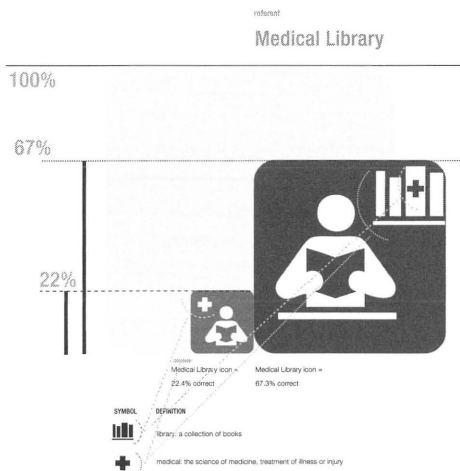


04 Improving Icon Design: Through Focus on the Role of Individual Symbols in the Construction of Meaning

MIKE ZENDER & G. MAURICIO MEJÍA

ABSTRACT

Despite the fact that icons are widely relied upon for communication, designers have few principles to guide icon design. This paper reports a study of the role individual symbols play on the construction of meaning from icons. An experiment compared two sets of four icons, each made of a different set of discrete symbols. It finds that the interaction of the right number of symbols for the referent, and a more apt combination of individual symbols for the referent, can significantly improve the construction of an icon that communicates what was intended. The rules of thumb proposed here are applicable to construction of any visual communication that uses symbols.



Icons today are ubiquitous and utilitarian. They shimmy on iPhones, bounce on computer screens, spin on cable TV's, and hang out on restroom doors. Icons are useful because they facilitate succinct communication. While their form is simple, their comprehension can be extensive. Indeed, nearly all communication happens through the interaction of symbols. Icons, ancestors of the earliest known forms of writing, have been a functional part of daily life since the pyramids were built so why study them now?

A sufficient reason would be that many icons are not understood as intended. The ISO (2007) and ANSI (2007) recommend 85% correct comprehension for all warning symbols. A 2010 USA Today article titled "One third of drivers can't recognize this idiot light" (Woodyard, 2010) reported that a tire inflation pressure warning icon mandated by law, was not understood by 60% of drivers: 46% couldn't even identify the symbol as a tire! Our own icon comprehension studies show depressingly similar results. Only eight of a set of 54 medical icons that were carefully designed to cross language and cultural barriers achieve 85% comprehension by subjects in the USA, and just 3 of those icons were comprehended at the 85% level by subjects in Tanzania. Indeed, fewer than 1 in 10 Tanzanians, many of whom had advanced medical training, could correctly identify 19 of the 54 medical icons. That's a failure rate of 90%.

Despite the common failure of icons, little is written about how they work from either a theoretical or a practical 'how-to' perspective. Beginning with Dreyfuss' *Symbol Sourcebook* (Dreyfuss, 1972) there has been steady parade of books that exhibit the latest symbols and icons, but few if any of these tomes explain how visual symbols work or how they might be made to work better. That is the gap our icon research seeks to fill. This paper describes a research study that measured the impact different combinations of symbols have on the comprehension of four icons. Based on this we identify some patterns, sketch some initial hypotheses for how people construct meaning from symbols, and propose some how-to rules of thumb to guide the design of more effective icons.

SYMBOLS AND ICONS

Besides being ubiquitous and utilitarian, icons are significant objects of study. Icons have simplicity of form compared with many other communication materials whose visual forms are much more complex. Icon's lean visual form reduces interpretive complexity. Icons also tend to have a very definite intended meaning: the referent...

..... Compared with other visual communications such as print advertising where the message can be nuanced, icon's intended message is clear, unambiguous and generally well-defined from the start of the design process.

This gives icons an established measure of comprehension success. Icons are typically created in a consistent graphic style. Since standardization efforts in the 1970's, notably the US Department of Transportation's commission of the AIGA to produce a standard symbol set, icons for a wide range of referents have followed the highly abstract round head and mitten hands familiar on restroom doors. Thus a wide variety of subject matter is available in a consistent visual style, facilitating study. While we are aware of one study that explores the effectiveness of this common style (Marom-Tock & Goldschmidt, 2011), similarity of style — however effective — has the benefit of reducing the number of variables in comprehension testing.

Despite the apparent simplicity, clarity, and consistency of icons, they are actually combinations of symbols. To clarify, a short digression into semantics may help. There are several descriptive words used for what we have been calling icons, variously called pictograms, symbols and signs. The *New Oxford American Dictionary* defines icons from a religious perspective: "a painting of Jesus Christ or another holy figure...". That dictionary's etymology of 'icon' traces it to the ancient Greek, eikon - 'likeness, image.' Biblical Greek translates eikon as image in Colossians 1:15 describing Jesus as "the image of the invisible God, the firstborn of all creation." Christian theology of the incarnation teaches that Jesus physically embodied the characteristics of God. Thus the early use of icon, or image, implies not mere physical resemblance but embodiment of the concept in visual form. Because historically 'icon' stands for this combination of physical representation that embodies concept, and because this integration of representation and concept can be widely seen today, we adopt icon as our preferred term for our study objects. We'll use the following taxonomy for this paper:

SYMBOL: *image referring to something else — a referent*

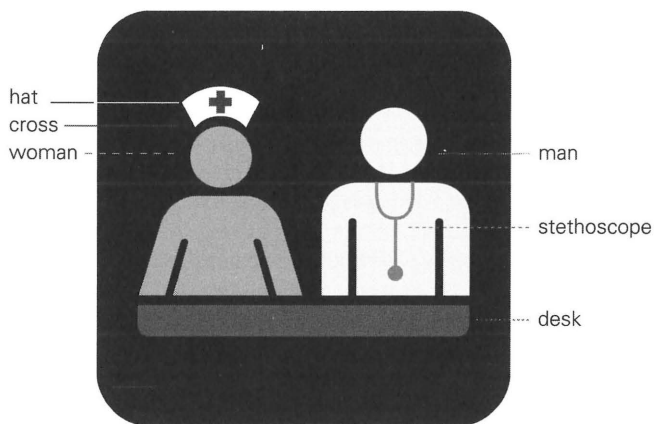
SIGN/GLYPH: *non-representational symbol, arbitrarily assigned with a wholly learned connection to a referent*

ICON: *representational image requiring no special learning for a categorical referent*

PICTOGRAM: *a combination of symbols and/or icons and/or glyphs to communicate a narrative or story or data set*

PICTURE: *representational image requiring no special learning for a particular referent*

Implied in our taxonomy is the concept that a single icon may integrate several different symbols to construct meaning. For example, the icon shown in Figure 1 combines six symbols: hat; cross; woman; man; stethoscope; desk; intended to communicate the referent "medical services." Sidestepping for the moment gender stereotypes, the man plus the stethoscope represent "doctor," the woman with the hat and cross on it represent "nurse," the desk represents "office or workstation." Taken together these individual symbols were designed to say in effect: "nurses and doctors (medical personnel) behind a desk to give information (not to treat patients)" which is the definition of the referent, "health services." An icon that combines multiple symbols to construct meaning is not an exception but the norm. In an article previously presented in this journal (Zender, 2006), we noted that many of the 50 icons in the 1974 AIGA/DOT Symbol /Sign system were combinations of symbols. In fact, 35 of the 50 AIGA /DOT icons use more than one symbol: that's 70%. In the medical icon system noted previously, only 7% of the icons use a single symbol, all the rest 93% (50 of 54), combine two or more symbols to convey a referent.



man + stethoscope = doctor

woman + hat + cross = nurse

doctor + nurse + desk = **medical services**

FIGURE 1 Six individual symbols interact to form a single icon to represent Medical Services.

To better understand how multi-symbol icons work, we conducted a study to explore the impact of changes in individual symbols within an icon on the construction of meaning. We aimed to improve the design of icons so that they are comprehended more accurately. Our first hypothesis was that the addition of more symbols to an icon could improve icon comprehension. We reason that comprehension of an icon's referent would be aided if the icon incorporated a symbol for each aspect of the referent definition. We also reason that a symbol more carefully matched to the referent definition would improve icon comprehension. In the previously mentioned 2006 article, context was identified as the key concept for decoding the meaning of symbols. The study reported here focuses on the 'immediate context' of symbols interacting within a single icon. It does not study the Environmental Context: the environment in which the images function, or the Proximate Context: the field of interaction where symbols in a system interact with other symbols in the same system to construct meaning. These important contexts were respected in this study by the methods used but were not the subject of the study. This study focuses instead on the impact of a change within the immediate context of a single icon brought about by adding or changing one of the individual symbols within that icon...

From this study we hope to gain some knowledge of how visual symbols interact to form meaning in icons and by extension, how symbols interact to form meaning across the entire field of visual communication.

STUDY

BACKGROUND

In 2009 the University of Cincinnati joined a five-school consortium brought together by SEGD and Hablamos Juntos, an organization devoted to improving healthcare for Latinos, to develop universal icons for healthcare environments. These icons, designed to cross language and literacy barriers in hospitals and clinics, were to supplement a previously developed set of healthcare icons that fit generally within the style of the 1974/1979 AIGA/DOT Symbol system. Teams of undergraduate design students at each institution developed candidate healthcare icons that were tested at four of the five schools using the ISO comprehension estimation protocol. (ISO, 2007) Test subjects spoke five different languages to insure universal comprehension independent of language and across cultural contexts. The result was a tested group of candidate symbols that were given to noted symbol designer Mies Hora for final development.

Mies recommended a number of refinements to icons including several that eliminated or replaced one or more symbols from an icon (see figure 2). For example, for the referent “Medical Library” the top scoring icon consisted of a dominant image of a person holding a book, at a desk represented by a single line waist high, with a shelf of books one with a medical cross in the upper right. The proposed revision simplified the icon by eliminating the desk and bookshelf. For another icon, Mies proposed keeping but changing contextual symbol: replacing the crescent moon symbol in the Inpatient Care icon with a clock symbol. Based on our previous work, we hypothesized that the elimination or alteration of the supporting symbols would affect comprehension. We specifically hypothesized that the removal of the smaller desk and bookshelf symbols would lower comprehension because it removed context; and that switching the moon and clock symbols would hurt comprehension because the clock contextual symbol less definitely suggested night, and therefore would not suggest overnight stay as well as the moon symbol.



FIGURE 2 Left column: the revised icons with fewer or less apt symbols (Group 2)
Right column: the original consortium icons with more or more apt symbols (Group 1)
Our study compared comprehension these icons.

STUDY DESIGN + AIM

The icon alternatives just noted provided us the opportunity to conduct a study of how individual symbol changes within an icon impact comprehension.

We labeled the two kinds of symbol change in a multi-symbol icon: quantity and quality. An icon could contain more or fewer symbols relative to the referent definition; quantity; or an icon could have a given number of symbols but those symbols could be more or less well matched to the referent definition: quality. We anticipated that if the multi-symbol icon lacked a symbol that was part of the referent definition that comprehension would suffer. Similarly, we anticipated that if a multi-symbol icon had a symbol that was less well-matched to the referent definition – less apt – that comprehension of that icon would likewise suffer. We suspected that absence of a symbol (quantity) would have a greater impact on comprehension than an imperfectly matched symbol (quality), reasoning from linguistics that an incomplete definition communicates less well than a definition with poor wording but that includes all the necessary concepts. We tested both kinds of change: quantity and quality. We selected icons for two referents: Medical Library and Health Services for the quantity test. Each had had a symbol removed. For the quality test we selected icons for two other referents: Inpatient Care and Nutrition each of which had a symbol replaced by different symbol. We designed our study to measure the effects of symbol change on comprehension. We expected that the icons with more symbols and more apt symbols would be better understood.

METHODS

This is a mixed methods qualitative and quantitative study. We wrote a modification of the existing *Hablamos Juntos* IRB protocol changing from a comprehension estimation method to the recommended ANSI Open-ended Comprehension Test method (ANSI, 2007) because this test is currently the most valid instrument for evaluation of icon comprehension. The test instrument asks a subject to imagine that he or she is in a given place (a health care facility for this study) and then answer two open-ended questions: the meaning of the icon and the actions that would be taken in response to the icon. The first question probes understanding at the level of abstract concept, the second at concrete action. Taken together, the two questions give an evaluator ample evidence to use to measure subject comprehension. An expert panel of three evaluators then independently scored subjects' qualitative answers. The quantitative scores were then analyzed.

We prepared two open-ended comprehension survey instruments to compare the two conditions between subjects: subjects in group 1 had icons with more and more apt symbols, while subjects in group 2 had icons with fewer and less apt symbols. Each test contained 4 icons, one each to represent the same 4 referents: Health Services; Medical Library; Nutrition; Inpatient. Figure 2 shows the eight icons used in the study. Two icons in group 1 contained more contextual information by way of additional symbols: the Health Services icon had an additional horizontal line symbol representing a desk; the Medical Library icon had a additional horizontal line symbol representing a desk and additional bookshelf symbol with a book having a medical cross symbol. The icons for group 2 lacked these symbols in the Health Services and Medical Library icons. The two other icons in group 1 had more apt symbols: the Nutrition icon for group 1 had a man instead of a triangle for group 2; the Inpatient icon for group 1 had a crescent moon symbol instead of a clock symbol for group 2. Our reasoning for these symbol substitutions was that the human symbol would better convey nutrition as a human health topic rather than a mere balance of foods, and that the crescent moon would better convey night than a clock which more generally represents time. Each test instrument had an identical instruction sheet that described the environmental context and gave an example of how to answer. After going over the instructions with the test administrator, each subject was allowed all the time necessary to write his or her answers on the test sheet, usually 15 – 20 minutes. The instrument administrator was present during the study only to answer necessary questions.

In spring 2010 a first cohort of subjects ($n=30$) was recruited, 15 randomly assigned to group 1, and 15 to group 2. Graduate students and Latino clients of a NGO composed this cohort. In autumn of that same year the study was repeated with a second cohort of 80 subjects ($n=80$), 40 randomly assigned to group 1, and 40 to group 2. Students visiting in a university student center composed this cohort. The total number of study subjects was 110 ($n=110$). Both cohorts were exposed to the same four referents to isolate the context variable for the study. The researchers discussed the classification of amount of contextual information and adequate context and used heuristics for decisions. Evaluators used a scoring sheet to independently score completed test instruments. A sample from the scoring sheet for Medical Library:

ICON 2: *Medical Library*

RESPONSE MUST INCLUDE: *medical or health or healthcare or hospital or clinic or doctor's office or care/care center, etc. plus library or books or book collection or reading room/area or information place/source, etc.*

Evaluators considered responses to both questions: what does it mean... what would you do... as a single answer to determine whether the subject correctly comprehended the icon. One of just three scores were assigned to each subject answer: correct, partially correct, incorrect. An example of a partially correct response would be a subject mention of library or books but not medical or healthcare or hospital for Medical Library. Scoring difficulties discussed elsewhere (Zender, Han, & Fernández, 2011) were largely overcome by use of three evaluators and by combining three forms of analysis. First, quantitative data analysis used t tests to compare comprehension means between groups. A value of 1.0 to correct responses, 0.5 to partially correct responses, and 0.0 to incorrect responses were assigned. Second, we used visualization to analyze the results and make comparisons. Third, we coded and analyzed the qualitative verbal answers. Using these methods we measured the changes in comprehension, if any.

FINDINGS

Comprehension was significantly better for the group 1 Medical Library and Inpatient icons that had the additional symbols. Comprehension was better, but not in a statistically significant way for the group 1 Health Services icon that had the more apt symbol, with no significant difference between the group 1 and group 2 Nutrition icons.

QUANTITATIVE ANALYSIS

In the quantitative data analysis, there was a significant difference in comprehension between the groups in the icons for Medical Library and Inpatient. In the Medical Library icon, the group exposed to more contextual information ($M = 0.736$, $SD = 0.407$) understood more than the group with less contextual information ($M = 0.400$, $SD = 0.400$). The mean difference was significant, $t(110) = 5.068$, $p = 0.001$. In the Inpatient icon, the group exposed to more apt contextual information ($M = 0.736$, $SD = 0.331$) understood more than the group with less apt contextual information ($M = 0.536$, $SD = 0.316$). The mean difference was significant, $t(110) = 3.774$, $p = 0.001$. There were no statistically significant comprehension differences between groups in Health Services and Nutrition icons.

Comparing the overall understanding including the four icon scores in each subject, the group exposed to more and more apt contextual information ($M = 2.618$, $SD = 0.938$) understood more than the group with less and less apt contextual information ($M = 2.009$, $SD = 0.781$). The mean difference was significant, $t(110) = 4.411$, $p = 0.001$. Therefore, contextual information is determinant of quality of icon comprehension as measured by understanding of meaning and actions that icons propose.

In comparing the means between the first cohort (n=30) and the second cohort of subjects (n=80), the trends were similar except for two slight differences. As shown in Figure 5, the Health Services icon in the first cohort had a lower mean (M = 0.450) than the second cohort (M = 0.650). In Nutrition icon, the first cohort had a higher mean (M = 0.600) than the second cohort (M = 0.482).

VISUAL ANALYSIS

Visualization of data is a proven method for gaining understanding. (Card, Mackinlay, & Schneiderman, 2003) The visualization in Figures 3.1 and 3.2 clarify the effect of two ways of quantitatively analyzing the data: the mean of the numeric rating of scores (1.0 = correct; 0.5 = partially correct; 0.0 incorrect), and the mean percent (%) of correct scores. The numeric rating scale incorporates partially correct scores with correct scores in a single number. As you can see in Figure 3.2, mean numeric scores inclusion of partially correct answers changes the picture: the advantage of the group 1 Medical Library icon is less pronounced when partial scores are included. Conversely, the superiority of the Medical Library icon is more pronounced when only considering % correct answers. The Nutrition icon shifts slightly from group 1 being better when only correct scores are considered to being slightly worse than group 2 icon when the numeric system accounts for the partially correct answers. One observation is that the simple question, "which icon is better" is more complex than it seems. Our results were significant, but clearly how you create those results has an important impact on the answer. These details matter. It is important for design research to operate at a level of nuance.

It is also extremely important to compare the full context when converting data to knowledge. The visualization in Figure 4 facilitates comparison by placing correct answers in context of all answers: partial and incorrect. Several interesting points stand out. The Medical Library group 1 icon clearly outperformed because of a very low number incorrect scores, while the group 2 icon for Medical Library had a very large proportion of incorrect scores. The group 1 Nutrition icon with the man instead of the triangle had either correct or incorrect scores, with very few partially correct, while the group 2 Nutrition icon had a significant number of partially correct scores. The group 1 Nutrition icon has a strong correct score but seeing that the number of incorrect scores is nearly equal to the correct suggests that subjects either get it, or they don't. The even dispersal of group 2 Nutrition icon scores suggests vagueness. A new design direction might build on group 1 seeking to clarify what some people get and others don't get. Both group 1 and 2 Health Services icons had a very large number of partially correct scores. This suggests the symbols integrated

in the Health Services icon are close to the right ones but not quite. In design language, this design needs 'tweaking' not a new direction. The same can be said of Inpatient group 2, except that the large number of correct answers for group 1 makes tweaking unnecessary – the right direction is clear: the moon tipped the icon from having a large number of partially correct to a large number of correct answers. None of these observations are obvious if you simply analyze the mean numeric scores or the % correct (figure 3).

mean numeric score

mean % correct

cohorts 1 + 2 (n=110)

cohorts 1 + 2 (n=110)

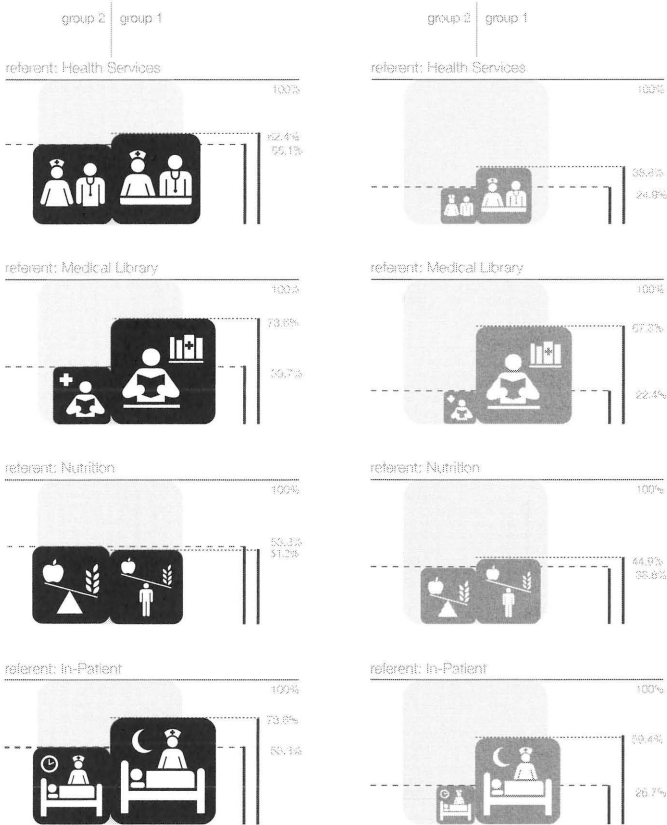


FIGURE 3.1 The mean numeric score for each symbol is shown in the column on the left. This score is the result of assigning a numeric value of 1.0 for correct, 0.5 for partially correct, and 0.0 for incorrect, adding the scores together and calculating the mean. The percent of correct icons is represented in the right column. This is simply the correct answers as a percent of all answers.

comparison: mean numeric and % correct

cohorts 1 + 2 (n=110)

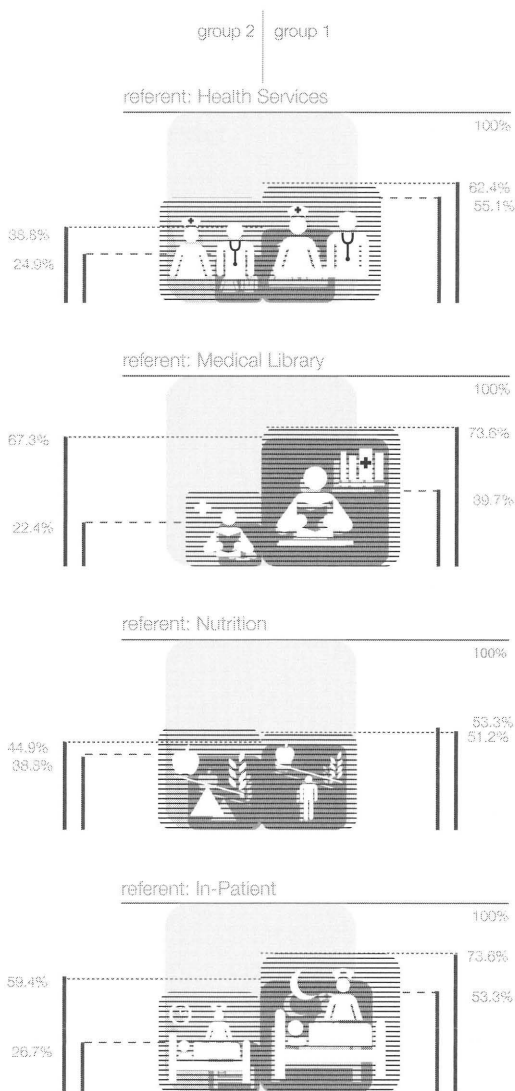


FIGURE 3.2 A comparison of mean numeric scores (transparent line pattern) and the % correct (solid gray fill) is shown. As you can see, numeric scoring inflates the apparent success of poorly performing icons. Nevertheless, more generous numeric scores were used to calculate statistical significance.

To verify the consistency of our data we also visually compared answers from the two different cohorts of subjects: cohort 1 with 30 subjects in spring 2010 and cohort 2 with 80 subjects in fall 2010 (*right column figure 4*). We found it interesting that cohorts 1 and 2 had similar results for the group 1 and 2 Nutrition icons, with the % correct being proportional for both icons, while the quantities correct were substantially different. Cohort 1 had eleven to eighteen percent more correct answers than cohort 2! This situation was reversed for the Health Services group 1 and 2 icons. While we were heartened that there was consistency of pattern between the two cohorts across the board, we have no clear explanation as to why the quantities were different between cohorts for Health Services and Nutrition icons. We can only speculate that cohort 1's lower number has more deviation and that the presence of numerous Spanish speaking subjects in that cohort skewed the results.

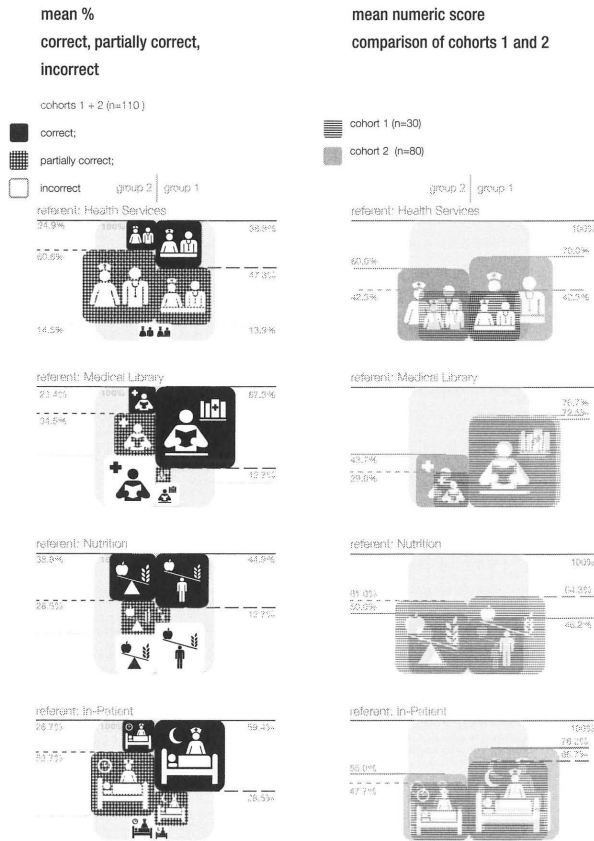


FIGURE 4 Visual analysis showing the full context of scores: correct, partially correct and incorrect, is shown in the left column. Visual comparison of mean numeric scores of the two subject cohorts: cohort 1 spring 2010, cohort 2 fall 2010, is shown in the right column.

QUALITATIVE TEXT ANALYSIS

Using a general inductive approach (Thomas, 2006), we analyzed the text of the raw subject answers to gain additional insight as to why some icons failed.

WHY WAS THE BOOKSHELF BETTER?

At the beginning of this paper we said our research program seeks to explain how visual symbols work and how they might be made to work better. As we have studied icons we have found that analysis of incorrect answers provides both insight for hypothesis formation and ideas for practical improvement. Our Analysis of subjects' written answers for the Medical Library icon shows that 8 of the 17 partially correct answers for the less successful "Medical Library" icon used the words 'read' or 'reading': 'reading room'; 'silent reading'; 'books available to read' are some examples. Clearly the general idea of reading was suggested by the man symbol holding reading material symbol. The problem is that a man can read a variety of things almost anywhere. The importance of this is illustrated by the incorrect answers to the same icon. Five of the 12 incorrect answers were 'waiting room.' Reading is often done both in waiting rooms and in libraries. The problem is that 'reading' is a more general concept that's related to, but not synonymous with, the more specific concept 'library.' Medical books however are generally only found in a medical library. As a result, we believed the key symbol in the icon was the medical bookshelf. We believed in this case the symbol of the man reading was functioning as a supporting symbol that aided interpretation of the rectangles and thin line as a bookshelf. Taken together, these symbols interacted to reliably communicate medical library. We believed that the rule under the man representing a table was supportive of the man-reading symbol, combining to communicate 'reading at a table' as opposed to 'reading in a lounge chair' which would suggest the different referent found in many incorrect answers: medical lounge /waiting room. We believed the reading at table better supported library.

As part of our analysis we considered comprehension of the book-shelf symbol with and without the reading man, and the reading man at the table with a reading man in a lounge chair. We believe the bookshelf symbol was the primary carrier of meaning, supported by the man reading symbol and the desk symbol. You might call the bookshelf primary, the man-reading secondary, and the desk tertiary symbols. All % correct scores are correct to the referent. As in all studies reported here, subjects were given the context of a medical facility.

To test this, we designed and conducted a follow-up study to measure the role of the medical bookshelf symbol in comprehension. As Figure 5 illustrates, the symbol for the medical bookshelf symbol consists of 4 vertical rectangles (1) + a horizontal line (2) + a cross

symbol (3). We tested each of these objects independently to see if and at what point the individual objects (1) + (2) + (3) communicated medical library. If the medical bookshelf alone communicated medical library, we wanted to know if it communicated as well as the icon that included the man-at-the-desk. It did, almost. The (1) + (2) + (3) bookshelf symbol alone, without the man, was comprehended correctly by 51% of 78 subjects, increasing to 90% when correct and partially correct scores were combined. This compares favorably with 67.3% correct and 80% combined score for the group 1 icon that combines bookshelf with the man-at-the-desk. This affirmed our suspicion that the key to success of the group 1 Medical Library icon was the bookshelf symbol.

follow-up study
role of bookshelf symbol in
Medical Library icon



FIGURE 5 As part of our analysis we considered comprehension of the bookshelf symbol with and without the reading man, and the reading man at the table with a reading man in a lounge chair. We believe the bookshelf symbol was the primary carrier of meaning, supported by the man reading symbol and the desk symbol. You might call the bookshelf primary, the man reading secondary, and the desk tertiary symbols. All % correct scores are correct to the referent. As in all studies reported here, subjects were given the context of a medical facility.

To help gauge the role of desk symbol (the line) in the man-at-desk symbol in the group 1 Medical Library icon, we tested an alternative icon with a man sitting in a lounge chair instead. This icon scored 50% correct. This is slightly lower than both the bookshelf alone at 51% and much lower than the bookshelf-with-man-at-desk icon at 67%. Most of the incorrect responses to man-in-lounge chair icon gave waiting room or medical lounge. This suggests that when users think of a medical library they picture in their mind primarily shelves of books and secondarily of people reading at tables. People-reading-in-a-lounge-chair suggests a different referent to many. We believe the group 1 Medical Library icon was more successful because it combined individual symbols each of which aligned with the mental image of the referent definition: Medical Library – a collection of medical books. This further suggests that the foundation of a successful icon is a clear understanding of the users' mental image of the referent definition and the creation of symbols that match it.

We identified the key symbol in the Medical Library icon as the bookshelf: a shelf on which books are stored. The bookshelf symbol is itself made of three symbols: (1) books, (2) a shelf, (3) a cross-on-book-spine. Like the Medical Library icon, we believed the bookshelf symbol/icon would work when it fully reflected the bookshelf definition. To test this we studied the bookshelf symbols separately and in combination. Seen alone, the books symbol (1) (4 rectangles) was a failure, only 7% correctly reading it as medical library or as books. Seen alone, symbol (2) shelf-under-books also performed dismally, at 17% correct. The 'shelf' line under the rectangles improved reading only slightly, but the addition of the cross on the 'spine' clarified the reading significantly: 51% correctly identified this as medical library. Clearly the combination of symbols (1) + (2) + (3) was necessary to communicate bookshelf.

As noted, library is a narrower concept than reading, and medical library is a specific type or subset of library, suggesting that for narrowly defined referents more symbols with narrower meanings are needed for accurate comprehension.

WHY WAS THE MOON BETTER?

Adding symbols to fit mental concepts improved comprehension, and exchanging symbols improved comprehension as well. The moon was better than the clock for communicating inpatient. We believed this is because moons are most apparent at night, while clocks are seen at all times and represent time generally, not just 'night' time. Indeed, text analysis of subjects' answers indicate that 'night' was more than four times as frequent in the answers to the icon with the moon symbol than in answers

to the icon with the clock symbol. Conversely, references to 'time' and clock' were many times as frequent in answers to the icon with the clock symbol. This is understandable since the icon included a clock symbol! The problem is that most peoples' mental image of inpatient is equivalent to an overnight stay, not a period of time measured by hours on a clock. Overnight is a more specific idea than time, it is a subset of time: night-time. We believe the clock symbol was too general, producing various misleading impressions.

As before, analysis of incorrect answers was revealing in this regard: the most frequent incorrect answer to the icon with the clock was 'visiting hours' or something similar. Note the association with time: hours. An interesting incorrect answer for this icon was 'nurse timing nap.' This reinforces an observation we have had in other studies that when people are deciphering a set of symbols in an icon they can be quite literal, particularly when they are having trouble getting the meaning. The icon with the clock contains 4 symbols: a man, in a bed, a nurse, and a clock. To that symbol set the interpretation 'nurse timing a nap' is completely rational...

..... *This reinforces that individual symbols in an icon interact to form meaning, and that changing one symbol in the group changes the meaning of the whole.*

We have one final humorous note. The more successful Inpatient icon with the moon had three incorrect answers, two of which were: 'dead body – avoid it'; 'dead people.' Perhaps the current preoccupation with zombies and vampires, think Dracula and the dark of the moon, provided a context where a moon was associated with death! In the future perhaps designers will routinely check users' mental images against the most popular interactive games, movies, and TV shows before selecting symbols for their icons so as to avoid unwanted gruesome associations!

SUMMARY CONCLUSIONS: A THEORY AND FOUR RULES OF THUMB

In this study we aimed to explain how symbols impact comprehension of an icon...

..... *We believe we have shown that icons embody multiple symbols, and that the interaction of those symbols evokes a specific concept in people's minds.*

This may seem like common sense but we don't think this has been stated this way before or demonstrated to this degree.

If icons can be seen as a collection of symbols contained in a common visual space to evoke a specific concept, then our findings might provide guidance for all symbol-based visual communication from concert posters to highway signs. We've shown that integrating symbols that more closely match the details of the referent definition can increase comprehension. This suggests both a possible theoretical foundation for symbol-based visual communication, and some practical rules of thumb for communication design practice.

POSSIBLE THEORETICAL FOUNDATION

There have been attempts to apply language theory to graphic design to lay a foundation for visual communication, semiology and semiotic theory being a notable example. One problem with applying semiology to visual communication is that Saussure (a founder of semiology) was a linguist. He appropriately emphasized the social construction of signs and the arbitrary assignment of meaning. This framework is a poor fit for symbol-based visual communication whose representational visual symbols have meanings directly linked to physical objects, not arbitrarily assigned to words. Without lapsing too deeply into digression we note that the relations between visual symbols also depend on the intended interpretive framework. Whole icons, as well as individual symbols that make them can be interpreted literally or metaphorically. W. T. J. Mitchell says, "The picture of an eagle in a Northwest Indian petroglyph may be a signature of a warrior, an emblem of a tribe, a symbol of courage, or – just a picture of an eagle." (Mitchell, 1987) It may be difficult to know the intended meaning of a particular visual symbol. Granted that there is considerable room for interpretation of visual symbols, nevertheless, a more appropriate foundation for visual communication might be founded on visual rather than linguistic processing.

Perceptual psychologist Stephen Kosslyn in *The Case for Mental Imagery* (Kosslyn, Thompson, & Ganis, 2006) explored how our brain uses, processes, stores, and manipulates mental visual representations to think and solve problems. He finds that humans store simple representations of familiar objects in a specific brain region and that these 'brain icons' can be recalled and pressed into visual working memory to think and solve problems. The existence of simple stored representations of common objects opens the possibility that people share a similar, or the same, 'brain icon' (our term, not Kosslyn's) for common objects. One of us explored this concept in 2011 in a series of workshops in India. In one workshop, each of the 75 participants was asked to draw a simple representation of a chair. The participants were then arbitrarily divided into groups

of 3 of 4 people. Each group was asked to draw a consensus version of the most representational chair, using their 4 individual chairs as the basis. Four groups were randomly selected to come and simultaneously draw their 'chair' icon, one on each side of 2 freestanding white boards at the front of the room: 1 'chair' on each side of 2 white boards: 4 chair icons. The class was surprised as they watched 2 identical chair drawings emerge as they were drawn on the whiteboards facing the class. They were amazed when the boards were turned around to reveal that the other 2 'chairs' were also identical to the 2 they had just watched unfold! The 4 drawings were identical: not only in having 4 legs, a seat and a back, but in the having the same 'canonical' point of view. This test has been done in various classrooms and in various locations with the same result: a typical chair with a canonical perspective.

Kosslyn's findings and our anecdotal experience exploring them in India and elsewhere suggests that a theory of how visual symbols are used by designers to construct meaning should account for the processes of visual perception, particularly the memory of simple iconic representations of familiar objects and concepts. This is a departure from the linguistic theoretical models offered previously, and appropriately so since it is now clear that language and visual perception use different neural processes. (Mayer, 2001) Our findings suggest that a successful icon is one whose symbols are most closely aligned with the mental images of users. This suggestion, built as it is on an established theory of visual processing and thinking, could lead to a theoretical basis for icon design.

FOUR RULES OF THUMB FOR ICON DESIGN

Our study is admittedly narrowly focused on the comparative success of 8 specific icons. Such a small study and such narrow findings are clearly insufficient to build a theory supporting specific principles or laws. However, we can tentatively offer three rules of thumb for icon design based on our study. Further research could raise these rules to the level of design principles.

1. MATCH SYMBOL TO DEFINITION

Individual symbols in the most successful icons were closer not only to people's mental images but also to their definitions of the referents...

⋮ *Designers should learn how people define the referent*
⋮ *then select the most accurate symbol for each concept*
⋮ *in that definition.*

Think for example of moon instead of clock for “overnight stay” or bookshelf rather than a man reading to represent “a collection of books”. Designers should first recognize that an icon is a collection of symbols and then match symbols in an icon to individual concepts of the referent definition. This means that designers need to expend effort to understand how people define the referent and then to conceive of symbols to represent each concept in the definition. Not every likely symbol will do. Sloppy symbol selection produces uncertain comprehension.

2. ADD SYMBOLS TO NARROW FOCUS

The first and second rules of thumb are related but where the first emphasizes symbol quality, the second refers to quantity. Fewer symbols are not better. The old saw “less is more” does not apply to icon design where a symbol for each referent concept is critical to success. Think for example of the Medical Library icon where adding the bookshelf symbol, making the icon more complex, improved it. Clearly the bookshelf and desk symbols were closer to “library” for the Medical Library icon than a “man reading” symbol alone. The “man-reading” symbol added important context for the “bookshelf” symbol but was not the key symbol. Both were needed for high comprehension...

⋮ *Designers should add specific symbols to help people shape*
⋮ *the correct referent meaning. More specific symbols provide*
⋮ *more substantial clues.*

3. CREATE SYMBOL HIERARCHY

In the follow-on studying of the Medical Library icon we found that the “bookshelf” was the primary symbol and that the “man-at-desk” was the supporting symbol. This suggests that a visual hierarchy of symbols exists within an icon and that placing the primary symbol at the top of visual hierarchy and supporting symbols at secondary levels might produce more accurate comprehension. In some icons a sequence of concepts is an important feature of the referent definition. For example, a referent saying “2 pills with meals,” is very different from “2 meals with pills.” For this kind of referent the symbol hierarchy could be particularly important to icon success. In other studies we have constructed visual stories using symbols in which reading sequence was absolutely critical to success...

: *Designers should consider the desirable reading sequence of*
: *symbols in relation to the referent definition, and then design*
: *an icon whose symbols have a corresponding*
: *visual hierarchy.*

4. STUDY FAILURE TO IMPROVE SUCCESS

The last rule of thumb does not deal with icon design per se, but with icon evaluation. First, we recommend that testing to evaluate communication be woven into the design process. Evaluation of comprehension is complex. (For difficulties in measurement of even simple icons, see Zender et al., 2011) However precise our current ability to evaluate symbols may be, it is clear that unsuccessful icons designs provide rich data to inspire improvement. We use the word inspire intentionally. Certainly, knowing a particular design approach fails helps to guide a designer toward one that succeeds. But we found that analyzing the answers to failing icons provides designers with more than guidance, it inspires designers to new creative possibilities they had not seen before. Our text analysis of raw user data of failed icons provided clear inspiration for improvement...

: *Designers wanting to improve comprehension would*
: *do well to study failing icons in order to gain insight for*
: *creative improvement.*

Designers seeking the best icon might even intentionally design various inaccurate icons to confirm right directions, validate suspicions, and inspire new directions. Little can be learned about improvement from a situation in which an icon succeeds.

IMPLICATIONS FOR FURTHER STUDY

This study evaluated icons, each constructed from group of individual symbols. We are so early in understanding of the interaction of visual symbols to construct meaning that it is appropriate to devote space to speculate about implications and suggest rich avenues for further study.

SEMANTIC RANGE

One idea comes from the study observations for the Medical Library icon. The symbol combination of the man holding reading material generated a wide range of answers from 'waiting area' to 'reading room.' Adding the bookshelf and desk symbols improved comprehension of 'medical library.' This suggests that the symbol

'reading material' or 'book' was too broad conceptually to communicate 'library,' but that the more narrow concept 'book shelf' significantly improved comprehension of 'library.' Borrowing a concept from linguistics: some visual symbols apparently have a narrow Semantic Range, while others have a broad Semantic Range. In relation to the Medical Library icon, the *New Oxford American Dictionary* gives the word 'book' 3 noun meanings indicating a broad semantic range (3 different meanings), while the word 'bookshelf' has a single noun meaning indicating narrow semantic range (1 meaning). Further study to explore the relation between semantic range of individual symbols and their interaction might lead to a model for identification of the semantic range of a desired referent and of candidate symbols the designer is considering to communicate it. Based on the study just reported, you might hypothesize that for accurate comprehension a narrower referent would need more symbols rather than fewer. This knowledge could be used to guide designers to more successful icons.

CONCEPTUAL FIT

Another outcome of the study was that the Inpatient icon was less well comprehended with a clock symbol representing the concept of 'time,' while the moon symbol to represent 'night' was comprehended better. In terms of semantic range the word 'clock' has one noun definition, and so does 'moon.' With differences in comprehension but similar semantic ranges, the failure of 'clock' and success of 'moon' suggests something in addition to semantic range is in play. The concept of good Conceptual Fit of each symbol to the referent is an idea that grows out of this study that might provide significant insight into icon communication. Since most icons combine several symbols with each symbol influencing the interpretation of the others, the role of each symbol should be weighed in relation to its Conceptual Fit to the referent. We noted for the Medical Library icon that the desk supported the man reading and that the man reading aided interpretation of the bookshelf symbol, and that the bookshelf symbol was the key symbol. Further study to explore the conceptual complexity of individual symbols and interactions might lead to a model for identification of the conceptual nuance of a desired referent and of candidate symbols the designer is considering to communicate that referent. This knowledge could be used to guide designers to more successful icons as well.

Icons are ubiquitous and utilitarian. They are good subjects for study because they promise to inform visual communication. We hope this report sheds some useful light toward the future design of icons as well as the full spectrum of visual communication.

.....
.....

REFERENCES

American National Standard Criteria for Safety Symbols, Z535.3 C.F.R. (2007).

Card, Stuart K., Mackinlay, Jock D., & Schneiderman, Ben. (2003). *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufman.

Dreyfuss, Henry. (1972). *Symbol Sourcebook An authoritative Guide to International Graphic Symbols*. New York: McGraw-Hill.

Graphical symbols - Test methods Part 1: Methods for testing comprehensibility, 9186-1 C.F.R. (2007).

Kosslyn, Stephen M., Thompson, William L., & Ganis, Giorgio. (2006). *The Case for Mental Imagery*. New York, NY: Oxford.

Marom-Tock, Yael, & Goldschmidt, Gabriela. (2011). Design for Safety: Symbol Genre in Safety Signs. Paper presented at the IASDR 2011 Diversity and Unity, Delft, Netherlands.

Mayer, Richard E. (2001). *Multimedia Learning*. Cambridge, UK: Cambridge University Press.

Mitchell, W. T. J. (1987). *Iconology Image, Text, Ideology* (paperback ed.). Chicago and London: University of Chicago Press.

Thomas, David R. (2006). A general Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, 27(2), 237.

Woodyard, Chris. (2010, Aug 03, 2010). One third of drivers can't recognize this idiot light, *USA Today*.

Zender, Mike. (2006). Icon Systems for Global Non Verbal Communication. *Visible Language*, 40(2), 177-206.

Zender, Mike, Han, Bing, & Fernández, Oscar. (2011). Did it Work? Evaluative Research Methods for Graphic Design. Paper presented at the IASDR 2011, Delft, Netherlands.