To appear in *Proceedings of Diagrams 2004: Third Int'l Conf. on the Theory & Application of Diagrams*, A volume in the Lecture Notes in Computer Science Series,

Predictors of success in diagrammatic problem solving

Daesub Yoon & N. Hari Narayanan

Intelligent and Interactive Systems Laboratory, Department of Computer Science
& Software Engineering, Auburn University, Auburn, AL 36849, USA
{yoondae, narayan}@eng.auburn.edu

Abstract. We conducted an eye-tracking study of mechanical problem solving from cross-sectional diagrams of devices. Response time, accuracy and eye movement data were collected and analyzed for 72 problem-solving episodes (9 subjects solving 8 problems each). Results indicate that longer response times and visually attending to more components of a device do not necessarily lead to increased accuracy. However, more focus shifts, visually attending to components in the order of causal propagation, and longer durations of visual attention allocated to critical components of the devices appear to be characteristics that separate successful problem solvers from unsuccessful ones. These findings throw light on effective diagrammatic reasoning strategies, provide empirical support to a cognitive model of comprehension, and suggest ideas for the design of information displays that support causal reasoning.

1 Introduction

How people reason about and solve graphically presented problems drawn from visual, spatial and causal domains has been a topic of interest in diagrammatic reasoning research. This interest stems from the fact that certain characteristics of systems in such domains lend themselves to graphical representations in which the visual properties and spatial distribution of components aid the problem solver in directing his or her reasoning along paths of causal propagation. These characteristics are: (1) components of a system are spatially distributed; (2) systems are dynamic, i.e. components and their properties change over time; (3) system components causally interact with each other; (4) such interactions can be traced along chains of cause-effect relationships (which we call lines of action) that branch and merge in spatial and temporal dimensions; and (5) predicting the operation of a system requires reasoning from a given set of initial conditions to infer these causal chains of events. Reasoning about mechanical devices from cross-sectional diagrams is a case in point. Other examples of problems from visual, spatial and causal domains are circuit design, weather forecasting, emergency response coordination and military course-of-action planning.

Understanding the cognitive processes underlying such reasoning, especially strategies that separate successful problem solvers from unsuccessful ones, can pro-

vide insights into the design of information displays that actively aid the problem solver and enhance his or her performance. In the context of a research program on designing and evaluating such information displays (see Section 4 for details), we report on an experiment that investigated how people make predictions about the operation of a mechanical device, when given a labeled cross-sectional diagram of the device and an initial condition – specified as the behavior of one of the device's components. In addition to determining the accuracy of their answers, we measured their response times and collected data on their eye movements across the stimulus display. Our goal was to understand the relations among accuracy, response time and patterns of visual attention allocation across the display.

The rest of this paper is organized as follows. In Section 2 we summarize earlier work on mechanical reasoning from diagrams that has a bearing on the present research. Section 3 describes the experiment and its results. In the final section, conclusions that can be drawn from the experiment's results, their implications and our future research are discussed.

2 Related Research

Larkin and Simon (1987) undertook a computational analysis of diagrammatic versus sentential representations, and described features of diagrams that aid reasoning. Extending this line of enquiry, Cheng (1996) proposed twelve functional roles of diagrams in problem solving: (1) showing spatial structure and organization, (2) capturing physical relations, (3) showing physical assembly, (4) defining and distinguishing variables, terms and components, (5) displaying values, (6) depicting states, (7) depicting state spaces, (8) encoding temporal sequences and processes, (9) abstracting process flow and control, (10) capturing laws, (11) doing computations, and (12) computation sequencing. These analyses suggest that diagrams can aid a problem solver by explicating and facilitating inferences about the components, structure, states and spatio-temporal sequences of causally connected events of systems in visual, causal and spatial domains.

Other research has delved into details of diagrammatic reasoning. Hegarty (1992) provides an account, based on reaction time and eye-fixation data, of how people infer the motions of mechanical devices from diagrams. She found evidence for an incremental reasoning process: the device is decomposed into its components and their behaviors are mentally animated in the direction of causality. However, mental animation is constrained by working memory capacity, such that people are only able to mentally animate one or two component motions at a given time. Furthermore, the eye-fixation data indicated that mental animation is accompanied by inspection of the relevant parts of the diagram.

In earlier research Narayanan, Suwa and Motoda (1994) investigated how people solved mechanical reasoning problems presented as diagrams. Analysis of subjects' verbal and gestural protocols supported the incremental reasoning model. It also suggested that shifts of the problem solver's focus from component to component were mediated to a great extent by the connectivity of components, internal visuali-

zation of component behaviors, propagation of causality and search for information in the diagram.

They further analyzed the intermediate hypotheses (extracted from verbal protocols) of subjects who were reasoning about an "impossible" mechanical device (see problem 6, Figure 1, later in the paper). In this device's operation, the causal chains of events branch and merge in the spatial dimension (within the device's physical structure) and the temporal dimension (events occur concurrently as well as sequentially). This analysis revealed that the trajectory of reasoning was mediated by the lines of action (Narayanan, Suwa & Motoda, 1995). However, instead of following a systematic strategy such as depth-first (traverse each branch fully before starting on another) or breadth-first (traverse all branches in an alternating fashion), subjects took a mixed approach with elements of both. They also retraced their reasoning paths multiple times, especially near the merging points of the lines of action.

3 Experiment

Prior research has thus identified several characteristics of mechanical reasoning from diagrams. Three important ones are decomposing the device into its components and attending to each individually, reasoning along the lines of action, and focus shifts mediated by several factors. In the experiment reported here we investigated whether successful and unsuccessful problem solvers could be separated in terms of these characteristics.

Response time, accuracy and eye movement data were collected and analyzed for 72 problem solving episodes: 9 subjects solving 8 problems of mechanical reasoning from cross-sectional and labeled diagrams. Components attended to, the direction of reasoning and focus shifts were determined from eye movement data.

Our interpretation of eye movement data is based on the premise that fixations are likely indicators of what the problem solver is (or has been) thinking about. The assumption (called the eye-mind assumption) is that the locus of eye fixations corresponds to the information being processed by the cognitive system. The eye-mind assumption is supported by two independent lines of research. Just and Carpenter (1976) discuss extensive evidence supporting this assumption for goal-directed tasks that require information to be encoded and processed from the visual environment. The problem solving tasks in our experiment were goal-directed ("given the initial motion of one component, predict the motion of another component further downstream"). Information needed to carry out these tasks had to be encoded and processed from a visual stimulus displayed on a computer monitor. Additional evidence that eye movement traces carry information about cognitive processes that underlie mechanical reasoning from device diagrams appears in (Rozenblit, Spivey & Wojslawowicz, 1998). These researchers collected eye movement traces of subjects making predictions about mechanical devices presented as diagrams and gave the traces to independent raters. The traces alone were sufficient for the raters (who did not see the device diagrams) to reliably predict both the principal axes of orientation of the devices subjects saw, and whether the subjects solved each problem correctly.

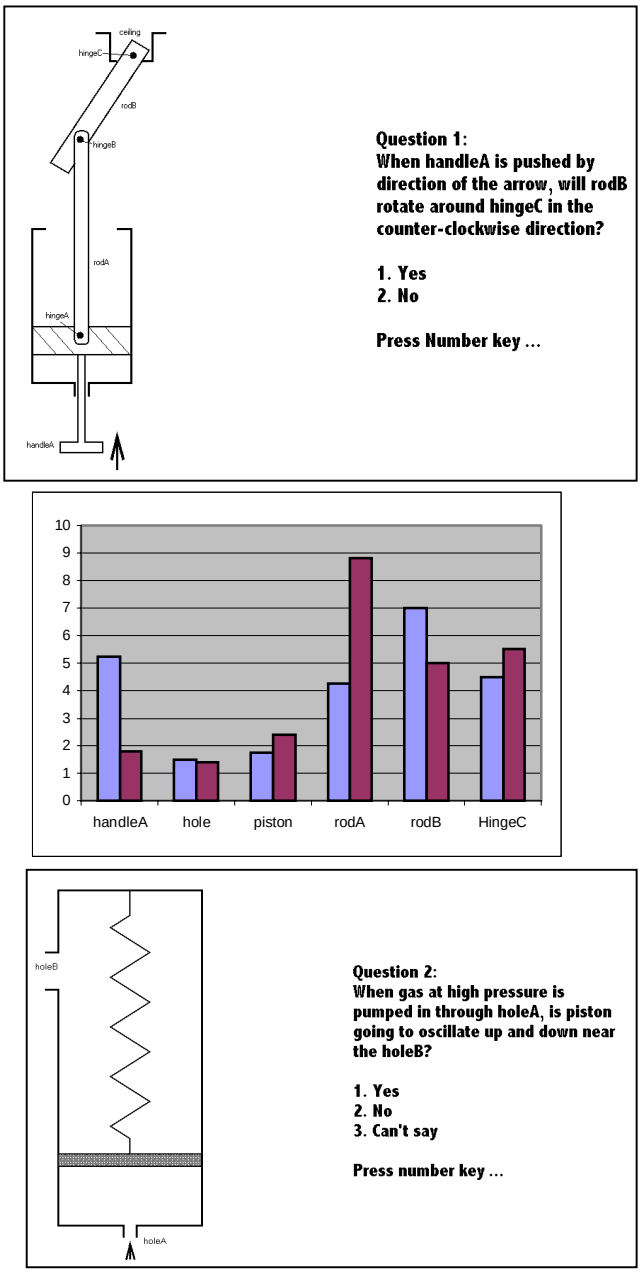
3.1 Procedure

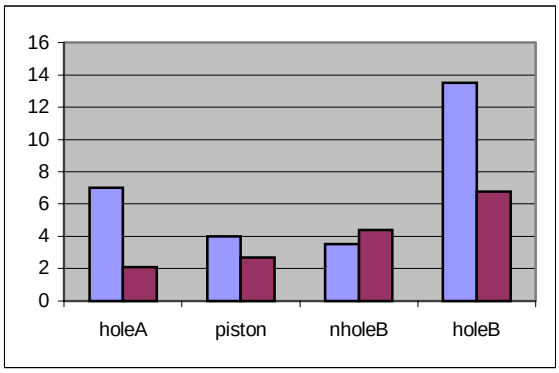
Nine engineering graduate students volunteered to participate. They were compensated with a payment of $10 each. Each subject solved 8 problems. Each problem was displayed as a labeled cross-sectional diagram of a device with an accompanying question and possible answers (Figure 1 shows the 8 stimuli exactly as displayed to subjects). The first three problems involve simple mechanical devices. The fourth is a Rube Goldberg-like device for frying an egg. These four problems have been used in prior research (Narayanan, Suwa & Motoda, 1994). The fifth is a pulley system previously used by Hegarty in her experiments (1992). The sixth is an "impossible" problem, also used in prior research (Narayanan, Suwa & Motoda, 1995). It involves branching and merging causal event chains, unlike the previous five problems. The seventh and eighth problems are about the flushing cistern, the most complex of all seven devices used in this experiment. The operation of this device also involves branching and merging causal event chains. Furthermore, in previous studies of comprehension of this device from interactive graphical presentations, it was found that while subjects were able to infer behaviors of components within each causal chain, they had difficulty integrating information between the two causal chains (Hegarty, Quilici, Narayanan, Holmquist & Moreno, 1999).

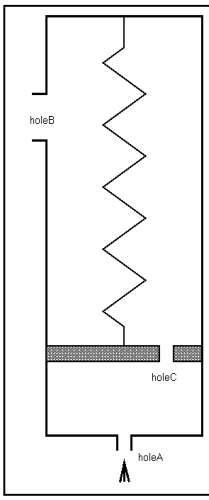
The experiment was conducted one subject at a time in an eye tracking laboratory equipped with a head-mounted eye tracker, eye tracking computer and a stimulus display computer. The eye tracker we used is the Eye Link model from SMI Inc. It consists of a headband, to which two infrared sources and cameras (one for each eye) are attached. It is a video-based eye tracker that detects pupil and corneal reflections from infrared illumination to compute screen coordinates of the subject's gaze point on the stimulus display monitor once every 4 milliseconds. The headband is attached by cable to a PC which functions as an experiment control station as well as carries out the necessary computations. This PC communicates with the stimulus display computer via an Ethernet link. The problems were presented as static pictures on the monitor of the stimulus display computer. Subjects sat in a high-backed chair, and viewed the problem on a 20-inch monitor mounted on a wall at eye level, at a distance of approximately 3 feet. The experimenter sat behind the subject and controlled the experiment through the eye-tracking computer.

First, a subject solved three practice problems with piston, pulley and gear systems. The devices in the practice problems were much simpler than the devices in the experimental problems. A 10-minute break and calibration of the eye tracker followed. The actual experiment began by the subject clicking the left mouse button to display the first stimulus. When ready to make the prediction, the subject depressed a number key on the keyboard to indicate his or her answer. This automatically brought up the next stimulus. The key presses were used to compute and record response times and determine accuracy. The specific problem to be solved, in the form of a question, appeared as part of each stimulus. This text also specified the number keys corresponding to possible answers. Each problem had two or three possible answers, of which only one was correct. Eye movement data was collected and recorded for all problems.

Fig. 1. Eight experimental problems. The bar chart below each problem shows the mean duration (as a percentage of total time on task) of visual attention on each component of the device in that problem by successful (light bars) and unsuccessful (dark bars) subjects.



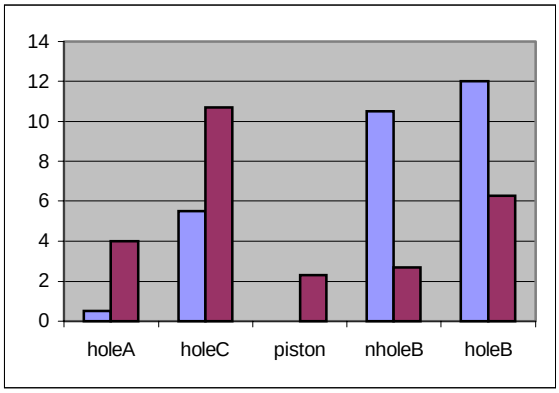


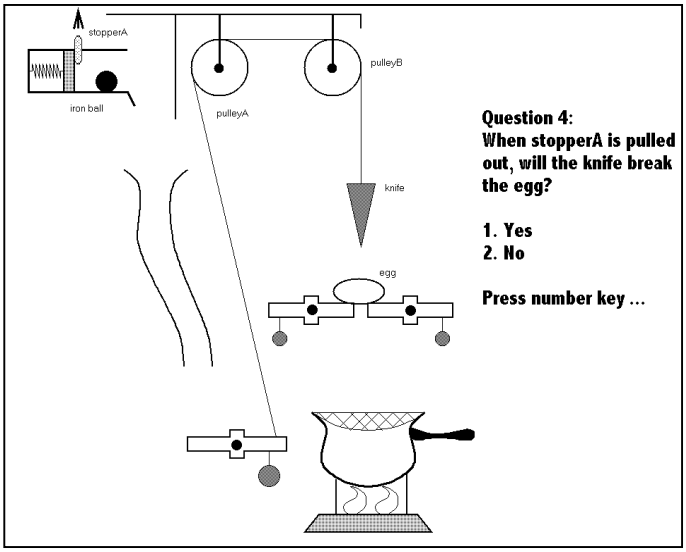


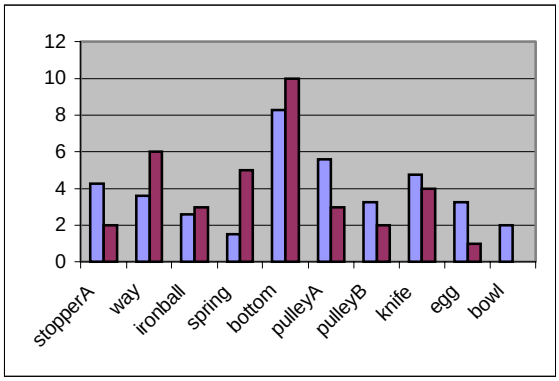
Question 3:
When gas at high pressure is pumped in through holeA, is piston going to oacillate up and down near the holeB?

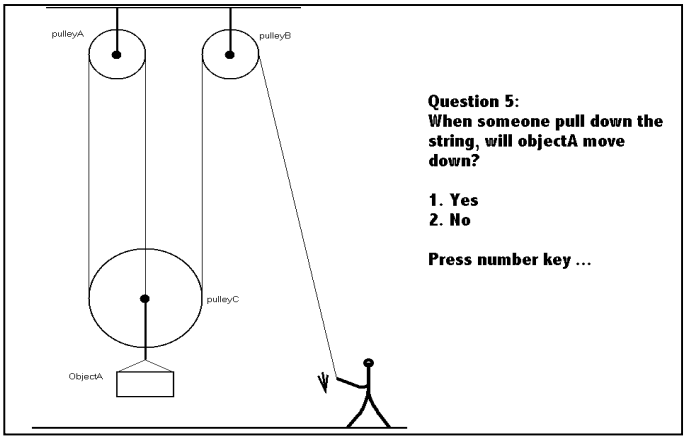
1. Yes
2. No
3. Can't say

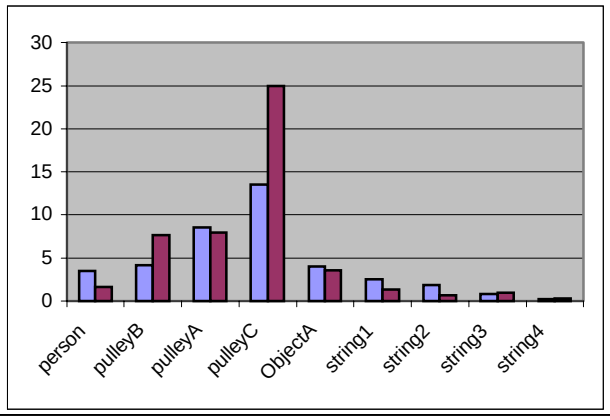
Press number key ...

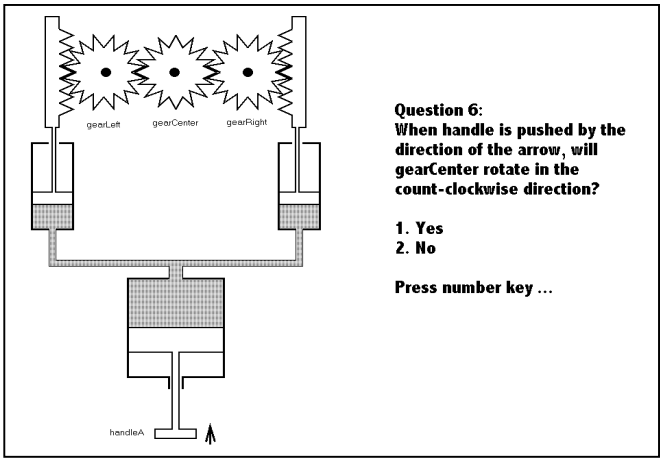


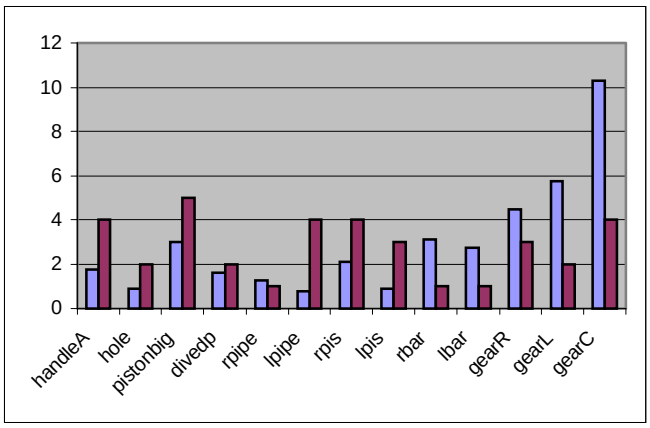

stopperA
iron ball
pulleyA
pulleyB
knife
egg
Question 4:
When stopperA is pulled out, will the knife break the egg?
1. Yes
2. No
Press number key ...

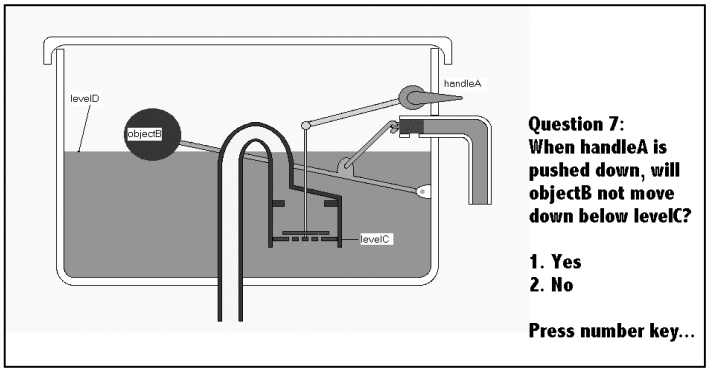

12
10
8
6
4
2
0
stopperA
way
ironball
spring
bottom
pulleyA
pulleyB
knife
egg
bowl

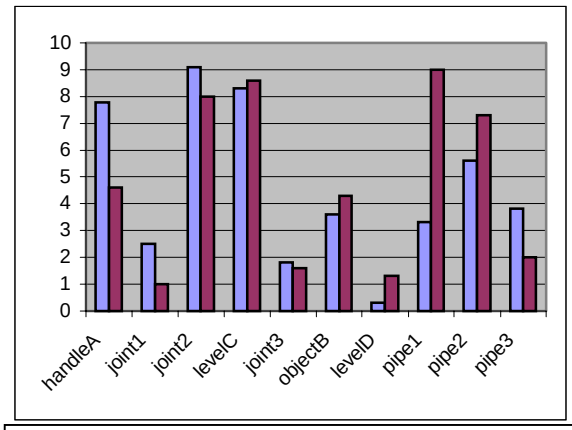

pulleyA
pulleyB
pulleyC
ObjectA
Question 5:
When someone pull down the string, will objectA move down?
1. Yes
2. No
Press number key ...


30
25
20
15
10
5
0
person
pulleyB
pulleyA
pulleyC
ObjectA
string1
string2
string3
string4

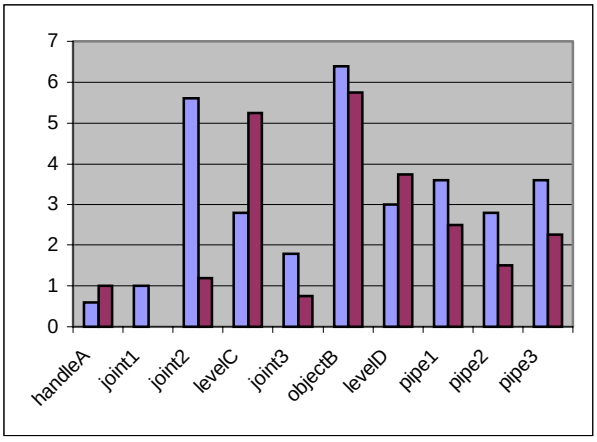

gearLeft
gearCenter
gearRight
handleA
Question 6:
When handle is pushed by the direction of the arrow, will gearCenter rotate in the count-clockwise direction?
1. Yes
2. No
Press number key ...


12
10
8
6
4
2
0
handleA
hole
pistonbig
divedp
rpipe
lpipe
rpis
lpis
rbar
lbar
gearR
gearL
gearC





The eighth problem showed the same device as in Problem 7, a flushing cistern, with the question: "After answering Question 7, will the water level rise until objectB reaches levelD?"



3.2 Results

We computed five dependent measures: *accuracy*, *response time*, *coverage*, *number of focus shifts* and (causal) *order*. Accuracy is a nominal variable categorizing a subject's answer to a problem as correct or incorrect. The correct answers to the 8 problems are: Question 1 – No; Question 2 – Yes; Question 3 – Can't say; Question 4 – Yes; Question 5 – No; Question 6 – No; Question 7 – No; Question 8 – Yes. The accuracies of the nine subjects were 62.5% (the first subject was correct in 5 out of the 8 problems), 50% (4/8), 50% (4/8), 62.5% (5/8), 12.5% (1/8), 87.5% (7/8), 50% (4/8), 62.5% (5/8) and 75% (6/8) respectively. The accuracies for each of the 8 problems were 44.4% (4 subjects out of 9 were correct), 22.2% (2/9), 22.2% (2/9), 88.9% (8/9), 66.7% (6/9), 88.9% (8/9), 66.7% (6/9), and 55.6% (5/9) respectively. Response times were automatically recorded by the stimulus display computer using subjects' key presses.

In each device diagram, individual components were delineated by bounding boxes. Once the co-ordinates of these bounding boxes were determined, we could associate fixations with components, and calculate the duration of each subject's gaze on each component as a percentage of the total time the subject spent on that problem. Coverage for a problem and a subject was computed as the percentage of components of the corresponding device that attracted at least one fixation by the subject.

Using the bounding box technique, we could also determine a subject's shifts of visual focus from component to component in each problem from raw eye movement data. We derived *S*, an ordered sequence of device components that a subject attended to during a session, for each subject and each problem. This was done by aggregating consecutive fixations inside the bounding box of a component and detecting when another component was fixated upon. *S* begins with the first device component attended to, and ends with the last component attended to before the subject pressed a number key indicating his or her solution. Number of focus shifts by a subject for a problem is then the number of transitions in S, i.e., (the size of *S*) – 1. Fixations on the question and blank display regions were ignored in this analysis.

In the sequence *S*, if component *j* appears immediately after component *i*, and if *i* can causally influence *j* according to the lines of action in the device, then *i-j* represents a causal link in the sequence *S*. Consecutive causal links represent causal subsequences of *S*. The length of a causal subsequence is the number of causal links in it. The measure order is defined as the sum of squares of the lengths of causal subsequences in *S*. This captures the total number of correct cause-effect pairs of components that a subject considered, weighted by the length of unbroken lines of action that the subject considered (i.e. if subjects A and B both considered the same number of causal links, but if A looked at longer causal link chains than B, the value of order would be higher for A than B).

Analyses were conducted to compare response times, coverage, focus shifts and order of successful problem solvers with those of unsuccessful problem solvers. The

goal was to discover whether successful problem solvers could be characterized by longer response times, higher coverage, more focus shifts or larger values of order.

Statistical analyses (t-tests) were carried out to compare the mean values of response time, coverage, number of focus shifts and order. Two groups of problem solving episodes from the total of 72 (9 subjects X 8 problems) were compared: a group of 41 in which the subjects provided the correct answer and another group of 31 in which the subjects were wrong. We found no significant difference between the mean response times of successful and unsuccessful subjects across all 8 problems, $t(71) = 1.138$, $p < 0.26$. The top-left bar chart in Figure 2 shows the mean response times of successful and unsuccessful subjects for each problem. For problems 1, 2, 4 and 8 subjects who provided the correct answer had lower mean response times than subjects who were wrong. For problems 3, 5, 6 and 7 this reversed.

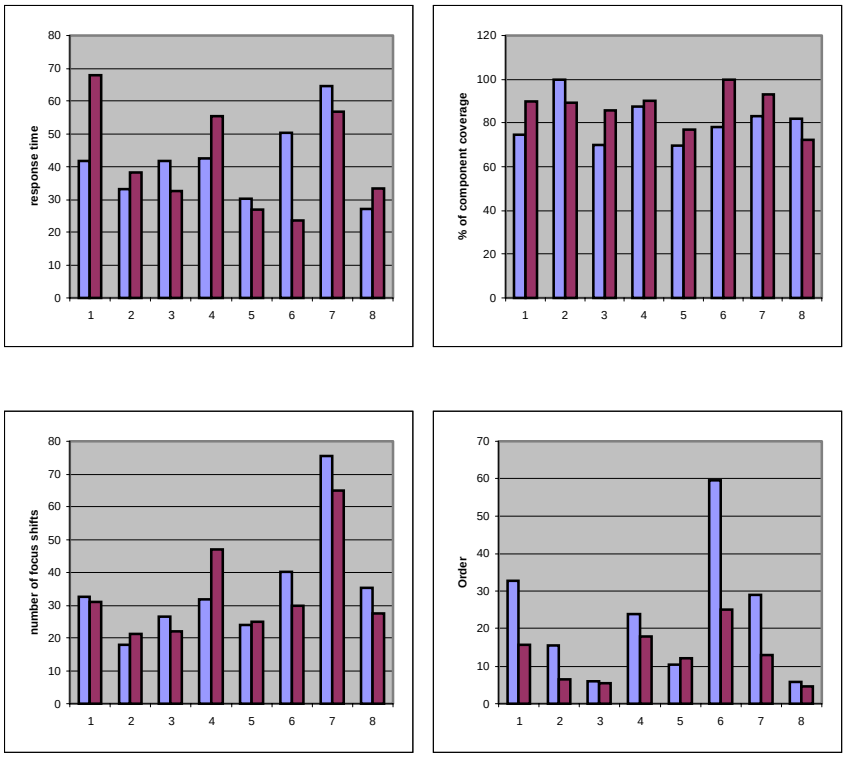


Fig. 2. These bar charts show, for each problem (on the x-axis), the means (on the y-axis) of response time (seconds), coverage, number of focus shifts and order for successful (light bars) and unsuccessful (dark bars) subjects.

There was no significant difference between the mean percentage of components that successful and unsuccessful subjects attended to across all 8 problems, $t(71) = -1.414$, $p < 0.16$. The top-right bar chart in Figure 2 shows the mean coverage of successful and unsuccessful subjects for each problem. For problems 1, 3, 4, 5, 6 and 7

subjects who provided the correct answer had lower mean component coverage than subjects who were wrong. For problems 2 and 8 successful subjects exhibited a higher mean coverage than unsuccessful ones.

We found marginal significance in the difference between the means of the number of focus shifts of successful and unsuccessful subjects across all 8 problems, $t(71) = 1.792$, $p < 0.08$. The bottom-left bar chart in Figure 2 shows the mean number of focus shifts of successful and unsuccessful subjects for each problem. For problems 2, 4 and 5 subjects who provided the correct answer exhibited lower number of focus shifts on average than subjects who were wrong. For problems 1, 3, 6, 7 and 8 successful subjects made more focus shifts on average than unsuccessful ones.

However, t-test comparisons of the mean value of order indicated that subjects who were accurate considered significantly more causal connections and longer lines of action than subjects who provided wrong answers did, across all 8 problems, $t(71) = 2.934$, $p < 0.0045$. The bottom-right bar chart in Figure 2 shows the mean values of order for successful and unsuccessful subjects for each problem. For all problems except the fifth, successful subjects had higher mean values of order than subjects who were inaccurate.

Next, we investigated the relation between gazing on particular components and accuracy. Even though no significant differences between successful and unsuccessful problem solvers were discovered in terms of their response times and the percentage of components of each device they visually attended to during problem solving, we explored this further by calculating the gaze durations of each subject on each component of each device as a percentage of that subject's response time for that device. From this data we calculated the mean gaze duration percentages of successful and unsuccessful problem solvers for each component of the eight problems. These are shown as bar charts appearing under each of the 8 problems in Figure 1. Note that gaze duration percentages are computed and shown in the bar charts for all major components of a device, even if only some of these components are labeled in the device's diagram that subjects saw. For example, the label "nholeB" in bar charts corresponding to problems 2 and 3 refers to the spring and surrounding area inside the cylinder below holeB in the corresponding two devices, though this label does not appear in the device diagrams.

The bar charts in Figure 1 illustrate one consistent pattern across all problems. Successful problem solvers are differentiated from unsuccessful ones by the fact that they spent more time, as a percentage of total time to solve the problem, on average on components that are critical to solving the problem. For problem 1, rodB is the important component. For problems 2 and 3, it is holeB. For problem 4, the last components in the line of action – pulleyB, knife and egg – are the critical ones. For problem 5 the various string segments are the critical components. For problem 6, its "impossibility" becomes evident when one considers the combination of three circular gears: gearRight, gearLeft and gearCenter. For problem 7 the critical components are in the region of the small arm of the siphon pipe (components joint1 and joint2). For correctly answering the eighth problem's question, the eventual closing of the inlet valve (component pipe3) is a critical inference. As can be seen in Figure 1, successful problem solvers spent a higher percentage of time fixating on these

components. Just and Carpenter (1976) suggest that gaze duration provides a measure of the time spent processing the corresponding symbol. Therefore it is reasonable to conclude that a longer duration (relative to time on task) of visual attention allocated to critical components of the devices is a characteristic that separates successful problem solvers from unsuccessful ones.

4 Discussion: Implications and Future Research

This paper described an experiment on diagrammatic reasoning about mechanical devices in which we investigated characteristics that differentiate successful and unsuccessful problem solvers. One outcome measure (accuracy) and five process measures (response time, coverage, number of focus shifts, causal order of processing and relative gaze durations on individual components), four of which were derived from eye movements, were analyzed to examine whether successful problem solvers could be characterized in these terms.

The results support several conclusions regarding effective diagrammatic reasoning strategies. Spending more time on the task and visually attending to more components do not necessarily lead to success in mechanical reasoning from device diagrams. On the other hand, considering more component pairs that are causally related and attending to longer causal chains of components can lead to better accuracy. Concentrating on critical components for relatively longer durations also appears to improve accuracy in problem solving. Increased shifting of one's focus of visual attention from component to component during problem solving is marginally positively related to accuracy, but clearly which components are attended to and in which order are more significant predictors.

These findings suggest that in training novices to qualitatively and successfully reason about mechanical systems from diagrams, it is important for instruction to focus on developing skills of determining and following causal chains of events in the operation of the device, and identifying components that are critical to solving the problem at hand. An important characteristic of the mechanical domain is that it consists of dynamic systems with spatially distributed components that causally interact and give rise to event chains that branch and merge in spatial and temporal dimensions. Several other domains share this characteristic. Therefore we postulate that the pedagogical implications of our findings extend to these domains, such as meteorology, as well.

Another implication of this research is that it has provided new empirical evidence supporting a previously reported cognitive process model of causal system comprehension from text and diagrams. Narayanan and Hegarty describe this cognitive model of multimodal comprehension (Narayanan & Hegarty, 1998), based on which they argue that an information display with the following six characteristics is likely to enhance comprehension (Narayanan & Hegarty, 2002). It should aid the viewer in decomposing the system being described, enable the viewer to invoke relevant prior knowledge, point out the common referents of external representations in different modalities, explain domain laws that govern the system, explicate

the lines of action in the operation of the system, and encourage mental animation. While they discuss empirical support for some of these characteristics, the efficacy of a display that supports reasoning along the lines of action has not been experimentally investigated. Our finding that successful problem solvers exhibit significantly higher values of the measure order, indicating consideration of more causal links along longer lines of action, does indeed suggest that a display that facilitates reasoning along the lines of action can enhance comprehension.

Understanding the cognitive processes underlying causal reasoning from visual displays, especially strategies that separate successful problem solvers from unsuccessful ones, can provide insights into the design of displays that actively aid the problem solver and improve his or her problem solving performance. For instance, Grant and Spivey (2002) studied Karl Duncker's radiation problem, by first showing a diagram of the problem to subjects and determining the part of the diagram that received most attention from successful problem solvers. In a subsequent experiment, they attracted subjects' attention to that part through a blinking action, and found that merely attracting the problem solver's attention to that part of the display dramatically improved accuracy. But can the display regions that are critical, and the optimal order of visual attention allocation, for a problem be determined a priori?

Our research suggests that, for problems drawn from domains with the five characteristics listed in the introduction of this paper, lines of action in the operation of the system provide spatial pathways of optimal visual attention allocation. Furthermore, components at the merge points in these pathways and components most closely associated with the problem being solved are critical, and allocating more attention to these can improve accuracy. This leads naturally to a research program on displays that track and guide a problem solver's visual attention, and provide relevant information at the right time and in the right place, to support causal reasoning. We term displays that thus exploit the trajectory of a viewer's focus shifts during problem solving "Reactive Information Displays" (Narayanan & Yoon, 2003). To be effective, such displays must have knowledge about the system/domain that is being displayed, knowledge about the problem solving task that the user is engaged in, knowledge regarding an applicable problem solving model and knowledge about the trajectory of the user's attention shifts.

An introduction to Reactive Information Displays and an empirical study of four reactive strategies are presented in (Narayanan & Yoon, 2003). This study showed that a display that guides the user's visual attention along paths of causal propagation while demonstrating potential behaviors of individual components significantly improved the accuracy of mechanical problem solving. In another experiment (Yoon and Narayanan, 2004), we discovered that when subjects were first shown a mechanical reasoning problem and then given a second problem on the same device, but without any diagram, about half had fixations on the blank region of the display that the device diagram occupied in the first problem. While these subjects were no more accurate than those who did not exhibit any eye movements on the blank part of the display, they looked at more "virtual" components and had significantly higher values of the measure order, indicating that their eye movements on the blank

region were systematic, along the lines of action (i.e. examining causally related chains of components). This suggests that Reactive Information Displays may be particularly useful for this kind of users, whose eye movements reflect cognitive processes even in the absence of an external stimulus. In current work, we are evaluating several reactive display strategies in the domain of algorithmic problem solving. Pursuing these lines of enquiry will remain the focus of our future research.

Acknowledgment This research is supported by the Office of Naval Research under grant N00014-03-10324 to Auburn University.

References

1. Cheng, P. C-H. (1996). Functional roles for the cognitive analysis of diagrams in problem solving. *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 207-212). Hillsdale, NJ: Lawrence Erlbaum Associates.
2. Grant, E. R., & Spivey, M. J. (2002). Guiding attention produces inferences in diagram-based problem solving. In M. Hegarty, B. Meyer & N. H. Narayanan (Eds.), *Multidisciplinary Studies of Diagrammatic Representation and Inference*, LNAI 2317, Berlin: Springer.
3. Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *JEP: Learning, Memory and Cognition*, 18(5), 1084-1102.
4. Hegarty, M., Quilici, J., Narayanan, N. H., Holmquist, S., & Moreno, R. (1999). Multimedia instruction: Lessons from evaluation of a theory based design. *Journal of Educational Multimedia and Hypermedia*, 8, 119-150.
5. Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441-480.
6. Larkin, J., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-100.
7. Narayanan, N. H., & Hegarty, M. (1998). On designing comprehensible interactive hypermedia manuals. *International Journal of Human-Computer Studies*, 48, 267-301.
8. Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies*, 57, 279-315.
9. Narayanan, N. H., Suwa, M., & Motoda, H. (1994). A study of diagrammatic reasoning from verbal and gestural data. *Proceedings of the 16th Annual Conference of the Cognitive Science Society* (pp. 652-657). Hillsdale, NJ: Lawrence Erlbaum Associates.
10. Narayanan, N. H., Suwa, M., & Motoda, H. (1995). Diagram-based problem solving: The case of an impossible problem. *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 206-211). Hillsdale, NJ: Lawrence Erlbaum Associates.
11. Narayanan, N. H. & Yoon, D. (2003). Reactive information displays. *Proceedings of INTERACT 2003: Ninth IFIP TC 13 International Conference on Human-Computer Interaction* (pp. 244-251). Amsterdam: IOS Press.
12. Rozenblit, L., Spivey, M., & Wojslawowicz, J. (1998). Mechanical reasoning about gear-and-belt diagrams: Do eye-movements predict performance? *Proceedings of Mind III: The Annual Conference of the Cognitive Science Society of Ireland* (pp. 158-165).
13. Yoon, D. & Narayanan, N. H. (2004). Mental imagery in problem solving: An eye tracking study. *Proceedings of the Third ACM Symposium on Eye Tracking Research & Applications*, ACM Press, To appear.