewer errors than (a) labels in a standard order with some nutrients missing or (b) labels with a complete set of nutrients in a random order. Implications of making lists compatible with expectations are described.

INTRODUCTION

The importance of being able to extract information from item lists is apparent in many contexts. Lists are found on food nutrition labels, in ingredient lists for edible and nonedible substances, and in various kinds of paper-based and video displays. Lists can be useful in providing information on a myriad of content domains, but finding and gaining the desired information can sometimes be time consuming and error prone (e.g., Tullis, 1990).

The ability to extract information from lists is important for a number of reasons. For example, attaining the correct information from lists of nutrients or ingredients on food products is important for individuals who have a health conditions affected by dietary content and who may need to consume more of some substances and avoid others. In this example, easy and consistent access to items listed on food products may be critically important. However, most of the research on characteristics of list structures have been based on *subjective preference* surveys (e.g., Heimbach, 1982; Heimbach and Stokes, 1982; Levy, Mathews, Stephenson, Tenney, and Schucker, 1985). Relatively few studies have examined people's *performance* in using various kinds of manipulated lists (e.g., Levy, Fein, and Schucker, 1992; Tullis, 1990) and most of the current research in this domain has concerned the design of computer menu systems (see Mayhew, 1992, for a review).

The purpose of the present study is to examine the influence of two factors, item order and completeness, on the performance (time and accuracy) to compare the item contents of pairs of fictitious nutrition labels. Food nutrition labels were selected as the vehicle to examine these factors, because of the considerable controversy that has surrounded the format of nutrition labels over the last few years. Despite recent implementation (U.S. Department of Health and Human Services, 1989, 1990; Sullivan, 1990) of stricter labeling requirements (National Label and Education Act of 1990), consumers are still confused with respect to the presence and quantity of nutrients because of the inconsistent ways nutrition information can be presented on the labels. Survey research conducted by the Opinion Research Corporation (1990) for the National Food Processors Association suggests that people prefer a standard ordering of items on the labels and that they prefer seeing all items listed even when the product contains zero or near-zero amounts of some items. However, at this point it is not clear whether these preference data actually reflect objective performance measures associated with the ease of reading the labels. Levy et al. (1992) present data demonstrating that preference and performance measures do not necessarily concur.

The factors of interest in this study, item order and completeness, are not only relevant to nutrition labels but also other kinds of list structures, including computer menu systems. Francik and Kane (1987) found that deletion of irrelevant items from menus produced faster selection times for expert computer operators using a keyboard input device compared to a condition in which the irrelevant items were displayed in a grayed out font (indicating that the items were temporarily inactive). However, Mayhew (1992) speculated that displaying grayed out items might be useful for less proficient operators using a mouse (or other input pointing device) because it would help these users develop a better mental model of the program by seeing what is potentially available. In other computer menu research, Somberg (1987) showed that experienced users were faster in making selections from a menu list in which the items were maintained in the same position compared to random ordering or two rule-based orderings.

The present study addressed two issues that are similar to the computer menu research mentioned above, but examined them with respect to evaluating nutrition labels: (1) constancy of the arranged order of items, and (2) completeness of the set of nutrients named on the label (including or excluding from the list certain nutrient items with zero or near zero quantities in a particular product). Performance was evaluated using speed and accuracy measures to compare two labels that differed in the quantity of particular nutrients. It was expected that: (1) people would find it easier to compare the labels in which all items were in a standard order as compared to a random order, and (2) labels having all items included would produce better performance compared to labels having certain items omitted.

METHOD

Participants

Thirty-four Rensselaer Polytechnic Institute undergraduates enrolled in introductory psychology courses received credit for their participation.

Stimulus Materials

The food labels used were based on those found on tomato soup cans. The numerical quantities of nutrients listed on the experimental labels were modified slightly from the original label so that label pairs in the booklets were not identical. Label pairs presented to participants were manipulated according to: (1) arrangement of individual nutrients on the label (standard vs. random order) and (2) completeness of the set of nutrients listed (all nutrients present vs. certain nutrients with zero, or near zero, amounts omitted). This resulted in four label configurations: (1) *Standard-All*, in which all nutritional items were listed on the label according to a standard format—similar to the nutrition list on tomato soup products; (2) *Standard-Missing*, in which items containing zero amounts (or RDAs designated with an asterisk indicating less than 2%) were omitted from the label but it was otherwise identical to the Standard-All format; (3) *Random-All*, in which all nutritional items were present, but arranged in a random order; and (4) *Random-Missing*, in which items containing zero or near zero amounts were omitted and the remaining items were listed in a random order. In the conditions having labels with missing items, two or three nutrients were deleted from both labels of a pair. Figure 1 shows four label pairs representing each condition.

For each label pair, the amount of one of six *relevant* items (defined in the present study as modifications to calcium, cholesterol, fat, iron, sodium, and vitamin C) and the amount of one or two *irrelevant* items (changes to nutrients such as protein, vitamin A, niacin, thiamine) was modified for one label in each pair. The amount of change for the relevant and irrelevant items was identical (two units), but opposite in direction. Thus, if the amount of a relevant item listed on a label were decreased, one or two irrelevant items on the label were increased to offset this change so that the numbers totaled to the same amount on each label.

The purpose of manipulating an irrelevant item in the labels was to avoid the possibility that participants could scan the quantities without giving specific attention to the particular nutrients. This also made the task more realistic because few real labels differ by only a single quantity, and made the task more difficult for the purpose of increasing the sensitivity (power) of the experiment to find possible format effects. Otherwise the labels of each pair were identical.

Booklets containing 12 label pairs (one label pair per page) were constructed for each of the four conditions, for a total of 4 booklets and 48 label pairs. Label pairs were positioned vertically on the page. The top one was labeled "Product A" and the bottom one was labeled "Product B." In every booklet, each of the six relevant items differed twice. Product A was the correct choice for six comparisons, and for the remaining six comparisons Product B was the correct choice. (Participants were not informed of this fact.) Label pairs were randomly ordered within each booklet. As a further control for order effects, the booklets were arranged such that half the participants examined the label pairs in the forward order and half examined them in the reverse order.

Procedure

After signing a consent form, participants were given the following description, and were told to assume its truth throughout the study:

> "Your doctor has advised you that you are at risk for a serious disease. However, you may be able to control this condition through proper diet. The doctor recommends that you should *reduce* your intake of Fat, Sodium, and Cholesterol and that you should *increase* your intake of Calcium, Iron, and Vitamin C."

Figure 1

Representative labels for the four conditions. Note that for each of these pairs Product B is the correct answer because one of the relevant items, Cholesterol, is lower, which according to the nutrition instructions given to subjects should be decreased.

STANDARD ALL

Product A
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Calories	90
Protein	2g
Carbohydrate	17g
Fat	2g
Cholesterol	2g
Sodium	680mg

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Protein	2
Vitamin A	10
Vitamin C	45
Thiamine	2
Riboflavin	*
Niacin	4
Calcium	*
Iron	2

* Contains less than 2% of US RDA of these nutrients

Product B
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Calories	90
Protein	0g
Carbohydrate	17g
Fat	2g
Cholesterol	4g
Sodium	680mg

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Protein	2
Vitamin A	10
Vitamin C	45
Thiamine	2
Riboflavin	*
Niacin	4
Calcium	*
Iron	2

* Contains less than 2% of US RDA of these nutrients

STANDARD MISSING

Product A
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Calories	90
Protein	1g
Carbohydrate	17g
Fat	2g
Cholesterol	3g
Sodium	680mg

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Protein	2
Vitamin A	10
Vitamin C	45
Thiamine	2
Niacin	4
Iron	2

* Contains less than 2% of US RDA of these nutrients

Product B
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Calories	90
Protein	2g
Carbohydrate	18g
Fat	2g
Cholesterol	1g
Sodium	680mg

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Protein	2
Vitamin A	10
Vitamin C	45
Thiamine	2
Niacin	4
Iron	2

* Contains less than 2% of US RDA of these nutrients

RANDOM ALL

Product A
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Protein	2g
Calories	90
Carbohydrate	17g
Fat	2g
Sodium	680mg
Cholesterol	2g

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Vitamin A	10
Vitamin C	45
Protein	2
Riboflavin	*
Thiamine	2
Niacin	4
Calcium	*
Iron	2

* Contains less than 2% of US RDA of these nutrients

Product B
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Calories	90
Carbohydrate	17g
Fat	2g
Protein	0g
Sodium	680mg
Cholesterol	4g

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Protein	2
Vitamin A	10
Thiamine	2
Riboflavin	*
Niacin	4
Calcium	*
Iron	2
Vitamin C	45

* Contains less than 2% of US RDA of these nutrients

RANDOM MISSING

Product A
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Calories	90
Fat	2g
Carbohydrate	17g
Cholesterol	3g
Sodium	680mg
Protein	1g

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Vitamin A	10
Niacin	4
Vitamin C	45
Thiamine	2
Protein	2
Iron	2

* Contains less than 2% of US RDA of these nutrients

Product B
TOMATO SOUP
NUTRITION INFORMATION
PER SERVING

Serving size	4oz condensed
Servings per container	3
Protein	2g
Carbohydrate	18g
Calories	90
Fat	2g
Sodium	680mg
Cholesterol	1g

PERCENTAGE OF US RECOMMENDED DAILY ALLOWANCES (US RDA)

Vitamin C	45
Thiamine	2
Iron	2
Niacin	4
Protein	2
Vitamin A	10

* Contains less than 2% of US RDA of these nutrients

To assist them in their task, participants were given the information depicted in Figure 2 on an index card. Participants were told to become familiar with the relevant items and to memorize the information on the card. They were given as much time as they needed to commit the items to memory. When participants indicated that they knew the information on the card, they were shown two practice label pairs to ensure that they understood the task.

Before the participants received the experimental materials, they were told: (a) to examine every label pair in the booklets to determine which label better fit the prescribed diet, (b) to mark the answer on the response sheet, and (c) to continue through each booklet until they completed the set of labels. The instructions emphasized both accuracy and speed. Participants evaluated all four booklets, one of each format. Participants were given a short rest period between booklets. Time to complete each booklet was measured.

Different random orders of booklets were given to participants to control for effects of learning and fatigue. After all four booklets had been completed, each participant was debriefed and thanked for his or her participation.

RESULTS

The time and error data were analyzed in separate 2 (Arrangement: standard versus random order) X 2 (Completeness: all versus missing items) repeated-measure analyses of variance (ANOVAs). Mean time and proportion error per label pair for each condition are shown in Table 1.

Using the completion time scores, the ANOVA showed a significant effect of Arrangement, $F(1, 33) = 56.25$, $p < .0001$. Participants viewing the standard arrangement labels ($M = 16.75$ s) performed significantly faster than participants viewing the random arrangement labels ($M = 22.82$ s). There was no significant main effect of Completeness, nor was the interaction significant (both Fs < 1.0).

Figure 2

Depiction of Information Presented on Index Card Showing the Desired Direction of Relevant Items to Assist Participants in Making Appropriate Label Choice Decisions.

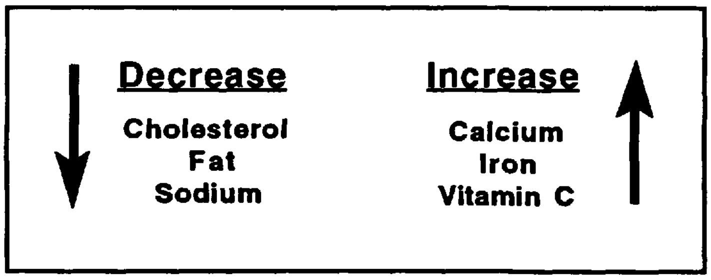

Table 1

Mean Time (in sec) and Proportion Errors Per Label Pair as a Function of Arrangement and List Completeness

	Arrangement			
	Standard		Random	
	Time	Error	Time	Error
Completeness				
All	16.13	.020	22.66	.053
Missing	17.38	.051	22.98	.034

Table 1 shows that the error proportions were low across all conditions. The ANOVA on these scores showed no significant main effects (both Fs < 1.0) but did show a significant interaction of Arrangement and Completeness, $F(1, 33) = 17.79$, $p < .001$. Post hoc comparisons using Newman-Keuls multiple-range test showed that the Standard-All condition produced significantly fewer errors than either the Standard-Missing or the Random-All conditions (ps < .01). No other comparisons were significant (ps > .05).

DISCUSSION

The results show that a standard arrangement of items on nutrition labels produces significantly faster comparison times than a nonstandard (random) order. This finding supports preference data from survey research (Opinion Research Corporation, 1990) suggesting that nutritional elements should be ordered uniformly across products.

Though the completeness factor did not have an effect on performance time, it did enter into an interaction with item arrangement in the error analysis. The error results showed that the Standard-All nutrition labels produced fewer errors than the Standard-Missing and the Random-All labels. One possible reason for the better performance in the Standard-All condition is that it allowed participants to locate all of the relevant items quickly and easily because the relevant nutrients were always in the same place and context.

The lower performance for the Standard-Missing and the Random-All conditions can be explained by the fact that in both conditions the context of the list items changed from label to label which could have disrupted comparison processing. Surprising was the fact that no significant difference was found between the Standard-All and Random-Missing conditions, although the latter condition was not significantly different from the two least accurate conditions either (Standard-Missing and Random-All). The somewhat higher than expected accuracy in the Random-Missing condition might be due to a shift in visual scanning strategy while comparing these labels. Alternatively, it might be due to chance (sampling error). Additional investigation is necessary to determine the reliability of this trend in the data.

The results indicate that participants were biased more to accuracy than speed, although both were emphasized in the instructions. The high accuracy rate suggests that participants took the task seriously. Their performance is likely to reflect that of careful consumers comparing labels. However, it was rather surprising to find that participants spent an average of 19.8 s per label pair (and 16.1 s in the best condition, Standard-All). This suggests that these nutrient labels require extensive, controlled processing even when labels differ by only a few elements and the participants know what to look for. It is difficult to imagine consumers spending this much time making comparisons among several similar products at the grocery store on a regular basis. It is not clear whether the newer nutrition labels recently introduced into the U.S. market would fare any better.

Because only two or three items were deleted in the missing-items conditions, the test of completeness might not have been as powerful as it might have been with more items deleted. With more missing items, the contextual nature of the label substantially changes which, in turn, might degrade performance. However, fewer items to search through might enable faster search times. Also, different effects might be expected depending on whether the items missing from the label are relevant or irrelevant to the person. The absence of relevant items might increase response times because an exhaustive search is necessary, whereas the absence of irrelevant items might produce no effect or facilitate response times to the relevant items. Further investigation is necessary to verify these possibilities.

Some product containers are very small, and thus, there are practical limits to the number of items that can be placed on the label and still maintain legibility. Sometimes having complete lists may not be as important as readability, particularly if the product is comprised of only a few component nutrients. One solution that has already been partly implemented in the 1975 FDA regulations allows the option of listing a short set of nutrients. Also offered in the 1991 U.S. DHHS document is a simplified format if the food product contains insignificant amounts of a majority of the required nutrients to be displayed. Another option is to increase the surface area of the product container to allow presentation of a larger nutrient list (and perhaps, larger print). Several ways of increasing the surface area of very small containers (e.g., attaching a tag) are described in Barlow and Wogalter (1991) and Wogalter and Young (1994).

Two final comments regarding generality of the results should be made. One concerns the population sample used. A representative sample of all potential consumers was not employed in this study. However, we would not expect other groups to show a much different pattern of results among conditions. The reason for this is that population sub-groups generally do not interact with experimental factors in reaction time studies. A second comment concerns the fact that we used just one kind of list, nutrition labels, in this study. However, the results concur with other studies in the computer menu literature that also involve lists of items (Francik and Kane, 1987; Mayhew, 1992; Somberg, 1987). Furthermore, we believe the present findings have generalizability to other kinds of lists in which comparison-type processing takes place.

The factors examined in this study, arrangement and completeness, are two of the most fundamental design features of item lists. Future investigations will assist in delineating other factors influencing list-related performance.

ACKNOWLEDGMENTS

The authors would like to thank Aida Hatem for her help in all aspects of this research, from constructing the stimuli to running subjects. The authors would also like to thank Diane Jessen for her helpful comments on an earlier version of this manuscript. Funding for this study was provided in part by a grant to the first author from the Paul Beer Trust Fund.

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 *Proceedings of Interface '91*. Santa Monica, CA: Human Factors Society.

Francik, E. P., and Kane, R. M. (1987). Optimizing visual search and cursor movement in pull-down menus. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 722-726). Santa Monica, CA: Human Factors Society.

Heimbach, J.T. (1982). *Public understanding of food label information*. Washington DC: U.S. Food and Drug Administration.

Heimbach, J.T. and Stokes, R.C. (1982). Nutrition labeling and public health: Survey of American Institute of Nutrition members, food industry, and consumers. *American Journal of Clinical Nutrition, 30*, 700-708.

Levy, A.S., Fein, S. B., and Schucker, R.E. (1992). More effective nutrition label formats are not necessarily preferred. *Journal of the American Dietetic Association, 92*, 1230-1234.

Levy, A.S., Mathews, O., Stephenson, M., Tenney, J.E., and Schucker, R.E. (1985). The impact of a nutrition information program on food purchases. *Journal of Public Policy and Marketing, 4*, 1-13.

Mayhew, D. J. (1992). *Principles and guidelines in software user interface design*. Englewood Cliffs, NJ: Prentice Hall.

Opinion Research Corporation. (1990). *Food labeling and nutrition: What Americans want - Summary of findings*. Washington DC: National Food Processors Association.

Somberg, B. L. (1987). A comparison of rule-based and positionally constant arrangement of computer menu items. In *CHI '87 Proceedings* (pp. 150-153), ACM.

Sullivan, L.W. (1990, July 12). *Press Conference on Food Labeling*. Washington, DC: Hubert H. Humphrey Building.

Tullis, T. S. (1990). The formatting of alphanumeric displays: A review and analysis. In M. Venturino (Ed.) *Selected Readings in Human Factors* (pp. 371-396). Santa Monica, CA: Human Factors Society.

U.S. Department of Health and Human Services. (1989). Food labeling: Advance notice of proposed rulemaking. *Federal Register* (DHHS Pub. 21 CFR Ch. 1). Washington DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. (1991). Food Labeling; General Provisions; Nutrition Labeling; Nutrient Content Claims; Health Claims; Ingredient Labeling; State and Local Requirements; and Exemptions; Proposed Rules. *Federal Register* (DHHS Publication 21 CFR Part 101, et al.) Washington DC: U.S. Government Printing Office.

Wogalter, M. S., and Young, S. L. (1994). The effect of alternative product-label design on warning compliance. *Applied Ergonomics, 25*, 53-57.