Restricted Focus Viewer: A Tool for Tracking Visual Attention

To appear in Proceedings of Diagrams 2000, Publisher Springer-Verlag

Alan F. Blackwell[†], Anthony R. Jansen[‡], and Kim Marriott[‡]

[†] Computer Laboratory
University of Cambridge
Cambridge CB2 3QG, UK
Alan.Blackwell@cl.cam.ac.uk

[‡] School of Computer Science and Software Engineering
Monash University
Clayton, Victoria 3800 Australia
{tonyj,marriott}@csse.monash.edu.au

Abstract. Eye-tracking equipment has proven useful in examining the cognitive processes people use when understanding and reasoning with diagrams. However, eye-tracking has several drawbacks: accurate eye-tracking equipment is expensive, often awkward for participants, requires frequent re-calibration and the data can be difficult to interpret. We introduce an alternative tool for diagram research: the Restricted Focus Viewer (RFV). This is a computer program which takes an image, blurs it and displays it on a computer monitor, allowing the participant to see only a small region of the image in focus at any time. The region in focus can be moved using the computer mouse. The RFV records what the participant is focusing on at any point in time. It is cheap, non-intrusive, does not require calibration and provides accurate data about which region is being focused upon. We describe this tool, and also provide an experimental comparison with eye-tracking. We show that the RFV gives similar results to those obtained by Hegarty (1992) when using eye-tracking equipment to investigate reasoning about mechanical diagrams.

1 Introduction

Visual attention is an important component in understanding how humans reason with diagrams. Complex diagrams can rarely be taken in at a single glance, and thus following the focus of visual attention can provide important insight into the strategies used in diagrammatic reasoning. For this reason, eye-tracking equipment has been of great benefit to researchers in examining these strategies. This applies not just in the area of diagrammatic reasoning, but also reading [8], cartography [13], scene perception [10] and cognitive processes in general [6, 16].

Yet despite the benefits that traditional eye-tracking provides, it also has significant drawbacks. First is expense: typically, the better the resolution accuracy and sampling rate, the more expensive the system is, and high quality systems are very expensive indeed. Second, many systems are awkward for participants, using head mounted gear, or requiring chin rests or bite bars to suppress head movements. Most systems require frequent re-calibration, and often they cannot be used with participants who wear glasses. Also, blinks or glances away from the stimulus can cause spurious trajectories in the eye movement data, and finally, and perhaps most important, it can be impossible to determine if a participant is taking in a broad overview of a stimulus, or focusing on a specific region. It is not surprising then that in some papers that discuss eye-tracking data, the results of some participants can not be used due to an inability to accurately record their eye fixations (for example, see [1, 11]).

Here we describe an alternative computer based tool for tracking visual attention: the Restricted Focus Viewer (RFV). This allows visual attention directed towards an image presented on a computer monitor to be tracked. This tool is a cheap, easy to set up system, that has flexibility allowing it to be tailored to specific diagrammatic elements of interest. It is non-intrusive and requires no calibration. We believe it will have wide appeal in research on cognitive aspects of diagrammatic understanding and reasoning. It is important to understand that the RFV is not intended to be a replacement for eye-tracking. Rather it is a complementary technique with its own advantages and disadvantages.

The RFV uses image blurring to restrict how much of the image can be clearly seen, with only a small region in focus. The region of focus can be moved around using a computer mouse. The idea behind using restrictions in the visual field is not new. In research on reading for example, the number of characters that can be processed in one fixation has been examined by using visual restrictions [9, 7]. Image blurring has been used to understand how people take in information from software manuals [15]. Studies have also been conducted which involve the use of artificial scotomas to disrupt visual processing [5], and the notion of a movable window has been used before in examining visual search [12].

The main contribution of this paper is threefold. First, we describe our implementation of the RFV in detail. It is a generic computer based tool, explicitly designed for research into diagrammatic reasoning. Features include graded blurring, motion blur and a data replay program (see Section 2). Second, we provide an empirical validation of the RFV. Results obtained from a mental animation experiment using eye-tracking equipment are compared with results obtained using the RFV instead (Section 3). Our third contribution is a detailed discussion of the relative advantages and disadvantages of the RFV over eye-tracking (Section 4).

2 The Restricted Focus Viewer

The Restricted Focus Viewer (RFV) is a computer program which takes a diagram, blurs it and displays it on a computer monitor, allowing the participant to

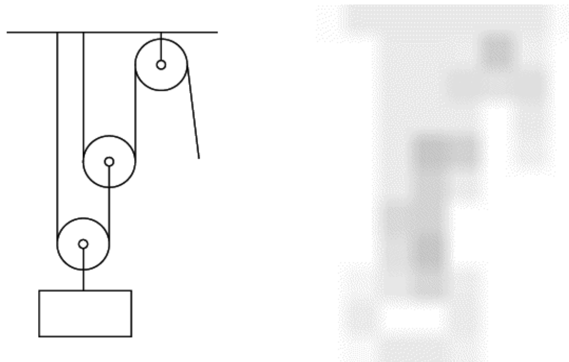


Fig. 1. Example of a visual stimulus and its corresponding blurred image.

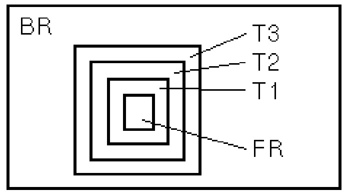
see only a small region of the diagram in focus at any time. The region in focus can be moved using the computer mouse. The RFV records what the participant is focusing on at any point in time, and the data can be played back using a replayer. We now describe the RFV and the replayer in more detail.

2.1 Description of the RFV

The human visual system can only focus on objects at the centre of the visual field. The region surrounding this area of sharp focus is still perceived, but the further from the centre of the visual field an object is, the more coarse is the perception of it [14]. Most of the time, visual attention is directed to the centre of the visual field, although this is not always the case as it is possible to covertly attend to other locations [2].

The RFV has been designed to reflect these aspects of the human visual system. The key idea is that only the part of the image under the focus window is in focus, with the remainder of the image out of focus. To keep most of the diagram out of focus, the RFV uses image blurring. Figure 1 gives an example of this with a pulley system diagram as the stimulus. The original system is shown on the left, with a blurred representation of it on the right. The blurred image still allows the general form of the diagram to be perceived, thus allowing the user to move directly from one region of the image to another. However, it does not reveal the finer details of diagram, with the individual components being indiscernible unless the user specifically focuses on them with the focus window. (Note that the blurred images in this paper have been made slightly darker so that they are clearer when printed. Some detail has been lost due to a reduced greyscale when printing.)

It is clear from the example figure that large structural features of the diagram are suggested in the blurred image. Our objective is that broad structure should be perceived as easily in the blurred image as by someone glancing momentarily at the original diagram. The degree of blurring required for a specific type of diagram depends on the size and visual characteristics of the diagram elements of interest. This is discussed in more detail later in this section.



FR – Focus Region
T1 – Transition Region 1
T2 – Transition Region 2
T3 – Transition Region 3
BR – Blurred Region

Fig. 2. Regions of the stimulus used to achieve the graded blurring effect.

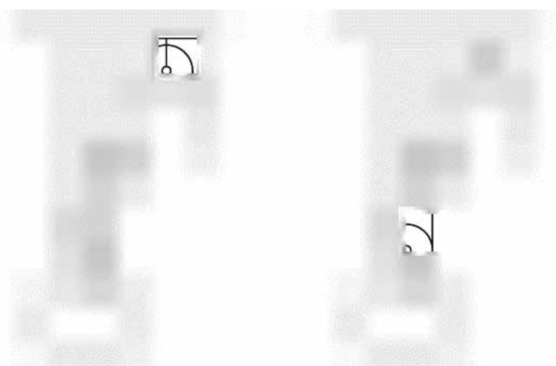


Fig. 3. Two examples of the focus window on different regions of the stimulus.

The *focus window* of the RFV is the region in which the stimulus is visible in full detail. Two important issues concerning the focus window for the RFV are firstly, how 'natural' it will look, and secondly, how big it should be.

Initial experience with the RFV led to a clear conclusion about the focus window: it is not sufficient to simply have a box on the screen in which the stimulus is in focus, while the rest of the image is blurred. The boundary between the two regions is too distinct, leading to a very unnatural effect. Also, a sharp cut-off will enable the participant to guess about neighbouring parts of the image from Gestalt continuation effects. A graded blurring effect, such that the transition from blurred to focus appears smooth and seamless, is needed to prevent this.

A graded blurring effect is achieved by the technique illustrated in Figure 2. The outer rectangle defines the stimulus area which is fully blurred. The innermost box is the region of focus. Surrounding this focus region are three transition regions. Each transition region is slightly more blurred than the last, so that there is only a subtle difference between neighbouring regions. The overall result is the appearance of a smooth transition from the region of the image in focus, to the region which is fully blurred. Using the mouse to move the focus window therefore moves not only the focus region, but also the transition regions.

Figure 3 gives two examples of the focus window positioned over different regions of the stimulus shown in Figure 1. The size of the focus window is determined not only by the dimensions of the focus region, but also by those

Focus Region Size	Goal	Should be slightly smaller than bounding box of typical element of diagram
	Lower Limit	Must allow recognition of any one element of the diagram when region is centred over the element
	Upper Limit	Should prevent simultaneous recognition of two neighbouring elements when placed between them
Transition Region Size	Goal	Should indicate direction of neighbouring connected elements
Level of Blurring	Lower Limit	Should be sufficient that any two elements are indistinguishable and that overall connectivity cannot be established
	Upper Limit	Should allow identification of diagram boundaries (at least convex hull)

Table 1. Guidelines for setting RFV parameters.

of the three transition regions. However, the two most important sets of values to be considered are the size of the focus region box, and the size of outermost transition region box. For the experiment discussed in this paper, the focus region and transition region boxes are all square, and centred around the current mouse co-ordinates, although this need not be the case. At the centre of the focus window there is also a small grey dot, to allow users to keep track of the focus window location when it is centred on an empty region of the image.

The degree of blurring and the size of the focus region should be adjusted in accordance with the size of syntactic elements in the diagram. In our experience with the RFV to date, the rules of thumb described in Table 1 should be applied in determining these factors. Note that although these parameters can be adjusted easily for a given experiment, the RFV program does not allow them to be changed during an individual trial. This is to prevent inconsistencies in the way that the participant views the stimulus.

Another feature that was implemented so that the RFV would more accurately mimic the way humans perceive visual stimuli is *motion blur*. This is based on the fact that during saccadic eye movements, visual information is not processed [9]. Thus, if the user of the RFV moves the mouse at high speed (that is, over a large distance on the screen in a small amount of time), the focus window will not achieve full focus. Once the user reduces the speed of the mouse motion back to below a certain threshold, or stops moving the mouse completely, full focus in the focus window will return. This feature helps in defining the temporal boundary between fixations and movements.

When the focus window is stationary or moving slowly, all of the regions listed in Figure 2 are present. During motion blur however, only the outermost transition region is added to the blurred stimulus. Because this region has less blurring than the rest of the image, the user is still able to track the location of the focus window on the stimulus. However, it is not possible to determine the finer details of that location without slowing or stopping the mouse. Only then will full focus be available. Table 2 describes guidelines for appropriate motion blur settings. That is, how fast the mouse needs to move before motion blur occurs.

Motion Blur Onset	Goal	Should allow separation between fixation and movement
	Lower Threshold	Should not allow "brass rubbing" strategy, that is, identification of diagram by waving window rapidly over it
	Upper Threshold	Should allow slow navigation with continuous focus over a connected diagram when the task requires it

Table 2. Guidelines for the RFV motion blur setting.

For each stimulus, the RFV outputs a data file that records the motion of the focus window. The file consists of a brief header to specify which trial the data pertains to. For the remainder of the file, each line contains the details for each updated mouse movement. The lines are composed of four data. First is the time elapsed (in milliseconds) since the RFV was initiated for that particular trial. The next two values are the x and y co-ordinates of the centre of the focus window with respect to the top left corner of the stimulus image. The final piece of information is a flag to indicate whether or not the focus window was motion blurred or not. This data allows the experimenter to exactly replicate the state of the RFV while the participant was performing the task.

2.2 Data Replayer

The data replayer is a companion program to the RFV that can read in a data file generated by the RFV, and replay the way that the RFV participant moved the focus window over the stimulus. This has two important benefits.

Firstly, it can be used as an experimental analysis tool by the experimenters when reviewing participant actions. For example, it can play back the focus trace at faster than real-time. The data replayer also has another function that is useful to experimenters. It can draw a scan-path line over the original stimulus based on the locations of the centre of the focus box. Figure 4 gives an example of this for the pulley system diagram. It can be seen that in this instance, the person started at the free end of the rope on the right of the diagram. From there they moved to the top pulley, continued to the middle pulley, and finally moved down to the bottom pulley. At each pulley, time was spent examining that region of the diagram.

The second important benefit is that the data replayer can be used during experimental sessions to elicit retrospective verbal reports from participants, about their strategy and actions during an experimental trial. As noted by Ericsson and Simon [3], it can be very difficult for participants to verbalise concurrently while carrying out a complex problem solving task. We have used the data replayer

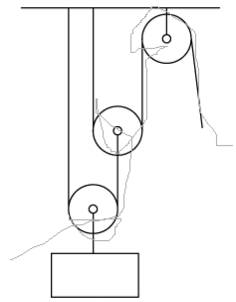


Fig. 4. Example of a data replayer scan-path.

to play back focus movements slower than real-time, so that the participant can record a verbal protocol describing their actions after completing the main task. Where Ericsson and Simon report that it is difficult to report retrospectively on a problem taking longer than 10 seconds to solve, the use of the data replayer with the RFV reminds the participant of his or her actions, and allows more extensive protocols than can usually be obtained.

3 Validation of the RFV

The key question before we can consider using the RFV in experimental work is to determine if it interferes with the strategies used by humans who perform diagrammatic reasoning. Clearly, there is some overhead since the participant must use a computer mouse rather than simple eye and head movements to change their focus. However, it is important to the validity of the RFV technique that this should not affect the strategies used. To test this, we have replicated a classic experiment on diagram interpretation [4], originally done with eye-tracking equipment, and instead used the RFV. Our hypothesis is that participants will use the same strategy with the RFV as with eye-tracking.

3.1 Method

Participants Eleven participants successfully completed the experiment. All were graduate or undergraduate students from the Computer Science and Psychology Departments. All participants were volunteers with normal or corrected-to-normal vision. Data from one additional participant was excluded due to excessive error rates, however the errors were due to a misinterpretation of the instructions, and were not attributable to the RFV.

Materials and Design Two diagrams of pulley systems were constructed based on those used in Experiment 1 in Hegarty [4] (see Figure 5). They each consisted of three pulleys, a weight, braces from the ceiling to support some pulleys and

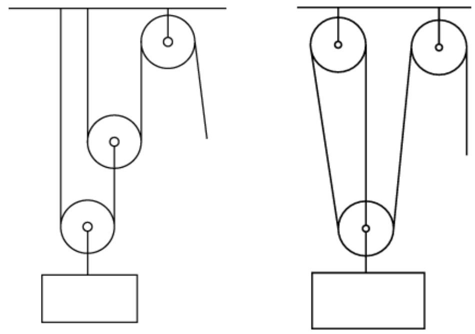


Fig. 5. The two pulley systems used.

sections of rope that were attached to the ceiling or weight, and went over or under the pulleys. In each system there was also a free end to the rope, and participants were required to infer the motion of the system components when the free end of the rope was pulled. The mirror images of these pulley systems were also used giving a total of four pulley system images.

For each pulley image there were twelve statements, each about the motion of one of the systems components. Six of these statements were true and six were false. For this experiment only kinematic statements were used (referring to the system in motion), with none of the statements addressing static aspects of the pulley system.

When the free end of the rope is pulled in any pulley system, a causal chain of inferences can be made about the motion of the components. For example, in pulley system 1, pulling the rope causes the rope to move right over the upper pulley, turning it clockwise. From this knowledge we can infer the motion of the middle pulley and so on. In this way, we can define each pulley in the pulley systems as being at the beginning, middle or end of the causal chain of events. The statements about the motion of the pulley system components are equally divided among the pulleys at each of these three locations in the causal chain. In pulley system 2, “the rope moves to the right under the lower pulley” is an example of a true statement about a kinematic event at the middle of the causal chain, and “the upper right pulley turns counterclockwise” is an example of a false statement about a kinematic event at the beginning of the causal chain.

Each statement was presented as a single line of text. A stimulus was composed of a statement appearing on the left, with the diagram of a pulley system on the right. This gave a total of 48 stimuli. Each participant was shown all 48 stimuli twice, with a rest between the two blocks. In each block, the stimuli were presented in a different pseudo-random order.

Procedure Participants were seated comfortably in an isolated booth. Items were displayed in black on a white background on a 17” monitor at a resolution of 1024×768, controlled by an IBM compatible computer running a version of the RFV program whose settings had been tailored to this experiment. The original eye-tracking experiment was conducted using an Iscan corneal-reflectance and pupil-center eye-tracker (Model RK-426), that has a resolution of less than one degree of visual angle, and samples the participants' gaze every 16 milliseconds.

The size of the images were 200×293 pixels for pulley system 1 and 160×300 pixels for pulley system 2. The text statements were on average 338 pixels across (range 244–417) and 16 pixels high. The RFV focus box had an edge length of 36 pixels, while the outermost transition box had an edge length of 50 pixels. The motion blur was set to a high tolerance, allowing for full focus to be maintained even during moderately fast movements of the mouse.

Participants were given a brief statement of instructions before the experiment began, and were shown diagrams which labelled all of the pulley components referred to in the statements. They were then given some practice items involving a very simple (only two pulleys) pulley system. The practice items allowed the participants to become familiar with the RFV program.

The only interface mechanism used by the participants was a standard computer mouse. Progress was self-paced, with each trial initiated by a single click with the mouse on a button at the bottom of a blank screen containing the prompt "Press the button to continue". This action started the timer (to provide a reference time for the rest of that trial), and a blurred image of the stimulus appeared on the screen. On the left was a statement and on the right was a diagram of a pulley system. The participant could move the window of focus over different regions of the stimulus by moving the mouse. By doing this, the participant could then read the text of the statement and look at the attributes of the diagram.

The participants were required to determine if the statement was true or false with respect to the pulley system presented. At the bottom of the screen were two buttons, one labelled "TRUE" and the other "FALSE". When the participant had decided on the validity of the statement, they single clicked on the appropriate button. This stopped the timer, and the program recorded the response given and the time taken. No feedback was given to the participant.

Participants were instructed to try and respond as quickly as possible, while still trying not to make too many errors. The experiment consisted of two blocks of 48, with a brief rest period in between. The full set of stimuli was shown in each block, resulting in two repetitions for each item in the experiment. The experiment took approximately forty minutes to complete.

Data Treatment To reduce the unwanted effects of outlying data points, an absolute upper cut-off was applied to response latencies, such that responses longer than 30 seconds were excluded from the response time data analysis and designated as an error. So as to be consistent with the original eye-tracking experiment, the data analysis was only conducted over the items that contained true statements, with the false items only acting as fillers in the experiment.

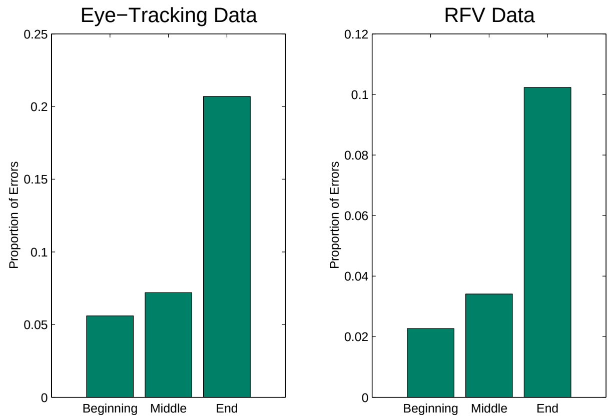


Fig. 6. Comparison of eye-tracking data from Hegarty's original experiment and RFV data from our replicated experiment, examining the proportion of errors at different causal chain positions.

For the one participant whose results were excluded, this was due to getting no correct responses at a particular causal chain location in one of the pulley systems, and thus not allowing the calculation of a mean response time at that position.

3.2 Results and Discussion

The aim of this experiment is to compare the results obtained using the RFV with eye-tracking results, to determine if the RFV affects the strategies used by humans who are reasoning with diagrams, in a manner different to the way that eye-tracking equipment might affect strategy. The main focus of this analysis is therefore to examine key data trends and significant results obtained in the original eye-tracking experiment, and see if the RFV results correlate.

Errors The overall error rate was only 5.3%. This is much lower than the error rate of the participants in the original eye-tracking experiment. However, this is not surprising given that the participants in the original experiment were all psychology undergraduates, and many of the participants in this experiment had more experience in dealing with technical diagrams.

The data comparisons were conducted using two-way ANOVAs (causal chain position $\times$ repetition), carried out over participant data. In the original eye-tracking experiment, the position in the causal chain of the component referred to in the statement had a significant effect on error rates. This effect was also present in the RFV data ($F(2, 20) = 5.33$, $p < 0.05$). This can be clearly seen in Figure 6, where the data from this experiment on the right is compared with the data from the original eye-tracking experiment on the left. There was also a trend for participants to make fewer errors on the second repetition of the stimuli (3.0%, sd = 5.4) than on the first repetition (7.6%, sd = 12.5), however

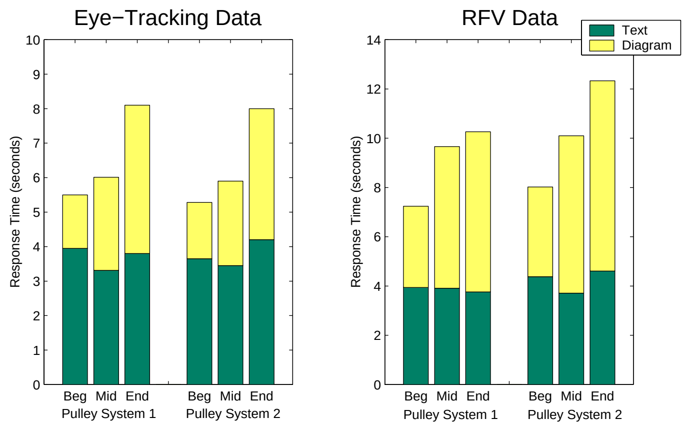


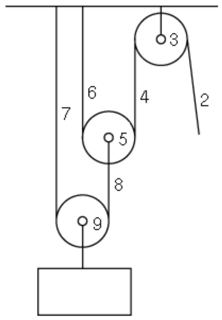
Fig. 7. Comparison of eye-tracking data from Hegarty's original experiment and RFV data from our replicated experiment, examining the mean response times for different trial types.

this trend was not statistically significant ($F(1,\ 10) = 4.52$, $p = 0.059$). This trend was also apparent in the eye-tracking experiment.

Response Times Figure 7 shows the mean reaction times (overall height of the bars) for each pulley system, for statements referring to components at different positions in the causal chain. As with the error graphs, the data from this experiment is shown on the right, with the eye-tracking data on the left for comparison. The times have been divided into the time spent reading the statement, and the time spent inspecting the diagram.

Response times for the two pulley systems were analysed separately, to allow for differences in the configurations. In the original eye-tracking experiment, the repetition caused a practice effect which resulted in participants responding significantly faster on the second repetition of the stimuli than on the first. The same effect was seen in the RFV data. The response advantage was 3.65 seconds for pulley system 1 ($F(1,\ 10) = 44.40$, $p < 0.01$), and 4.05 seconds for pulley system 2 ($F(1, 10) = 38.07$, $p < 0.01$). The original experiment also showed that the position in the causal chain of the component referred to in the statement had a significant effect on response times. Again, the RFV data corresponds to the data from the original experiment, with position in the causal chain significantly influencing response times ($F(2,\ 20) = 18.87$, $p < 0.01$ for pulley system 1; $F(2,\ 20) = 24.15$, $p < 0.01$ for pulley system 2). This effect can be seen in Figure 7.

These results are clearly in agreement with the eye-tracking experiment. However, the participants using the RFV take approximately 50% longer to respond. Response time is likely to be affected by the fact that the participants interface with the stimulus using a computer mouse. This is due to the difference in the ballistic speed of human motor control of the arm and the eye. Also, the blurring of the image with only a small region of focus means that one would expect



Gaze	Object Fixated	Gaze Duration
1	Statement	1884 ms
2	Pull rope	500
3	Upper pulley	595
4	Right upper rope	345
5	Middle pulley	1160
6	Left upper rope	180
7	Left lower rope	150
8	Right lower rope	61
9	Lower pulley	5917

Fig. 8. Example of an RFV eye-fixation protocol for the statement *The lower pulley turns counterclockwise*

participants to take longer in moving from one area of the diagram to another area. Despite the extra time taken, the overall trend in the data is very similar.

Further data analysis that was done in the eye-tracking experiment involved examining how long participants were inspecting different components of the pulley systems. In particular, for each statement, the components in the diagram were divided into those whose motion occurred before the component referred to in the statement, the referent itself, and those components whose motion occurred after the referent. This allows for a further breakdown of response time spent viewing the diagram. Rectangular bounding boxes were used to enclose regions of the diagrams containing pulleys, rope strands, the ceiling and the weight, just as in the original eye-tracking experiment. This allowed the order of fixations on the components of the diagrams to be determined, along with how long those fixations were. Figure 8 gives an example of an eye-fixation protocol taken from the RFV data.

The gaze duration is defined as the total time spent fixating on components, in certain locations in the causal chain with respect to the referent in the statement. The graphs of the gaze duration data are shown in Figure 9. Again, the data from this experiment is on the right, with the eye-tracking data on the left. The original eye-tracking experiment showed that when looking at the pulley system, participants spend most of their time inspecting the referent and the components whose motion precedes that of the referent in the causal chain of events. The RFV data shows the same result for both pulley system 1 (91%, sd = 4.9) and pulley system 2 (92%, sd = 4.6).

Due to the fact many participants in our experiment had more experience with technical diagrams than the participants in the original eye-tracking experiment, we would expect some minor differences in strategy. However, the results of this experiment indicate that the response time, accuracy and gaze duration trends obtained from the original experiment using eye-tracking techniques, are

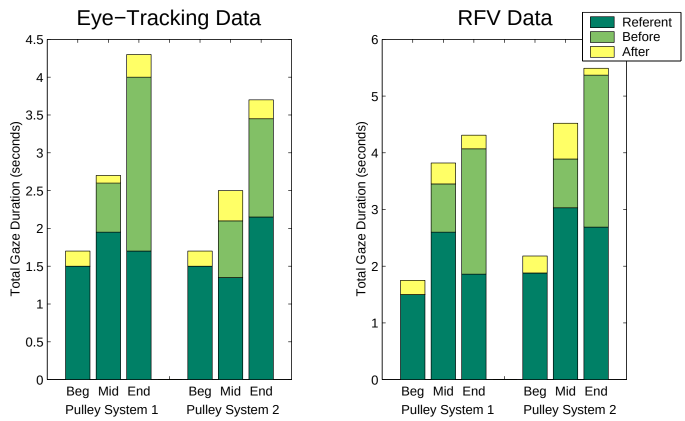


Fig. 9. Comparison of eye-tracking data from Hegarty's original experiment and RFV data from our replicated experiment, examining the breakdown of gaze duration on different components of the pulley systems.

the same as those obtained using the RFV. There were no key significant results from the original eye-tracking experiment that the RFV failed to obtain.

4 Comparison of the RFV with Eye-Tracking

The Restricted Focus Viewer (RFV) is designed to collect data about visual attention, as are eye-trackers. However, the RFV is not intended as a replacement for eye-tracking techniques. Rather it is an alternative experimental technique and apparatus that adds to the toolbox available to cognitive scientists as they try to understand the processes of diagrammatic reasoning.

Here, we explain the relative advantages and disadvantages of the RFV and eye-tracking. One primary concern is whether the RFV affects the high level strategy used by participants. We note that this issue is not just confined to the RFV: in eye-tracking also, repeated calibrations and head gear (such as head mounted cameras and infra-red reflectance devices) may influence participant strategy. The previous experiment suggests that high level strategy is not changed by using the RFV instead of eye-tracking equipment. However, this depends on the task and the choice of RFV parameters. In particular, there are two important issues to consider.

The main issue is that with the RFV, the experimenter can modify the size of the focus window and the amount of blurring, so as to ensure that the participant must explicitly focus on those components of the diagram they are interested in. Such changes however, may change the strategy used by the participant.

Consider for example, a mathematical equation as the visual stimulus. It is preferable that the size of the equation not be excessively large, since this would make any task involving it seem less natural to the participant. Trying to determine the specific symbols that are being focused on would be practically impossible using eye-tracking, since each eye fixation would take in a large collection of symbols. With the RFV, the size of the focus window can be reduced so that only a few or even only one symbol can be viewed at a time. This would allow visual attention to be recorded at a level of detail not available using eye-tracking techniques. However, the reduction in focus window size also reduces the amount of information available to the participant at any given moment. This could affect the way a given task is approached. It appears that the more accurately you record the focus of the participants attention (by reducing the size of the focus window), the more likely you are to affect the strategy that they would normally use. However, this is not a defect of the RFV, but rather an aspect that experimenters need to be aware of in experiment design.

The second issue is that a computer mouse is used rather than eye and head movements to change the direction of attention. Thus, participants in the experiments need to be confident in the use of a mouse. Also, response times will be slower due to the use of the arm and hand. For some tasks that require fast responses, this may present a problem. However, we believe that for most tasks this should not affect data trends or the significance levels of experimental results. One difference between the RFV and eye-tracking is what happens when the participant is not explicitly focusing on any part of the diagram, because their attention is directed towards internal processing. At this time, their gaze may drift across the stimulus. With eye-tracking this means that spurious fixations may appear in the data, while with the RFV the object last in focus will have a longer gaze duration.

By comparison, we have experienced far more significant difficulties when using eye-tracking equipment (ASL 5000) for experimental tasks to which we had easily applied the RFV. These problems will be familiar to experienced users of eye-tracking equipment, but are nevertheless significant obstacles to new researchers. Eye-trackers are expensive. They require substantial expertise in calibration and adjustment. They do not work reliably in strong daylight conditions. They can be unreliable with participants who have shiny skin, watery eyes or contact lenses. Output data is often subject to positional drift, in addition to local nonlinear uncertainty. Vertical resolution is poor compared to horizontal resolution, making them less useful with detailed two-dimensional diagrams. Analysis of data requires subjective classification of fixation and saccade thresholds. Fixations are often at a point between two display elements, leaving it unclear whether the participant is defocused or viewing both elements as a unit. The tracker can lose stable gaze identification during the experiment, leading to invalid trials. By comparison to experiments conducted with the RFV using the same stimuli, eye-tracking results provided very little useful data.

Overall, the RFV has several advantages over traditional eye-tracking techniques. The system is cheap and easy to set up, providing accurate data about the region that is being focused upon. It is non-intrusive, requiring no special gear to be worn or restrictions on the movement of participants. It does not require any calibration and can be used by participants who wear glasses. The RFV data is not corrupted by blinks or glances away from the stimulus, and the replayer provides a useful tool for immediate feedback on participant performance. Finally, the RFV has flexibility in its parameter settings, allowing it to be tailored to meet specific goals.

5 Future Work

We have described a new tool, the RFV, for tracking visual attention during diagrammatic reasoning. We have also provided an experimental comparison with eye-tracking equipment. The primary direction for future work is to expand this experimental comparison by considering other classes of diagrams. We also intend to extend the RFV and data replayer by adding extra features. Such features include modifying the data replayer to allow it to replay multiple scan-paths at once. This would be useful in allowing for a direct comparison of the strategies used by different participants. The RFV is available in the public domain for other researchers to use at

`http://www.csse.monash.edu.au/~tonyj/RFV/`

Acknowledgements

The authors wish to thank Mary Hegarty for her helpful information and discussions on the original work with the pulley systems using eye-tracking equipment. We also wish to thank Greg Yelland for his suggestions during the development of the RFV program. Alan Blackwell's research is funded by the Engineering and Physical Sciences Research Council under EPSRC grant GR/M16924 “New paradigms for visual interaction”.

References

1. Patricia A. Carpenter and Priti Shah. A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4(2):75–100, 1998.
2. Stanley Coren, Lawrence M. Ward, and James T. Enns. *Sensation and Perception*. Harcourt Brace and Co., fourth edition, 1994.
3. K. Anders Ericsson and Herbert A. Simon. *Protocol Analysis: Verbal Reports as Data*. MIT Press, Cambridge, Massachusetts, revised edition, 1993.
4. Mary Hegarty. Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(5):1084–1102, 1992.
5. John M. Henderson, Karen K. McClure, Steven Pierce, and Gary Schrock. Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics*, 59(3):323–346, 1997.
6. Marcel Adam Just and Patricia A. Carpenter. Eye fixations and cognitive processes. *Cognitive Psychology*, 8:441–480, 1976.
7. Naoyuki Osaka and Koichi Oda. Moving window generator for reading experiments. *Behavior Research Methods, Instruments & Computers*, 26(1):49–53, 1994.
8. Keith Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3):372–422, 1998.

9. Keith Rayner and Alexander Pollatsek. *The Psychology of Reading.* Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
10. Keith Rayner and Alexander Pollatsek. Eye movements and scene perception. *Canadian Journal of Psychology*, 46:342–376, 1992.
11. Susan K. Schnipke and Marc W. Todd. Trials and tribulations of using an eye-tracking system. In *CHI 2000 Extended Abstracts*, pages 273–274, 2000.
12. L. W. Stark, K. Ezumi, T. Nguyen, R. Paul, G. Tharp, and H. I. Yamashita. Visual search in virtual environments. *Proceedings of the SPIE: Human Vision, Visual Processing, and Digital Display III*, 1666:577–589, 1992.
13. Theodore R. Steinke. Eye movement studies in cartography and related fields. *Cartographica*, 24(2):40–73, 1987. Studies in Cartography, Monograph 37.
14. Martin J. Tovée. *An Introduction to the Visual System.* Cambridge University Press, 1996.
15. Nicole Ummelen. *Procedural and declarative information in software manuals.* PhD thesis, Universiteit Twente, Utrecht, 1997.
16. Alfred L. Yarbus. *Eye Movements and Vision.* Plenum Press, New York, 1967.