On the Prediction of Pictorial Comprehension

Kevin E. Hicks
Department of Psychology
North Carolina State University
Raleigh, NC 27695-7801 USA

Jennifer L. Bell
Department of Curriculum & Instruction
North Carolina State University
Raleigh, NC 27695-7801 USA

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

Abstract

Pictorial development and testing can be a costly and inefficient process. The process of designing and testing pictorial symbols could benefit from a precursor test to determine the likelihood that a concept will permit the design of a successful symbol (according to subsequent comprehension testing). This study examines whether ratings of the concepts of to-be-designed symbols could be useful in the prediction of comprehension of the ensuing symbols for those concepts. Participants rated 50 text descriptions (referent plus further verbal context) on: (1) how concrete is this concept? (2) how easy is it to visualize this concept?; and (3) how effective would a simple picture be in conveying this concept? These ratings served as predictors of the population estimates and open-ended comprehension scores obtained from previous research by Young and Wogalter (2001). Results showed that there was a high correlation between the measures used in this study with both population estimates and open-ended comprehension scores. Ease of visualization of a concept had the highest predictive value with concreteness, being the second highest predictor. Measures of the ease of visualization or imaging a concept may be useful preliminary tools for pictorial designers.

Introduction

Manufacturers, designers, and researchers have the arduous task of effectively communicating instructions, warnings, and hazardous information through pictorials to a variety of individuals from varied backgrounds and across many contexts and situations. Comprehension of a pictorial concept is critical to its ability to convey the appropriate information. Open-ended comprehension testing is a common method recommended by the American National Standards Institute (ANSI) to test pictorial understanding and is considered the "gold standard" for testing pictorial symbols (Dewar, 1998). In ANSI Z535.3-1998, open-ended testing using 50 participants, who are representative of the target population, are shown a series of pictorials and asked to respond with a short definition for each pictorial. These responses are then judged to be "correct" or "incorrect," with scores of 1 or 0 given respectively. Pictorials are judged acceptable if there is at least 85% correct and no more then 5% critical confusions (acting in direct opposition to the intended referent). These pictorials can be used in the intended setting with no additional textual Information needed. Pictorials receiving lower than 85% comprehension scores and are intended to be used alone (without text) must be re-designed and re-tested. The International Organization for Standardization (ISO 3864-1984) has somewhat different guidelines (i.e., 67% comprehension scores are judged acceptable). These organizations provide procedures and standards such as these to ensure safety pictorials are comprehended by a substantial percentage of the population.

However, iterative testing and re-design can be an expensive and time-consuming process (Wolff & Wogalter, 1993). Recent research has been conducted to discover alternatives to open-ended comprehension testing. Alternatives that could better predict a pictorials comprehension rating prior to open-ended testing. Zwaga (1989) used a method in which participants were shown a series of pictorials and asked to estimate the percentage of the population that would comprehend the pictorial. They completed the standard open-ended comprehension procedures as well, and the results of the two were compared. Estimates of population comprehension were significantly correlated to

open-ended comprehension scores. Brugger (1994) also used estimates of population comprehension scores compared to open-ended comprehension testing and found significant correlations; further demonstrating the utility of using preliminary estimations of population comprehension as predictors of open-end comprehension. Population estimates, while not necessarily recommended as a total replacement to open-ended comprehension testing (Wolff and Wogalter, 1998) are useful in narrowing the field of pictorials for further testing. Testing only the pictorials best suited for open-ended comprehension (those most likely to pass) could allow designers and researchers to save time and money.

McDougall, Curry, and de Bruijn (1999) studied some additional factors influencing pictorial comprehension and found that participants indicated a pictorial's concreteness and complexity were different component factors. McDougall and Curry 2000 performed additional research to assess the concreteness, complexity, and distinctiveness of safety pictorials and their relationship on pictorial comprehension. Participants were shown a series of verbal referents (i.e., no food or drink, wear hard hat, slippery surface). Each referent was presented with a set of nine pictorials in varying forms of concreteness and complexity. Participants were instructed to select the correct pictorial among those presented which best corresponded to the verbal referent provided. Seventy nine percent of the participants made the correct referent selection when the pictorials were concrete, as opposed to 48% when the pictorials were abstract. Results of the study concluded that a pictorial's concreteness, rather than its complexity, determined how accurately and easily it was understood.

The concreteness of a pictorial's intended message (concept) could have a significant impact on future open-ended comprehension testing results. It may be possible to eliminate or refine very abstract concepts being developed early in the design phase. If the notion of concreteness could be applied in the initial stages of pictorial construction, then designers and researchers could predict whether a concept would even be a viable candidate for further development.

Young and Wogalter (2001) looked at additional factors that may influence pictorial comprehension along with open-ended comprehension testing and ratings of population estimation. Participants were shown referents and pictorials and asked to rate the correspondence between the two. Ratings of the correspondence between a referent and the associated pictorial were significantly correlated to open ended comprehension scores. This suggests the conceptual nature of the verbal referent is important. It is possible that these concepts (verbal descriptions) themselves could be examined in a preliminary round of testing before a pictorial is even developed.

The present study used 50 concepts from the Young and Wogalter (2001) pictorial study and examined them in the form of ratings. The three rating measures were: (1) how concrete is the concept, (2) how easy is concept to visualize, and (3) how effective would a simple picture is in illustrating the concept. The data was used as a predictor of open-ended comprehension and population estimation scores obtained from Young and Wogalter (2001).

Method

Participants

Forty participants, all students from North Carolina State University in Raleigh, North Carolina, were recruited for this study and were compensated for their participation with research credits.

Materials

A series of 50 descriptive statements were developed referencing a set of fifty pictorial symbols used in Young and Wogalter (2001). The 50 pictorials were originally chosen to represent a variety of sources including prescription drug labels, stickers, industrial safety pictorials, safety related clip-art, and instructional manuals. Each descriptive statement was created to give a full verbal description of the original pictorials. The descriptive statements contained at least: (a) the pictorial referent (e.g., slippery surface), (b) context in which the pictorial would be used (e.g., located in an industrial setting) and (c) statement of hazard (e.g., wet floor) and/or instruction (e.g., use in case of fire). See Table 1 for samples of the pictorials, the referents used, and the corresponding descriptive statements. We used the referent as the

root of our descriptive statement followed by the context; further instructions and/or hazards were added to more fully communicate the conditions of the referent.

Each descriptive statement was individually printed on a 3 x 3 inch (7.62 cm x 7.62 cm) with a randomly assigned identifying number (1 to 50), the labels were then placed on 3 x 5 inch (7.62 cm x 12.7 cm) hole-punched index cards, and bound to create a booklet. Five booklets were created containing 50 index cards per booklet. The order of descriptive statements for each of the booklets was randomized per participant to protect against order effects.

Each participant was given a 3-ring binder that contained a page of instructions and a set of 3 rating sheets. The response sheet had numbered blanks for the participant to record the rating for each of the 50 descriptive statements. The 3 ratings for each of the 50 descriptive statements were: (1) How abstract versus concrete is this concept? (2) How easy is it visualize (or imagine) the concept in this statement, (3) How effective or useful would a relatively simple picture be in showing this concept? (Hereinafter called picture simplicity). Each rating sheet contained one question with one associated 0 to 100-point percentage scale using named anchors.

The concreteness rating scale used four named anchors ranging from extremely concrete (0), very concrete (30), very abstract (70), and extremely abstract (100).

The visualization rating scale used five named anchors ranging from not at all easy to visualize (0), not very easy to visualize (30), easy to visualize (50), very easy to visualize (70), and extremely easy to visualize (100).

The picture simplicity rating scale used five named anchors ranging from not at all effective/useful (0), not very effective/useful (30), effective/useful (50), very effective/useful (70), and extremely effective/useful (100).

Procedure

Participants were run in groups of 3 to 5. Participants were provided with a informed consent form prior to participation. They were then given a binder containing the instructions, rating sheets, and a booklet containing the descriptive statements. Participants were instructed to use one rating page at a time for each of the 50 descriptive statements to ensure that each scale was completed for all 50 items before moving on to the next rating scale. The rating sheets were randomly ordered for each participant. Participants were also instructed that they could give any rating, between 0 to 100, for any of the pictorials across the 3 different ratings scales. After each participant verified they understood the task, they were instructed to proceed with the rating tasks. After completion of all 150 ratings, answer sheets were collected, participants were asked if they had any questions and were debriefed.

Results

Besides the data collected in this study data a previous study were used. Overall scores for open-ended comprehension and population comprehension estimates for the 50 pictorials were obtained from research conducted by Young and Wogalter (2001). Simple and multiple regression correlations were performed for each averaged rating scale between open-ended comprehension scores and again with population comprehension estimation scores. Open-ended comprehension scores and population comprehension estimation scores and were also based on 0 to 100 point scales.

Concreteness, ease of visualization, and picture simplicity (shown in Table 2) were all significantly correlated at $p <$. 0001. The substantial positive intercorrelations among these three measures suggest that they may be measuring the same underlying dimension.

Open-ended comprehension. Visualization of the concept had the highest correlation ($r = .53, p < .0001$) with open-ended comprehension scores. The correlations with other two ratings with open-ended comprehension were somewhat lower but were still highly correlated: concreteness ($r = .48, p < .01$) and picture simplicity ($r = .40, p < .01$).

Multiple regressions were performed to determine whether a combination of the three ratings would better predict open-ended comprehension scores. The coefficients from the analysis never reached higher then $r = .53$, which was the same numerical value of the correlation of open-ended comprehension and visualization alone.

Table 1. *Example abstract and concrete pictorial symbols, referents, descriptive statements, and mean ratings of visualization, concreteness, picture simplicity, open-ended comprehension scores, and population comprehension estimation scores*

Pictorial	*Referent*	*Context and Descriptive statement*	*Visualization*	*Concreteness*	*Picture simplicity*	*Open-ended comprehension*	*Population comprehension estimation*
Abstract 	Radioactive	*Laboratory - on a door* Area is safe of radioactive release. Use for emergency protection from radiation.	32	23	38	10	39
Abstract 	Carcinogen	*Hospital - on a door in a x-ray department* Cancer causing agent. Prolonged exposure can cause severe illness or death.	43	45	42	0	16
Concrete 	Slippery surface	*Industrial setting - near a newly washed floor* Wet floor You could slip and fall.	81	73	68	84	78
Concrete 	Fire Exit	*Hotel - on a door in the hallway* Use only in case of fire.	82	77	58	91	79

Population Comprehension Estimation. All three ratings significantly were correlated to population comprehension estimation scores. Correlations between ratings of visualization and population comprehension estimation scores again had the greatest relationship ($r = .62, p < .0001$). Ratings of concreteness ($r = .58, p < .0001$) and ratings of simple picture of concept ($r = .46, p < .01$) were somewhat lower but were also highly correlated to the population comprehension estimation scores.

A multiple regression analysis was performed to determine if the three ratings (visualization, concreteness, and picture simplicity) could provide a stronger predictor to population comprehension estimation scores than just the ease of visualization alone. The results show that beyond the prediction provided by visualization, the other two measures did not add significantly more variance accounted for the population comprehension estimation scores.

Table 2. *Pearson correlation coefficients of visualization, concreteness, and picture simplicity, together with open-ended comprehension scores, and population comprehension estimation scores form Young and Wogalter (2001)*

	Visualization	*Picture simplicity*	*Open-ended comprehension*	*Population comprehension estimation*
Concreteness	.88*	.74*	.48**	.58**
Visualization		.73*	.53**	.62**
Picture simplicity			.40**	.46**

Note. All p <.01 ,
** present study, ** data described in Young & Wogalter (2001)*

Discussion

Results showed that all three ratings were significantly correlated to both open-ended comprehension and population comprehension estimation scores. The ease of visualizing the concept had the highest correlations with open-ended comprehension scores compared to the other two ratings. Results suggest however that any of the three ratings by them selves could be used to predict population estimates and open-ended comprehension of pictorial symbols that may be considered in the future.

While visualization only captured about 28% of the variance of the open-ended comprehension, the other two ratings did not account for additional

variance in multiple regression analysis beyond that accounted by visualization alone. Since only about a quarter of the open-ended comprehension score variance was accounted for by visualization, there is room for improving the prediction in future research.

The process of designing and testing pictorial symbols could benefit from a precursor test to determine the likelihood that a concept will permit the design of a successful symbol. Previous research has focused on alternatives, such as population comprehension estimation scores, which can predict open-ended comprehension scores. These alternative methods still dictate that a pictorial be developed to be tested. Our measures involve tests that can be performed before pictorials are developed.

Research has also shown that pictorial design can be costly and frequently inefficient (Wolff & Wogalter, 1993). Ratings of the ease of visualization or concreteness could be used early in the design process. This assessment could indicate whether there is likely to be difficulties in the designing an understandable pictorial. Preliminary testing of concepts could predict potentially difficult or impossible to produce pictorials. This testing could provide a basis for decisions on whether to pursue time-consuming and expensive processes in an attempt to develop a pictorial symbol for the concept. A concept that can not be easily converted to a simple concrete (and visualizable) pictorial symbol, might be better served by the use of textual statements.

References

ANSI (1998). American National Standard. Criteria for Safety Symbols. ANSI Z535.3-1998. Rosslyn, VA: National Electrical Manufacturers Association.

Brugger, C. (1994). Public information symbols: A comparison of ISO testing procedures. *Proceedings of Public Graphics*. (pp. 26.1-26.11).

ISO 9186 (1995). Procedures for the development and testing of public information symbols, Revision 1995. Geneva: International Organization for Standardization.

McDougall, S., Curry, M.B., & de Bruijn, O., (1999). Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. *Behavior Research Methods, Instruments, & Computers*, 31, 487-519.

McDougall, S.J.P., & Curry, M.B. (2000). Exploring the effects of icon characteristics on user performance: The role of icon concreteness, complexity, and distinctiveness. *Journal of Experimental Psychology: Applied, 2000,* 6, 291-306.

Wolff, J.S., and Wogalter, M.S. (1993) Test and devolvement of pharmaceutical pictorials. *Interface '93*, (pp.187-192). Santa Monica, CA: Human Factors and Ergonomics Society.

Wolff, J.S., and Wogalter, M.S. (1998) Comprehension of pictorial symbols: Effects of context and test method. *Human Factors, 40,* 173-178.

Young, S.L., and Wogalter, M. S. (2001). Predictors of pictorial symbol comprehension. *Information Design Journal,* 124-132.

Dewar, R. (1998). Design and evaluation of public information symbols. In H.J.G. Zwaga, T.Boersema, H.Hoonhout, *Visual information for everyday use* (pp. 285-304). London, UK: Taylor & Francis.

Zwaga, H.J. (1989). Comprehensibility estimates of public information symbols: Their validity and use. *Proceedings of the Human Factors Society 33rd Annual Meeting.* (pp. 979-983). Santa Monica, CA: Human Factors Society.