

Cognitive Design Principles for Visualizations: Revealing and Instantiating

Julie Heiser (jheiser@psych.stanford.edu)¹
Barbara Tversky (bt@psych.stanford.edu)¹
Maneesh Agrawala (maneesh@graphics.stanford.edu)²
Pat Hanrahan (hanrahan@graphics.stanford.edu)³

¹Department of Psychology
Stanford University
Stanford CA, 94305-2130

²Microsoft Research
One Microsoft Way
Redmond WA 98052

³Department of Computer Science
Stanford University
Stanford CA, 94305

Abstract

The creation of visualizations can be automated, informed by design principles derived from cognitive research. We describe a collaboration between psychology and computer science to uncover design principles for assembly instructions and instantiate them in algorithms. Participants assembled an object and then constructed diagrammatic instructions; those instructions were first rated and then tested. It was found that effective visualizations contain step-by-step action diagrams with 3-D views that best reveal the assembly actions. The design principles were incorporated into algorithms that generate assembly instructions from object models. The general techniques as well as the specific principles extend to other visual explanations.

Introduction

Visualizations are cognitive tools used to communicate, depict, instruct, and record information in nearly every domain, from architecture to chemistry, meteorology to radiology, economics to history. When designed properly, visualizations can augment human information processing (Tversky, 2001). Their external nature reduces cognitive load, provides common ground in collaborative tasks, and enables modifications and updates. Their spatial nature encourages more efficient spatial reasoning, such as inferences of proximity, relatedness, and common destination. Effective visualizations abstract essential information, sometimes distorting it, and eliminate irrelevant details.

Perhaps due to their effectiveness in communication and inference, perhaps due to advances in computing, the demand for visualizations increases. Individualizing visualizations is a job for computers, but computers, like designers, need to be educated. In order to create effective visualizations, computers must “know” what information is essential and what is irrelevant; they must also “know” how to present that information so that people can perceive and understand it. Creating effective visualizations by computer then, requires a partnership between cognitive psychologists and computer scientists. The task of the cognitive psychologist is to uncover the desired mental representation and to evaluate the comprehensibility of the design; that is,

to uncover the specific design principles (e.g. Kosslyn, 1994; Mayer, 2001). The task of the computer scientist is to incorporate the design principles into algorithms. This may sound simpler than it is. For one thing, the design principles are likely to be in the form of tradeoffs, with the potential for inconsistency and conflict; for another, they are unlikely to be sufficient.

Line-drive (Agrawala & Stolte, 2001) is an example of such a partnership. This algorithm generates sketch-like maps rather than the highway maps with routes superimposed that websites typically produce. Sketch route maps differ from highway maps in a number of ways (Tversky & Lee, 1998). They show only the route, with some extra spatial information as protection against error. They take liberties with distance and direction information in the service of conveying the essential information clearly: paths and node, points of action. Thus, the visualizations produced by students and those produced by the algorithm distort distance, direction, and other information, much like the heralded London subway map. Despite the distortions, or perhaps because of them, the sketch maps produced by mapblast.com have met with praise by users.

There are many domains in need of improved visualizations as well as individualized ones. One notorious domain is assembly instructions. Optimistic consumers see attractive barbecues or desks in stores, bring them home in their boxes, pull out the “easy” instructions, and are all too often greeted with a cluttered exploded diagram, such as that in Figure 1, perhaps impressive as an image, but hardly functional as a set of instructions. Small parts can’t be discerned, nor can how and where they should be attached. What to do when is not evident. The authors have collected a corpus of assembly instructions, and these problems prove to be widespread and frequent.

Assembly instructions are of interest not only in themselves but also because they are representative of a larger class of visualizations that includes instructions on how to operate something, like a copier or a nuclear plant, and how complex systems work, like a heart or a corporate structure.

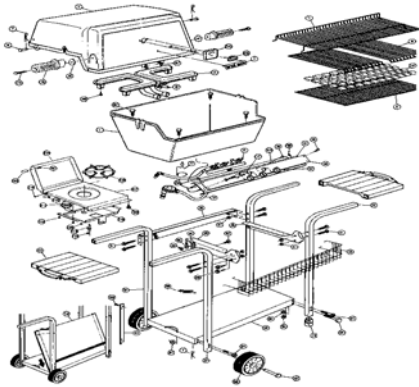


Figure 1. Instructions for assembling a barbeque.

How can cognitive design principles be revealed? For the project on route maps, knowing how people think about routes as well as how they produce sketch maps provided many design principles. For assembly of objects, it helps to know how people think about objects; people think about objects in terms of their parts, first and foremost, the parts that are perceptually salient and functionally significant (Tversky & Hemenway, 1984). Next, it helps to know how people think about assembly; people think about assembly as hierarchically organized sequences, where objects or significant object parts separate steps at the higher level and differentiated actions on the same object at the finer level (Zacks, Tversky, & Iyer, 2001). More suggestions come from a study of instructions for constructing origami figures, where instructions that show each step were more effective than instructions that only showed start and end states (Novick & Morse, 2000).

These guidelines are useful, but need to be supplemented by guidelines that are specific to assembly. To uncover people's conceptions of routes, Tversky and Lee (1998) asked participants to produce route maps and route directions. Both route maps and route directions schematized the information in ways that people found efficient and useful. The route sketches also provided guidelines for the information that should be captured in visualizations of routes and guidelines how to display it. We adopted the same strategy here, to simultaneously uncover the information that is essential to convey as well as how to convey it.

We adopted the methods of Tversky and Lee (1998) to extract people's conceptions of object assembly and effective visualizations of the assembly. This part of the work has three stages. In the first, participants assembled a simple object—a TV cart—using only a photo of the assembled object; then they produced instructions for others. Next, new participants assemble the TV cart, and rate the instructions produced by the first set of participants for quality and effectiveness. The design features that appeared in the highly rated instructions were abstracted. Finally, a third group of participants assembled the TV cart

using instructions with high or low ratings; their performance was evaluated in terms of the quality of the instructions.

Experiment 1: Generate Instructions

Students assembled a TV cart using only a photograph of the assembled object. Then, they produced instructions to assemble it.

Method

Participants. Forty-five Stanford University undergraduates participated for pay. The data of two participants were eliminated as they had participated more than once.

Spatial Ability Measures. Participants first completed a questionnaire on their previous experience assembling objects. Then they completed the Vandenburg and Kuse (1978) test of mental rotation and the Money Spatial Navigation Task (Money & Alexander, 1966), a 1-minute test that evaluates egocentric perspective transformations. These measures have correlated with performance in many tasks.

Object Assembly. Participants assembled a TV stand from parts, connectors, and tools provided using the photograph on the box as a guide (see Figure 2). Assembly required five major steps, each corresponding to a major part of the stand.



Figure 2. Picture of box and completed TV Stand.

Constructing Instructions. Upon completion of assembly, participants were asked to write instructions for a novice to “easily and efficiently assemble the TV stand.” They were told they could use text and diagrams; whatever they felt could convey the process most effectively.

Results

Performance on Tasks. All participants were able to assemble the TV stand without instructions. Performance on the assembly tasks was highly related to scores on the

individual differences measures. Experience also correlated with an average of the ability measures ($r = .61, p < .01$). Participants were divided into high and low spatial ability by a median split of the average of the two spatial ability tasks, yielding 21 low and 22 high spatial participants. Participants had to perform below average to be included in the low category, and above average in the high spatial category. Low spatial participants took 12.7 minutes ($SD = 3.56$ min) to assemble the TV stand, while high spatial participants completed the task in 7.3 minutes ($SD = 2.09$ min), $F(1,41) = 36.01, p < .01$. Low spatial participants made more errors during assembly, such as putting pieces in upside down, putting the unfinished side of the wood facing outward, etc. These errors were manifest in the instructions participants wrote where 86% of low spatial participants included an error of an “impossible action” in their instructions, whereas only 12% of instructions produced by high spatial participants had such errors, $t(1,41) = 5.9, p < .01$.

Instructions Produced. Differences between high and low spatial participants appeared in the sketches drawn in the instructions. High spatial participants produced 2.67 action drawings on average. Action drawings show how parts are assembled. By contrast, low spatial participants produced .64 action drawings per instruction set, $F(1,41) = 16, p < .01$. Conversely, low spatial participants included 1.45 drawings that depicted the structure of the system, but high spatial participants produced only .81 structural drawings per instruction (see Figure 3 for examples), though this difference was not significant due to high variance. Action diagrams necessarily depict structure, so the majority of drawings produced by the high spatial participants depicted both action and structure. Low spatial participants were more likely to include sketches of parts on their own (low had mean of 4.14 compared to high mean of 2.19, $F(1,41) = 5, p < .05$). More differences manifested, such as, high spatial participants were more likely to include complex 3D pictorial or exploded diagrams in instructions.

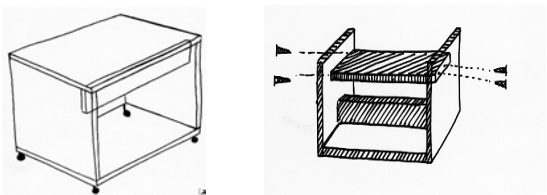


Figure 3: Diagrams Produced by Participants. A. Structural Diagram. B. Action Diagram

Discussion

Undergraduates successfully assembled a TV stand using a photograph of the completed stand as a guide. High spatial participants completed the task nearly twice as fast as low

spatial individuals. Then, they constructed instructions for assembly. The assembly steps adopted corresponded to the major components of the TV stand. High spatial participants produced instructions with rich and informative drawings, specifically, 3-D step-by-step perspective drawings that showed assembly actions. Low spatial participants produced fewer diagrams, and were more likely to produce 2-D pictorial lists of parts or structural diagrams. The efficacy of the diagrammatic techniques used by the high spatial participants is rated in the next study.

Experiment 2: Rate Instructions

Do other participants prefer the diagrammatic techniques spontaneously adopted by the high spatial participants in the first experiment? Given the large differences in drawings produced by high and low spatial participants, are there parallel differences in their preferences? Here, we address those questions by having a new set of participants rate a subset of the instructions produced in the first experiment.

Method

Participants. Twenty-one undergraduates from Stanford University participated in this experiment to fulfill a course requirement.

Ability Measures and TV Stand Assembly. The procedures for this part of the experiment were identical to those in the first experiment. After assembly, instead of producing instructions, participants rated the previous ones.

Rating Instructions. Thirty-nine sets of instructions from the first experiment were rated. Four sets were eliminated due to high similarity to other instructions in the sample. All written text was typed and sketches redrawn to equate clarity, legibility, and aesthetics. The order of instructions given to raters was randomized to avoid ordering effects. Participants rated the quality and effectiveness of instructions from very poor (1) to excellent (7). They were also asked to write comments explaining their ratings, specifically to describe the features of the sets of instructions they liked or disliked.

Results and Discussion

There was general agreement on the quality and the good and bad characteristics of the instruction sets across participants. Ratings of the 21 participants correlated highly, $r = .97, p < .001$, and there were no differences in preferences of high and low spatial/experience participants.

The characteristics of instructions that received the highest ratings were: action diagrams, or diagrams that showed assembly of parts: explicit ordering of assembly; use of extra-pictorial elements, such as arrows and guidelines to

convey directions of actions; step-by-step diagrams; precise use of text; clear portrayal of actions showing connector pieces and how to use them; and use of 3-D perspective to convey views depicting needed actions.

Instructions receiving low ratings included those that had no diagrams, just part menus, poor representation of proportions for parts, diagrams with lengthy text, instructions that left out steps, and instructions that did not specify the order of assembly. Participants commented that these features would lead to confusion and errors.

Experiment 3: Test Instructions

Results from Experiment 2 allowed us to recognize important features of instructions. In Experiment 3 we undertook preliminary testing of instruction sets selected from Experiment 1 & 2. These instruction sets were reconstructed in order to test different features of instructions, with the goal of extracting guidelines for the design of instructional visualizations.

Method

Participants. Forty-four undergraduates from Stanford University participated in this experiment to fulfill a course requirement. They were randomly assigned to one of the four instructions sets.

Ability Measures. As in Experiments 1 & 2, participants completed a questionnaire about their experience with assembling objects and two spatial ability tests.

Instructions. Four sets of instructions were selected that used the highly rated features, such as step-by-step instructions, use of action diagrams, clear indication of ordering, and explicit indications of part attachment. Each included various of the features that were preferred in varying degrees, such as use of explanatory text, presence of menu of parts, and integrated text and diagrams.

Assembly of TV Stand. Participants were told to use the instructions provided to assemble the TV stand. All participants were videotaped. After assembling the TV stand, they were queried as to the features of the instructions they found helpful or not.

Results and Discussion

Instructions improved performance of low spatial participants considerably compared to performance in the first experiment, unaided by instructions. Without instructions in the first experiment, low spatial participants took 12.6 minutes to assemble the stand; with instructions in this experiment they took only 9.4 minutes. High spatial participants showed no benefit from instructions,

performing as they had without, assembling the TV stand in 7 minutes. The videos revealed that high spatial participants consulted instructions only 3.3 times, whereas low spatial participants consulted them 9.1 times ($F(1,42), p < .01$). This suggests that for this relatively easy task, high spatial participants did not need instructions. With instructions, low spatial participants made fewer errors than in the first experiment though they made more errors than high ability participants ($F(1,42) = 5.8, p < .05$).

Although there were no differences in performance as a consequence of the four versions of the instructions, there were differences in the features regarded as important in the post-assembly interview. The lack of performance differences reflects the fact that all instructions contained the features most highly rated and most frequently mentioned.

Participants' total assembly time was divided into the time it took to complete each step. Interestingly, the only step that was significantly longer for low spatial/experience was the first step, where low participants took an average of 2.29 minutes and high participants took an average of 1.04 minutes. $F(1,42) = 7.17, p = .01$. Analysis of the videotapes revealed that low ability participants spent that extra minute orienting themselves to the parts, looking at the connector pieces, inspecting the instructions again, whereas the high ability/expertise participants began assembling almost immediately.

Responses in the post-assembly interview indicated that participants found extracting information from the instructions easier from action diagrams than from text explanations. Several participants reported that they never read the text that described actions portrayed in diagrams, but that they may have needed it if the assembly was more complicated.

Experiment 3 confirmed and added to what was learned from Experiments 1 & 2, as well as providing more data from which to extract design principles. For example, having a menu of parts will be helpful in orienting the low spatial/experience participants, text may be needed as a buffer, but clear, explicit diagrams are more helpful, and if text is used, it should be succinct.

Applying Design Principles

Overall, the experiments have provided cognitive guidelines that are directly relevant to assembly instructions and pertinent to other kinds of visualizations as well, particularly those that show processes over time. Here are some of the features that are important for effective visualizations:

- Step-by-step, one for each major step
- Clear and explicit order
- Parts added in each step should be visible
- Mode of attachment should be visible
- Action diagrams rather than structural
- Arrows and guidelines to indicate attachment

- Show perspective of attachment
- Avoid changing perspectives
- Show stable orientations

Note that some of these desiderata can conflict, notably the preference for showing the perspective of attachment and the preference for minimizing perspective switches. This illustrates just one of the challenges of instantiating cognitive design principles in algorithms. A second challenge comes from under-specification; the design principles are moot on many design decisions that must be taken, for example, whether to show shadows or texture. Yet another shortcoming is revealed in unpredictable emergent properties. Some of the perspectives are unstable, that is, they reverse depth like a Necker cube.

Despite the limitations, the design guidelines have been instantiated in a working program (Agrawala, Phan, Heiser, Haymaker, Klingner, Hanrahan, and Tversky, 2003).

Automated Assembly Instruction Design

We have developed an automated assembly instruction design system based on these principles. As input our system requires a geometric model of the assembly with each part in its final assembled configuration and a default orientation for the assembly. Users can optionally specify additional semantic information including:

Groupings: Labeling of parts that should be grouped together based on functional or geometric properties. The grouping labels include *fasteners*, *significant parts*, *symmetry*, and *similar-action*.

Ordering Constraints: Constraints on the order of assembly operations.

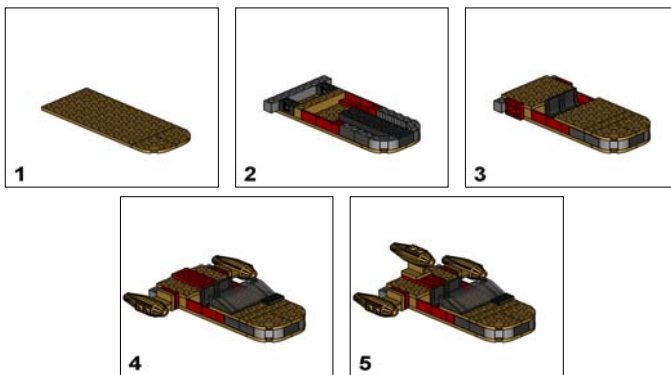


Figure 4. A structural assembly sequence of a Lego Landspeeder generated by our system.

Our system divides the task of generating assembly instructions into two phases; *assembly planning* and *diagram production*. In the assembly planning phase the system analyzes the geometry of the model to compute the sequence of steps required to build the object. Assembly

planning algorithms have been well-studied in robotics (Romney, Godard, Goldwasser and Ramkumar, 1995, Wilson, 1992). These algorithms first compute all geometrically feasible assembly sequences for the object by analyzing the blocking relationships between all pairs of parts. This analysis yields a directed acyclic graph encoding the removability constraints on each part and each valid topological sort of this removability graph produces a geometrically valid assembly sequence for the object.

The goal of these algorithms is to plan a sequence of instructions that a machine-tool could use to build the object. Each valid assembly sequence is evaluated for the particular machine-tool it will be implemented on and the most cost-effective sequence is chosen as the final assembly plan. But, these assembly planning techniques do not consider the requirements of human builders and therefore, the resulting instructions are usually very difficult for humans to follow. As our human-subject experiments show, for an assembly plan to be easy to follow it must be both geometrically feasible and adhere to the principles described in the previous section.

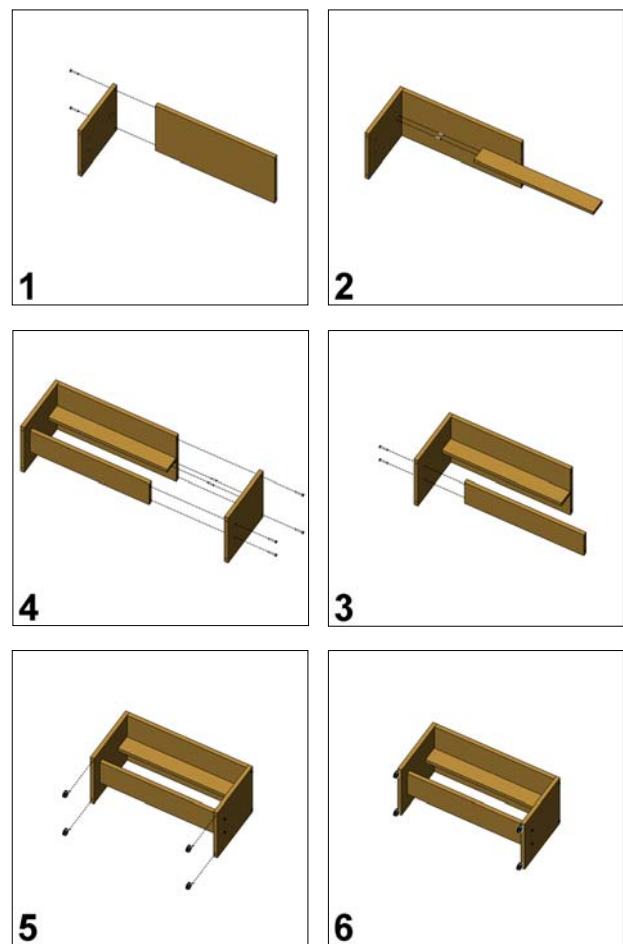


Figure 5. Assembly instructions for TV Stand generated the automated system described here.

Therefore, our system first computes the set of geometrically feasible assembly sequences using the removability analysis technique developed for robotics. But, unlike prior systems we evaluate the resulting assembly sequences based on the cognitive design principles revealed by our experiments. Thus our system is able to produce step-by-step assembly sequences that are well-designed for human builders.

Once we have computed the assembly plan we must generate a diagram showing each step in the plan. We can generate structural sketches by simply showing the set of parts added in each step in their final assembled positions. The assembly planning stage of our system ensures that all of the parts added in each step and their points of attachment will be visible in each step. The resulting instructions are similar to those included with Lego (see Figure 4).

The diagram production stage of our system is also capable of generating action diagrams (see Figure 5). Instead of placing each new part in its final assembled position, we offset the new parts from the earlier parts and use diagrammatic elements to indicate how they attach to one another. These dotted lines and arrows indicate the motion of the parts as well as the points of attachment between them.

Conclusions

Good visualizations for assembly make a difference. They improve performance especially for those in need, those lower in spatial ability and experience. This applies to visualizations in other domains as well. Visualizations can facilitate understanding, inference, and insight. Designing visualizations does not have to be an art at best or guesswork at worst. Cognitive experiments can take much of the art and guesswork out of design by revealing how people think about the system visualized and how they convey that information. From these, guidelines can be extracted that can be evaluated. Finding design principles is an iterative process, benefiting from both performance of experienced people as well as their evaluations of features that aided or hindered performance.

Principles for designing effective visualizations can be explicit enough to be instantiated in computer algorithms that generate individualized visualizations on demand. Previous work instantiated cognitive design principles for route maps into an algorithm that generates sketch-like route maps on demand. The present paper describes a project that reveals and implements design principles for assembly instructions.

Two implications emerge from this collaboration. Many of the specific guidelines for assembly are applicable for other visualizations, such as visualizations for operating a device or explaining a system. More generally, the procedures for eliciting cognitive design principles can be

applied to other cognitive tools, and could have great impact in educational domains as well. Understanding what types of visualizations elicit accurate and complete mental representations can provide insight into the nature of knowledge representation and information processing in human cognition.

Acknowledgments

Special thanks to Christina Vincent, Rachel Mackenzie, Heesoo Kim and Vincent Pham for help in data collection and analysis. This research is funded by ONR grants N00014-PP-1-O649, N000 14-01-1-0717, and N00014-02-1-0534

References

- Agrawala, M. & Stolte, C. (2001). Rendering Effective Route Maps: Improving Usability Through Generalization, *SIGGRAPH 2001*, August 2001.
- Agrawala, M., Phan, D., Heiser, J., Haymaker, J., Klingner, J., Hanrahan, P., and Tversky, B. (2003, in press). Designing effective step-by-step assembly instructions. *SIGGRAPH2003*, July 2003.
- Kosslyn, S. (1994) *Elements of Graph Design* NY: Freeman.
- Mayer, R. E. (2001). *Multimedia Learning*. New York: Cambridge University Press.
- Money, J., and Alexander, D. (1966). Turner's syndrome: further demonstration of the presence of specific cognitive deficiencies. *Journal of Medical Genetics*, 3, 47-48.
- Novick, L. and Morse, D.L. (2000). Folding a fish, making a mushrooms: The role of diagrams in executing assembly procedures. *Memory & Cognition*, 28(7), 1242-1256.
- Tversky, B. (2001). Spatial Schemas in Depictions. In M.Gattis, Editor. *Spatial Schemas and Abstract Thought*. Cambridge: MIT Press.
- Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113, 169-193.
- Tversky, B, Zacks, J., Lee, P. U., & Heiser, J. (2000). Lines, blobs, crosses, and arrows: Diagrammatic communication with schematic figures. In M. Anderson, P. Cheng, and V. Haarslev (Editors). *Theory and application of diagrams*. Pp. 221-230. Berlin: Springer.
- Vandenburg, S.G. and Kuse, A.R. (1978). Mental rotations. A group test of three-dimensional spatial visualization. *Perceptual Motor Skills*, 47, 599-604.
- Zacks, J., Tversky, B., and Iyer, G. (2001). Perceiving, remembering, and communicating structure in events *Journal of Experimental Psychology: General*, 130(1), 29-58.