Using Diagrams to Design Information Systems[1]

James E. Corter (corter@tc.edu) [a]
Jeffrey V. Nickerson (jnickerson@stevens.edu) [b]
Barbara Tversky (tversky@tc.edu) [a]
Doris Zahner (dzahner@stevens.edu) [b]
Yun Jin Rho (yjr2101@columbia.edu)[a]

a Department of Human Development, Teachers College, Columbia University
525 W. 120th St, New York, NY 10027 USA
b Howe School of Technology Management, Stevens Institute of Technology
Castle Point on Hudson, Hoboken, NJ 07030 USA

Abstract

Designers use sketches to offload memory and information processing and to promote discovery and inferences. Designers of information systems depend on topological connectivity rather than Euclidean distance. For such designers, understanding graph topology and manipulating graphs are essential skills, because graph manipulation facilitates the generation of alternative designs. We found that students of systems design have difficulties interpreting diagrams, revealing two biases: a sequential bias and a reading order bias. These biases lead to systematic errors of omission and commission while inferring connectivity from a diagram. The results have implications for teaching as well as diagram design.

Keywords: diagrams; design; graphs; topology

Introduction

Design entails arranging and rearranging real or virtual objects and parts and comparing and weighing their functions and goals. Although the mind seems to have almost unlimited space to passively store information, its space for actively manipulating information is highly limited. When the mind runs out of mental space, it often turns to externalizations, such as sketches, diagrams, charts, and models.

In many cases, diagrams are meant to represent more than just physical structure. Extending diagrams from representing structural information to representing functional or abstract information often requires the addition of symbolic conventions: lines, boxes, arrows, words, and more (e. g., Tversky et al. 2000). This information can be ambiguous; arrows, for example, can indicate causation, sequence, or flow, among many other meanings (cf. Nickerson 2005).

Sketches and diagrams, then, must be interpreted in order to be used. Their very Euclidean character, the metric properties of diagrams — distances, angles, sizes, shapes and their proximity — are difficult to ignore, even when irrelevant (e. g., Landy and Goldstone, 2007), and can encourage false inferences. Although diagrams and sketches present information in parallel and do not privilege any direction or location over any other, the mind does not process them in parallel; rather, they are interpreted sequentially. When there is a natural beginning, diagrams are "read" from there, but when there is not a natural beginning, diagrams tend to be scanned in reading order, in Western languages from left to right and top to bottom (e. g., Taylor & Tversky, 1992). The richness and complexity of diagrams render them more useful but also more problematic at the same time.

Systems designers use system diagrams to document the topology of a network, and to plan the flow of information. The links in a network may be hierarchically organized in subtle ways. For example, at the infrastructure level, network bridges and routers partition networks in order to control performance and security. The designer of an information system is expected to understand such network topologies and the diagrams used to represent them. In order to successfully create and interpret systems diagrams, students and practitioners must learn to suppress conscious or unconscious inferences based on Euclidean properties of diagrams, such as the planar distances among nodes, and learn to rely instead on properties of graphs: diagrams composed of nodes and edges. How well do novices and experts understand these conventions and the formal structure underlying them? Can they interpret and generate the paths that a topology implies? Can they suppress geometric intuitions when necessary?

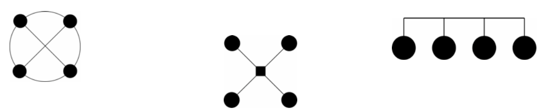
Study 1: Understanding and Producing Network Topologies

To understand how expert and novice students produce and understand systems diagrams, we presented design problems to students in a Master's level class in the design of systems (see Nickerson, 2006) at the beginning of the semester (Study 1) and at the end of the semester (Study 2). The students varied widely in prior programming and design experience.

We expected two biases in students' production and comprehension of diagrams: a sequential bias and a reading order bias. Because diagrams are presented in Euclidean

[1] This paper is abridged and adapted from Nickerson, et al. in press. Specifically, only three of the problems presented to student designers are discussed here and an additive tree analysis of errors is added.

space, students would be biased toward Euclidean interpretations and thus have difficulties comprehending the topological relations, indicated by lines. Specifically, they would have difficulty making interpretations based on *connectivity* rather than *proximity*. We also expected students to have difficulties with hierarchical concepts, in particular, comprehending and using a logical *bus*. A bus is a sub-cluster of components that are mutually interconnected. Most local area networks are organized in this way. By convention, busses are indicated by a line with satellite lines as in Figure 1c. Within a bus, all nodes are interconnected, even though this is only implied in 1c, unlike in Figure 1a where the interconnections are explicitly shown. Figure 1b shows a hub and spokes model. All three of these diagrams represent the same functional topology.

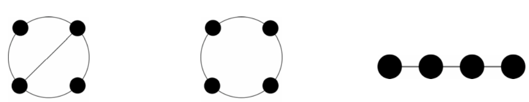


a) Complete Graph b) Hub and Spokes c) Logical bus

Figure 1. Alternative (correct) LAN representations

Presented with the bus diagram (Figure 1c), a student without knowledge of this diagram type might wrongly infer that a path between the far left node and the far right node must pass through the middle nodes, when in fact, the two extreme nodes can connect directly. Moreover, a student who misunderstands modern local area network technology may portray it using one of the inappropriate or obsolete structures shown in Figure 2.

A second expectation was that students would by default inspect and interpret diagrams in reading order, which would bias them to see paths more compatible with reading order and to miss paths less compatible with reading order.

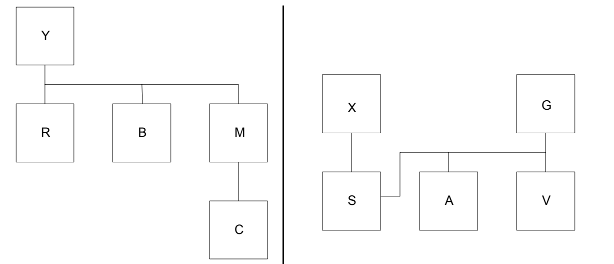


a) Incomplete Graph b) Ring c) Chain

Figure 2. Obsolete or incorrect LAN representations

Method

Sixty-eight students from four different sections of the same course were presented with a set of design problems to be answered in class (the 3 problems analyzed here are shown in Figure 3). For Problems 1 and 2, they were presented with a diagram of the configuration of a system and asked to generate all the shortest paths of information through the system. This is an important type of inference in systems design as it is a check that information flows according to systems constraints. Problems 1 and 2 had the same network topology, but a different embedding on the plane. Problem 3 asked students to generate a diagram from a verbal description of the same topology as Problems 1 and 2.



Problem 1 Problem 2

Imagine a network configuration in which Nodes A, B, C, and D are on a shared Local Area Network (LAN) and Node E is connected to node B through a different networking connection (assume machine B has two networking cards, one to connect to the shared LAN, and one to connect to node E). List all shortest paths between all pairs of nodes.

Problem 3

Figure 3. The topology test

All problems described two networks joined by a bridging node (M, S, and B respectively). Students were given an example and then asked to enumerate all shortest paths between the nodes in the graph. For all problems, there were twenty minimal paths (Problem 1: YR, YB, YM, YMC, RY, BY, MY, CMY, RB, RM, RMC, BR, BY, CMR, BM, BMC, MB, CMB, MC, CM). To answer correctly, students must understand (for example) that all shortest paths with C as a terminal in Problem 1 need to go through M, but that the shortest paths to Y from B, M, or C do not go through R. This presupposes they understand the diagramming convention, the concept of a shortest path, and the partitioning role that bridges play.

For the problems (1 and 2) requiring generation of shortest paths from given diagrams, two types of errors were possible. The first is a commission error, listing paths that were not shortest paths. Commission errors are a consequence of not understanding the essential concepts taught in the class, either the concept of shortest path, or the topology represented by the bus convention. For example, in Problem 1, listing YRBMC is a commission error because the shortest path between Y and C is YMC.

The second type of error is an omission error, failing to list one or more shortest paths. Omission errors can reflect conceptual confusions or simple carelessness. If students generate paths in reading order, starting at upper left and proceeding left to right and top to bottom, they will be more likely to omit backwards paths than forwards paths, where forwards means starting upper left and backwards means starting lower right.

For the third problem, the diagram generated should be diagnostic. Chains or rings would suggest a sequential bias, and should be associated with more commission errors.

Results

Problem solutions were coded for commission and omission errors and for solution strategy. The *reading order bias* predicts that students should begin by listing forward paths, (i.e. paths starting from the upper left) and list more forward than backward paths, because the latter may be inadvertently omitted. Some students, however, may have *presupposed* reversibility of paths; that is, they may have intentionally only listed forward paths and presupposed that each of them could be reversed to constitute a backwards path. Therefore, students who listed only forward paths were eliminated from analyses of the reading order bias (two students from Problem 1 and five from Problem 2).

Students tended to list paths in reading order. For example, in Problem 1 (see Figure 4) 90% of students listed YR as their first path. A paired-groups t test revealed that the mean rank of forward paths in subjects' path listing ($\bar{X}$=6.77, s = 2.11) was significantly lower than the rank of the corresponding backward path ($\bar{X}$=10.63, s = 3.30), t(65) = 10.96, p < .001, indicating that students indeed tended to list forward paths before backward ones.

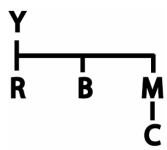


Figure 4. The topology of Problem 1.

Students also tended to omit more backwards paths. A dependent-groups t test revealed more backwards omissions than forward omissions for both Problem 1, $\bar{X}$ = 1.64, s = 2.24 for forward omissions versus $\bar{X}$ = 2.11, s = 2.71 for backwards omissions; t(65) = -2.98, p = .004, and for Problem 2 $\bar{X}$ = 1.60, s = 2.49, for forward omissions versus $\bar{X}$ = 2.02, s = 2.87 for backwards omissions, t(62) = -2.54, p = .014.

The *sequential bias* predicts that students should be influenced by the physical proximity of nodes in the diagram. One implication is that they are expected to make commission errors which introduce extraneous (intervening) nodes, for example listing YRBMC. Misunderstanding the concept of a bridge node leads to commission errors such as listing YC when YMC is correct because the bridge node M needs to be included. They could also combine these confusions, producing YRC. The first type of error, the introduction of an extraneous node, accounted for 93.7% of combined 298 commission errors in questions 1 and 2. Thus, the vast majority of commission errors are consistent with a sequential bias. The second type of error, the omission of the bridge node in a path crossing the bridge, accounted for only 2% of the errors, and omission of the bridge node combined with an extraneous node accounted for another 1.7% percent of the time. Thus, bridges were only omitted 3.7% of the time: for the most part, students did understand that information had to travel through the bridge node. There were other answers that fell into no obvious category – such as the inclusion of a node from the previous diagram, or single node paths – and these occurred 2.7% of the time. In principle, students could list nodes in any order. But the paths in general proceeded sequentially forwards or backwards. Of all the commission errors in questions 1 and 2, only 14, or 4.7% involved a change of direction, for example, BMY.

Analysis of associations among specific omission and commission errors provided additional evidence for this account of the conceptual gaps underlying errors in the paths task. Correlations were calculated among all 20 possible omission errors in Problem 1, and these correlations were then input to GTREE, an additive tree fitting program (Corter, 1998). The resulting tree is shown in Figure 5. We modified the additive tree graph by varying the thickness of the leaf arcs to represent the frequency with which the error occurred. The tree solution accounts for 80% of the variance in the correlations.

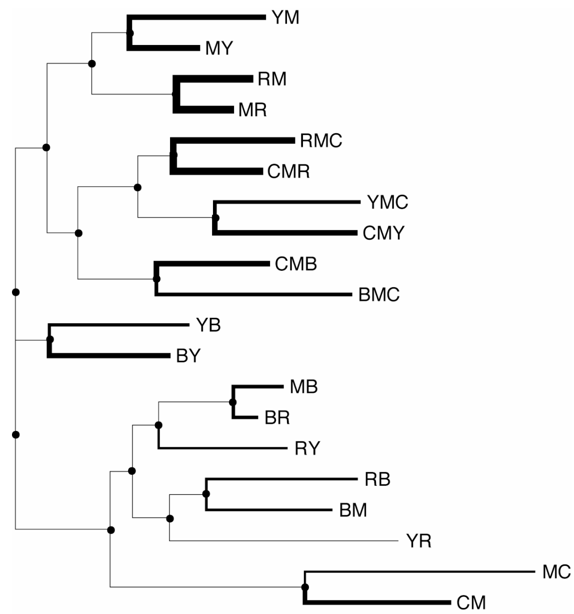


Figure 5. Additive tree of correlations among specific omission errors for Problem 1 (RSQ = .80, N=68).

Highly correlated errors tend to appear as close neighbors in the tree structure. For example, omission errors YM and MY are highly correlated, and they are neighbors in the tree. This frequent pattern indicates that when students fail to list a shortest path, they tend to omit its inverse as well.

The clusters at the top are made up of missed paths that are not between contiguous nodes, for example, path YM skips nodes R and B. This is true for the topmost eight errors (YM, MY, RM, MR, RMC, CMR, YMC, and CMY), which are also the most frequent omission errors. The bottom clusters consist of omissions of paths connecting contiguous nodes (e.g. YR). These types of errors are less common. To summarize, the clustering structure suggests that students make systematic omission errors, in part reflecting a tendency to miss the noncontiguous paths that is consistent with the hypothesized sequential bias.

A similar analysis was conducted for the specific commission errors made for Problem 1. The resulting tree is shown in Figure 6. It accounts for 59% of the variance in the correlations. Again, many branches consist of a path and its inverse (RBM and MBR). The larger clusters seem interpretable as well: the first cluster of 16 paths, including paths RBM through CMBYR, consists of non-minimal paths that are all consistent with a sequential bias. The next cluster, from path RYBM to CMRY, has a similar interpretation. The cluster comprised of paths RC, CY, YC, and BC groups paths that incorrectly omit the bridge node M. Three paths that are interspersed with this cluster are BMR, BMRY, and BMY. These three paths are three of the four commission errors that list nodes in an order that is inconsistent with a sequential (path-order) bias. Finally, the last cluster contains three paths that are BOTH non-minimal paths and omit the bridge node M (YBC, CBY, and YRC), and one additional path, RMBC, that is the remaining path that list nodes in a non-sequential order.

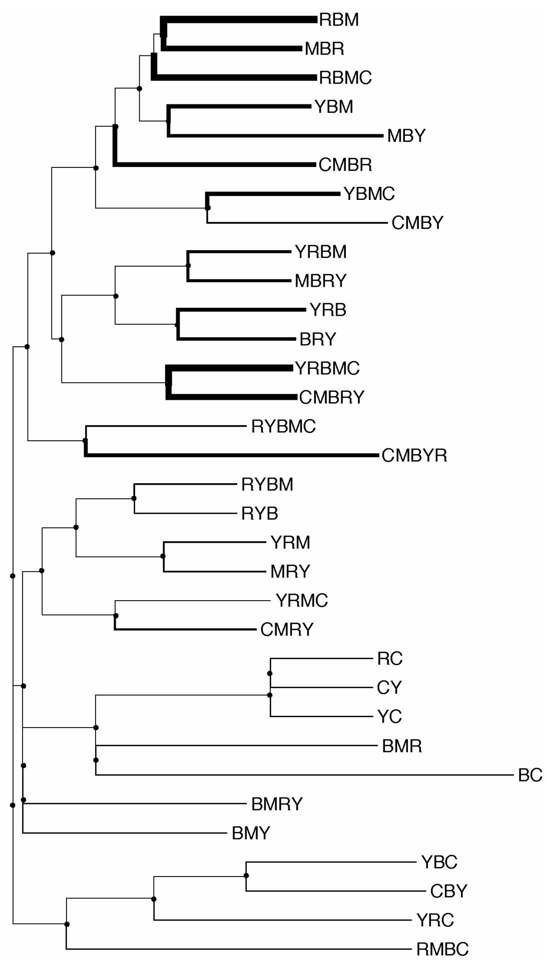


Figure 6. Additive tree of correlations among specific commission errors for Problem 1 (RSQ = .59, N=68).

In summary, the tree confirms that path errors are largely the result of a combination of two types of conceptual errors. The first type stems from believing that nodes on the path between the start and end nodes must be visited. The second type is a failure to recognize that all paths leading through a bridge node must visit it. Thus, the additive tree suggests the existence of underlying misconceptions that generate surface errors.

The diagrams generated for Problem 3 also provide data diagnostic of systematic errors. Because Problem 3 presented two opportunities for error, in translating the text to a diagram and in generating the shortest paths, there should be more errors on Problem 3 than Problems 1 and 2, even though all problems are identical in structure. This held for omission errors, with $\bar{X}$ = 5.45, s = 5.46) for Problem 3, more than for Problem 1 ($\bar{X}$ = 4.0, s = 5.03) and Problem 2 ($\bar{X}$ = 4.47, s = 5.94). This difference is significant, $F(1,66) = 7.58$, $p = .008$. However, there were not significantly more commission errors for Problem 3 ($\bar{X}$ = 2.51, s = 4.15) than for Problem 1 ($\bar{X}$ = 2.26, s = 4.09) and Problem 2 ($\bar{X}$ = 2.21, s = 4.21), $F(1,66) = 0.45$, $p = .504$.

The diagrams were also categorized for type and appropriateness according to the classes in Figures 1 and 2. Five students produced no diagram while attempting to answer Problem 3. Students who produced appropriate diagrams made fewer omission errors $\bar{X}$ = 2.84, s= 4.30, than students who produced inappropriate diagrams, $\bar{X}$ = 8.71, s = 4.56 or no diagram, $\bar{X}$ = 9.60, s = 7.13. This difference was significant in a between-subjects ANOVA, $F(1, 64) = 21.47$; $p < .001$. There was also no difference in number of omission errors between use of an inappropriate diagram and no diagram, $F(1,64) = 0.15$, $p = .696$.

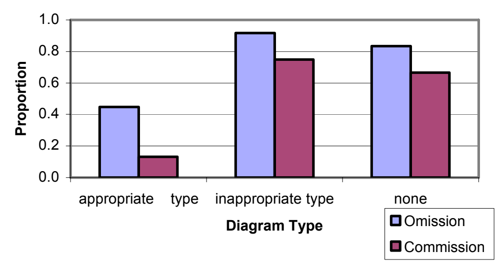


Figure 7. Proportion of students making omission or commission errors, by appropriateness of diagram type.

The same pattern held for commission errors. There were fewer errors for appropriate diagrams, $\bar{X}$ = .47, s = 1.55, than for inappropriate diagrams, $\bar{X}$ = 5.54, s = 5.19, or no diagram, $\bar{X}$ = 3.40, s = 3.28, $F(1, 64) = 15.67$; $p < .001$. The two ineffective strategies, using an inappropriate diagram type and using no diagram, did not differ in the number of commission errors, $F(1,64) = 1.61$, $p = .209$.

Discussion

Students in an introductory course in systems design were asked to solve three design problems. Two involved making inferences from a supplied diagram; a third entailed creating a diagram and making inferences from the created diagram. All problems required understanding network buses and network bridging – that is, understanding the topological structure of the problem and the conventions used to represent such structures. Students were expected to have difficulties interpreting and creating these kinds of diagrams. Specifically, they were expected to exhibit two biases, a sequential bias and a reading-order bias.

The sequential bias predicted the introduction of extraneous nodes in paths. This bias showed up strongly in our data set. The sequential bias also predicts that only commission errors reflecting nodes listed in their surface sequential order will occur: e.g., RBMC but not BRMC in Problem 1. Indeed, virtually all commission errors were sequential paths.

What might explain these types of errors? Because we live in a Euclidean world, we may tend to make inferences based on the proximity of objects. More specifically, we may read the diagram imagining that we are traversing the lines of the diagrams as if they were paths.

Because diagrams are too complex to be comprehended as wholes, they must be examined in sequence. The default sequence is reading-order. This order bias predicts more omission errors for inferences that don't correspond to reading order. This prediction, too, was borne out by the data. When students were asked to generate a diagram as well as generating the set of shortest paths, they used their diagram to generate the paths. That is, the type of diagram students generated predicted the errors they made.

The results indicate that diagrams are useful and actually used in making inferences in systems design. However, students have difficulties interpreting diagrams, especially when the diagrams portray a topological space that does not correspond to Euclidean intuitions. Can classroom instruction improve performance? The second study addresses this question.

Study 2: Posttest Diagram Generation

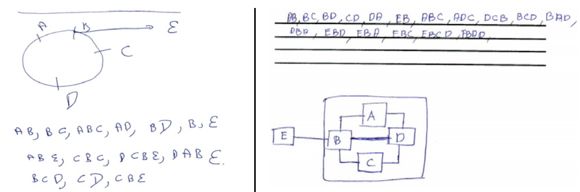
Method

Late in the course, 35 Master's level students from two of the four sections of the design course were asked to solve Problem 3 a second time (the "posttest"). The posttest was coded identically to the pretest version of Problem 3, and the results were compared. The expectation was improved diagrams and improved inferences as a consequence of the classroom instruction and exercises.

Results

Students did make fewer omission errors in the posttest, $\overline{X} = 3.85$, s = 5.02, than in the pretest, $\overline{X} = 4.48$, s = 4.90, indicating decreased reading order bias. However, this difference was not significant, t(32) = .73, p = .470.

Commission errors still occurred (Figure 8). However, students made fewer commission errors, indicating decreased sequential bias, in the posttest, $\overline{X} = .97$, s = 2.44, than in the pretest, $\overline{X} = 2.52$, s = 3.80, and this difference was significant, t(32) = 2.11, p = .042. Furthermore, of the 33 students who participated in both tests, 12 students (36%) gave a fully correct answer in the pretest, while 18 (55%) gave a fully correct answer in the posttest. This improvement in performance demonstrates increased understanding of network topology.



a) A ring representation with commission errors

b) An incomplete graph with commission errors

Figure 8. Examples of errors on posttest diagrams

Table 2 shows the frequency of producing bus, other, and no diagrams on the pretest and posttest. Those who used a bus on the pretest either used it again on the posttest or used no diagram on the posttest. Impressively, 13 of the 18 students who failed to use a bus on the pretest used a bus at posttest. This increase in use of the bus is marginally significant, $\chi^2(1) = 3.02$; p = 0.082.

TABLE 2. Changes in use of Diagram Types for Problem 3

		Posttest			
		bus	inappropriate	none	Total
Pretest	bus	11	0	4	15
	other appropriate	4	1	1	6
	inappropriate	9	3	0	12
	none	0	0	0	0
	Total	24	4	5	N=33

Do students who produce better diagrams also produce better solutions? Table 3 shows the effectiveness of the diagram types for promoting correct inferences. Fourteen out of the 26 students who drew the conventional bus diagram got the problem correct, while only one out of four students who tried to use another type of diagram did so. A chi-square test of independence for type of diagram used (bus versus other graph versus no graph) and solution correctness was marginally significant, $\chi^2(2)=5.55$; p = 0.062, suggesting that learning to choose the right diagram convention is key to solving the problem. A chi-square test of independence for an association between the use (or no use) of a diagram and solution correctness on the posttest was significant, $\chi^2(1) = 4.375$, p = 0.036.

TABLE 3. Posttest Problem 3: Use of appropriate (bus) and inappropriate diagram types, with solution correctness.

Diagram type	Correct answer	Incorrect answer	Proportion Correct	N
Bus	14	12	.54	26
inappropriate	1	3	.25	4
None	5	0	1.00	5
(Total)	20	15	.57	35

Discussion

Students in a systems design class (Nickerson, 2006) were asked to generate diagrams to solve design problems that included a logical bus topology both early in the semester and late in the semester. In the posttest, more students produced diagrams that represented a bus and were able to correctly produce all shortest paths. Students who produced appropriate diagrams also produced better solutions

By the end of the semester, commission errors decreased, suggesting that most students did master topological concepts. Omission errors remained fairly constant, suggesting that they need to be counteracted with a different form of instruction.

Conclusions

Sketches and diagrams are an essential component of design of information systems. Working with complex systems may overload limited capacity working memory, a problem solved by externalizing the structure (and perhaps function) of a system by committing it to paper. An external representation serves as a basis for inferences, a basis for generating new designs, and a visible product that can be shared with others. Diagrams and sketches facilitate inferences by capitalizing on their physical features, such as proximity, angle, and connectivity, and because they can capture complex relations among parts and wholes. They foster creativity by enabling alternatives, expansions, reductions, revisions.

But all these virtues depend on successful reading and interpretation of diagrams and sketches, skills that improve with expertise. Reading and interpreting diagrams are affected by habits and biases from reading and interpreting the visual world and other common external representations, such as maps. These habitual ways of interacting can lead to failures and to errors in using diagrams.

Here, we studied sketches produced by students of systems design in the service of problem solving, making inferences from the sketches. The systems portrayed contain a logical bus, a set of links that are mutually connected. The graphic convention for the bus shows components attached to a line. This convention causes difficulties for students. In the present experiments, students exhibited the difficulties by generating paths that include unneeded nodes – errors of commission. Because these extra nodes are virtually always listed in the order that they appear along the path connecting the endpoints, we call this bias a *sequential* bias. The second bias exhibited by students was a *reading order* bias. Students generated paths in the canonical reading order for European languages, from upper left to lower right. They often failed to generate all the correct paths (errors of omission), and there were more backwards omissions than forward omissions. These biases may be quite widespread, because similar errors can be found in reading maps (Taylor and Tversky 1992), and in constructing diagrams of cycles (Kessell and Tversky 2007).

Knowledge of these biases and the specific problems students have in constructing and using diagrams could be used to improve instruction in systems design. For example, teaching several alternative appropriate diagram types for a given problem (e.g., bus, hub-and-spoke, and complete graphs) might protect against errors of inference that tend to be made with one type of diagram. Specific practice with a number of different inference tasks might also deepen understanding of diagram conventions, and improve student understanding of the flow of information in networks. Although the interpretation of diagrams frequently seems obvious to experts — after all, the information is *there* — it is increasingly apparent that producing and interpreting diagrams are skills that need to be taught.

Acknowledgements

The authors gratefully acknowledge the support of the National Science Foundation (awards IIS-0725223 and REC-0440103) and the Stanford Regional Visualization and Analysis Center.

References

Corter, J. E. (1998). An efficient metric combinatorial algorithm for fitting additive trees. *Multivariate Behavioral Research, 33*, 249–272.

Kessell, A. & Tversky, B. (2007). To go in circles or forge straight ahead: Depicting cyclical processes. *Poster presented at the 48th meeting of the Psychonomic Society*.

Landy, D. and Goldstone, R. (2007). The alignment of order and space in arithmetic computation. *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society,* 382-387. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nickerson, J. V. (2005). The meaning of arrows: Diagrams and other facets in system sciences literature. *Proceedings of the 38th Hawaii International Conference on System Sciences*.

Nickerson, J.V. (2006). Teaching the integration of information systems technologies. *IEEE Transactions on Education,* 49, 271-277.

Nickerson, J. V., Corter, J. E., Tversky, B., Zahner, D., and Rho, Y.J. (in press). Diagrams as a tool in the design of information systems. In J. S. Gero and A. Goel (Eds.), *Design Computing and Cognition '08*, Dordrecht, Netherlands: Springer.

Taylor, H. A. & Tversky, B. (1992). Descriptions and depictions of environments, *Memory and Cognition*, *20*, 483–496.

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 & V. Haarslev (Eds.), *Theory and application of diagrams*. Berlin: Springer.