Judgment and Decision Making, Vol. 6, No. 8, December 2011, pp. 740–749

Applying the *decision moving window* to risky choice: Comparison of eye-tracking and mouse-tracing methods

Ana M. Franco-Watkins* Joseph G. Johnson†

Abstract

Currently, a disparity exists between the process-level models decision researchers use to describe and predict decision behavior and the methods implemented and metrics collected to test these models. The current work seeks to remedy this disparity by combining the advantages of work in decision research (mouse-tracing paradigms with contingent information display) and cognitive psychology (eye-tracking paradigms from reading and scene perception). In particular, we introduce a new *decision moving-window* paradigm that presents stimulus information contingent on eye fixations. We provide data from the first application of this method to risky decision making, and show how it compares to basic eye-tracking and mouse-tracing methods. We also enumerate the practical, theoretical, and analytic advantages this method offers above and beyond both mouse-tracing with occlusion and basic eye tracking of information without occlusion. We include the use of new metrics that offer more precision than those typically calculated on mouse-tracing data as well as those not possible or feasible within the mouse-tracing paradigm.

Keywords: decision making, eye tracking, process tracing, metrics.

1 Introduction

Decision researchers must often rely on outcome measures (i.e., choice, preference, etc.) to infer *how* decisions are made by individuals. The advent of increased computing power, availability, and usability has allowed decision researchers to develop methods to examine underlying processes (attention, information acquisition, deliberation, etc.) rather than relying solely on observable outcomes. Process-tracing paradigms such as information boards and mouse-tracing methods (e.g., Payne, Bettman, & Johnson, 1988, 1993), eye-tracking methods (e.g., Franco-Watkins & Johnson, 2011; Glöckner & Herbold, 2011; Horstmann, Ahlgrimm, & Glöckner, 2009; Russo & Rosen, 1975), and verbal protocols (see Ranyard & Svenson, 2011, for review) have been utilized in decision research as mechanisms for capturing the decision process. The need to verify theoretical advances that specify *how* (i.e., process) in addition to *what* (i.e., outcome) has produced a great interest in tracing methods (see Schulte-Mecklenbeck, Kuehberger, & Ranyard, 2010 for an overview).

The research was supported in part by grant SES–0851990 from the U.S. National Science Foundation.

*Department of Psychology, Auburn University. Email: afrancowatkins@auburn.edu.

†Department of Psychology, Miami University. Email: johnsojg@muohio.edu.

1.1 Process-tracing methods

In the ubiquitous mouse-tracing paradigm, information on a computer screen is occluded from the decision maker until the individual places a mouse cursor over a specific region to reveal the corresponding information. By occluding the information, researchers are able to determine how, when, and which information is acquired (revealed) during the decision process. Mouse-tracing paradigms have been successfully applied to different types of decisions such as probabilistic inference, probability, and marketing (see Norman & Schulte-Mecklenbeck, 2010 for a review). This technique was revolutionary given the *Zeitgeist* in which it was developed and has provided a precedent and recognition of the importance of process-tracing as an additional tool for the decision researcher. However, several decision researchers have noted theoretical and methodological limitations with the mouse-tracing paradigm (e.g., Glöckner & Betsch, 2008b; Johnson & Koop, 2010; Koop & Johnson, 2011; Norman & Schulte-Mecklenbeck, 2010). For example, recent research has shown that the mouse-tracing paradigm itself may affect the information search process, possibly introducing experimental artifacts and/or confounding measurement of attention (Glöckner&Betsch, 2008b).

Fortunately, some of the limitations inherent in mouse-tracing paradigms can be improved upon with eye-tracking technology. If one assumes that visual attention and eye movements are coupled (cf., Hoffman, 1998; Rayner, 1998), attentional shifts can be cap-

https://doi.org/10.1017/S1930297500004174 Published online by Cambridge University Press

tured by changes and patterns associated with eye movements. Converging evidence between mouse-tracing and eye-tracking methods were initially established to support the mouse-tracing methodology (Lohse & Johnson, 1996), however, recent successful applications of eye tracking have examined decision-making processes across a variety of decision tasks: probabilistic inference (Horstmann et al., 2009), risky decisions (Glöckner & Herbold, 2011), consumer decisions (Rayner, Rotello, Stewart, Keir, & Duffy, 200l), and advertisements (Wedel & Pieters, 2000). Here, we extend the use of eye-tracking via a new methodology, the *decision moving window,* that combines the advantages of work in decision research (occlusion with contingent information display) and cognitive psychology (eye-tracking paradigms from reading and scene perception).

1.2 Decision moving window

The decision moving-window (DMW) paradigm presents stimulus information contingent on eye fixations rather than cursor placements. Specifically, information is masked from the user until the decision maker fixates on a given area to reveal the corresponding information, one at a time. As soon as the eye (fovea) moves away from the area, the mask returns, and the previously viewed information is hidden again. The primary advantage of this paradigm is that the researcher has a more direct and reliable measure of overt attentional processing (i.e., selective attention) during decision making while allowing the subject to effortlessly determine how the information is revealed. In Franco-Watkins and Johnson (2011), we introduced the DMW methodology in greater detail as well as made comparisons to standard eye-tracking (ET) and mouse-tracing (MT) paradigms in a probabilistic inference task. We found that both eye-tracking methodologies appeared to have an advantage over MT by producing a greater number of fixations, of shorter duration, and were less susceptible to significant variability over the course of an experiment. Additionally, the DMW allows for more direct comparisons with MT paradigms often used in decision research. The current work represents the first application of the DMW to a risky decision making task and comparison to existing methods to further examine how attentional processing affects the acquisition of information. Additionally, it introduces new analyses (involving transition matrices and pupil dilation) for understanding dynamic processes in decision making. These important extensions serve to establish the robustness of the method across the most common experimental tasks in decision research and to augment the standard repertoire of analytic methods on process-tracing data in general, and eye-tracking data in particular.

2 Method

Subjects and Design. Seventy-eight undergraduate students participated in the experiment in exchange for extra credit towards a psychology course. Subjects were assigned to either the DMW (n = 26), ET (n = 26), or MT (n = 26) paradigm.

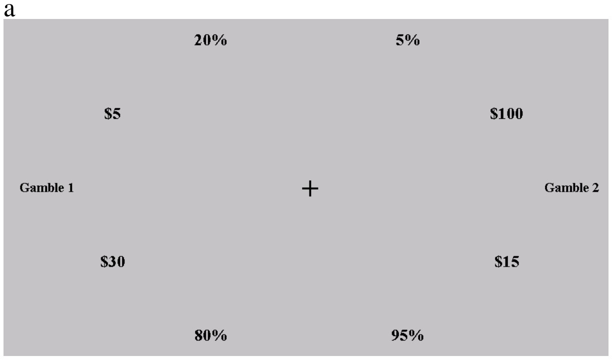
Materials, Apparatus, and Procedure. The decision task consisted of 40 pairs of gambles previously used with MT and ET paradigms (Glöckner & Betsch, 2008a; Glöckner & Herbold, 2011). The gambles reflected five decision task categories, abbreviated as: CERT_{PRO}, CERT_{CON}, SIM, $\text{MED}_{\text{ALM\_CERT}}$, and MED_{CERT}. The CERT_{PRO} category favored gamble A because of the zero outcome for gamble B and conversely CERT_{CON} category favored gamble B because of the zero outcome for gamble A. The SIM category reflects similar outcomes and values between both gambles. The $\text{MED}_{\text{ALM\_CERT}}$ and MED_{CERT} categories reflect medium outcomes that are almost certain (gamble B) and certain (gamble A), respectively. Each decision task category consisted of 8 trials (see Table 1 and Appendix of Glöckner & Herbold, 2011, for complete stimuli). We counterbalanced left-right presentation of Gamble within a pair and top-bottom presentation of Outcome within a gamble per category, and randomized presentation of all 40 trials for each subject.

Eye movements were recorded using the binocular tracking of the Tobii 1750 eye tracker (17" monitor with 1280 x 1024 pixels, sampling rate: 50Hz, spatial resolution: 0.5°; and calibration accuracy: 0.5°). To develop the DMW program, we used E-prime extensions for Tobii (Psychology Software Tools). Figure 1 (top panel) presents a sample decision task. For all paradigms, an area of interest (AOI) is a cell associated with each piece of gamble information (i.e., value or probability). We created a black mask (Figure 1, middle panel) that covered each AOI for the DMW and MT paradigms. When the subject placed their eye or mouse cursor on a specific AOI, the black mask immediately became transparent revealing the gamble information (Figure 1, bottom panel). Whenever, the subject's eye or cursor moved away from the AOI, the gamble information was immediately hidden. Hence, eye and cursor movements to a specific AOI revealed the corresponding information and all other information remained occluded. Furthermore, the labels Gamble 1 and Gamble 2 and the center cross were AOIs, however, this information remained visible on the screen. All AOIs were 150 x 100 pixels and positioned equally distant from the center cross.

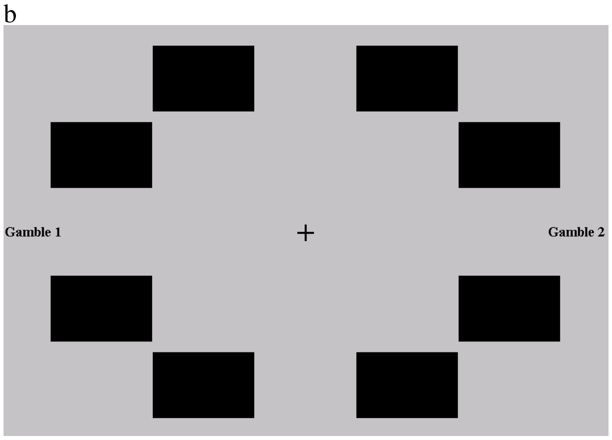
Upon arrival to the laboratory, subjects in the eye-tracking paradigms completed the eye-tracking calibration using Tobii Studio software. All subjects received onscreen instructions detailing how to complete the deci-

Figure 1: Sample screen shot of decision task: basic eye tracking (a) moving window or mouse tracing starting state where information is occluded (b) until the individual moves their eyes or cursor to specific AOI (c).

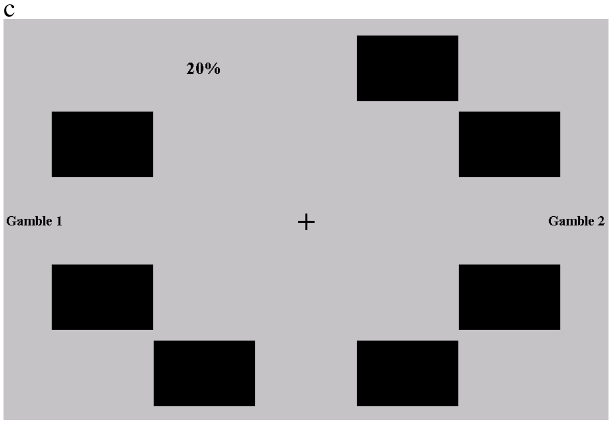
a



b

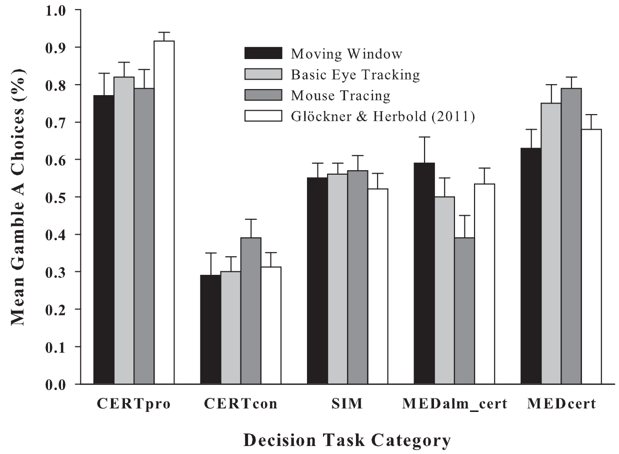


c



sion task (derived from Glöckner & Herbold, 2011). Next subjects practiced two trials of the decision task without occlusion. Subjects in the DMW and MT paradigms had two additional practice trials where they practiced moving their eyes or cursor to reveal gamble information. Practice stimuli were novel from experiment stimuli. The sequence for each trial consisted of a 1000ms screen with

Figure 2: Average (standard error) gamble A choices per decision category across paradigms.



the center cross alone, followed by gathering gamble information and making a choice (using the "1" and "2" keys) which terminated the trial. A rest screen was placed between trials and subjects initiated advancement to the next trial.

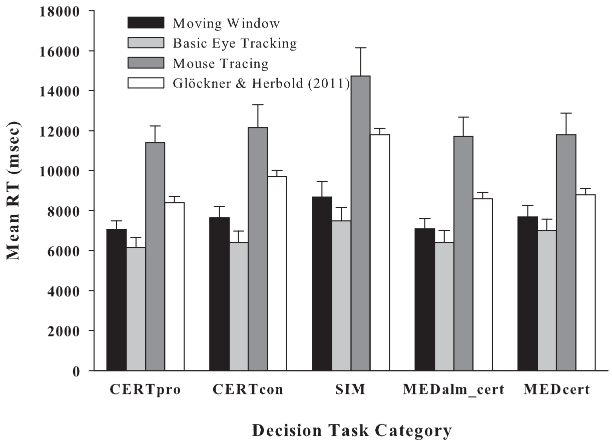
3 Results

We conducted a 3 (paradigm: DMW, ET, and MT) x 5 (decision category) mixed factorial ANOVA with paradigm as the between-subjects factor and decision category as the within-subject factor for all dependent variables (choice, RT, and process variables).

Choice. Figure 2 presents the average proportion of gamble A choices per decision category per paradigm. Glöckner and Herbold's (2011) results are depicted for comparison purposes. Analyses revealed a significant main effect of decision category, F (4, 300) = 36.30, $p < .01$, $\eta^2_p = .33$ demonstrating the subjects were sensitive to different gambles and outcomes associated with each category. The main effect of paradigm was not statistically significant ($p = .56$, $\eta^2_p = .01$), however, the interaction effect was approaching significance F (8, 300) = 1.74, p =.09, $\eta^2_p = .04$. The data in Figure 2 suggests this interaction is most likely due to differences in the latter two decision categories (MED_{CERT} and MED_{ALM_CERT}) between MT and DMW paradigms. Notably, post-hoc analyses revealed differences between MT and DMW for MED_{ALM_CERT} (t = 2.31, p = .02, d =.59) and MED_{CERT} (t = 2.41, p =.02, d = .68) decision categories. Inspection of Figure 2 indicates that the DMW subjects appeared to prefer the riskier option (59 % for Gamble A) relative to the MT subjects (39% for Gamble A) in the MED_{ALM_CERT} decision category where Gamble B

https://doi.org/10.1017/S1930297500004174 Published online by Cambridge University Press

Figure 3: Mean (standard error) decision reaction times per decision task category across paradigms.



is the safer option because the medium outcome is almost certain. Post-hoc comparisons did not reveal significant differences for the MED_{ALM_CERT} decision category between the two eye-tracking paradigms (p = .29). In the MED_{CERT} decision category where gamble A is the safer option (medium gamble with certainty), DMW subjects (37% for Gamble B) again selected the riskier choice more often than the MT subjects (21% for Gamble B). Additionally, post-hoc analyses revealed a trend approaching significance for the DMW subjects to select the riskier choice more often than those using basic eye-tracking (25% for Gamble B) subjects (p = .08, d =.46). In sum, it appears that risk preferences may differ between the two occlusion paradigms (DMW and MT) for some stimulus configurations (here, when gambles had two nonzero outcomes and were dissimilar), and to a lesser extent perhaps also between the two eye-tracking paradigms (DMW and ET). We did not anticipate *a priori* that different choice patterns would emerge between the two occlusion methods and refrain from speculating on these further. Our focus was on how these different methods reveal search behavior; we turn now to analyses of process measures.

Decision Time. Figure 3 presents mean decision times per decision task category across the different paradigms. The MT method resulted in longer decision latencies than the two eye-tracking-based methods leading to a significant effect of paradigm, F (2, 75) = 17.56, p < .01, η^2_{p} = .32. We again replicated the significant main effect of category from Glöckner and Herbold (2011), F (4, 300) = 17.40, p < .01, η^2_{p} = .19. The interaction was approaching statistical significance, F (8, 300) = 1.86, p = .07, η^2_{p} = .05. The SIM category resulted in the longest latencies compared to the other decision category conditions, and the difference between SIM and other categories was more prominent in the MT paradigm compared to the two eye-tracking paradigms. Post-hoc analyses revealed that the mouse-tracing paradigm took significantly longer than the two eye-tracking paradigms for all decision categories (ps < .01), however, the two eye-tracking paradigms did not differ from one another. This replicates our previous findings (Franco-Watkins & Johnson, 2011) where the mouse-tracing paradigm resulted in longer times.

Process Measures. For comparisons with prior work, we examined summary process measures such as fixations (total number of AOI acquisitions), average AOI duration (time per acquisition), the proportion of AOIs acquired, AOI reacquisition rate, and search index (described in Payne et al., 1988); Table 1 presents these measures. To simplify presentation, we focus on the methodological comparison central to the current work by presenting main effects of paradigm (collapsing decision categories) and planned contrasts between the three paradigms (DMW vs. ET, DMW vs. MT, and ET vs. MT). We reserve additional statistics and accompanying graphs separated by decision task category for the Appendix.

Fixations. We computed an AOI acquisition (fixation) from the raw eye-tracking or cursor-placement data generated in the experiment by first removing out-of-area (non-AOI) acquisitions and then sequencing the eye or cursor placement to each AOI associated with the gambles, from the onset of eye movement or cursor placement to a single AOI until the eye movement or cursor placement was displaced from the given AOI. Fixations to the center cross or labels were included in sequencing but are not included in the reported statistics. The fewer number of fixations observed in the DMW was significantly different from both the ET, t (50) = 3.30, p <. 01, d = 0.92, and MT, t (50) = 2.15, p =.04, d = 0.60, paradigms. The ET and MT paradigms did not differ from one another (*n.s.*). Focusing only on fixations alone might not indicate why such differences were observed; thus, we also examine how the paradigms differ in terms of fixation duration, as well as proportions of cells accessed and reacquired.

Duration. The average fixation duration, in milliseconds, is based on AOI fixations associated with gamble information. An overall pattern emerged where the longest fixation durations were observed in the MT, followed by the DMW and then ET paradigms. The MT paradigm was different from both the DMW, t (50) = 2.71 p <. 01, d = 0.75, and ET, t (50) = 9.52, p <. 01, d = 2.64, paradigms. Additionally, a significant difference

https://doi.org/10.1017/S1930297500004174 Published online by Cambridge University Press

Table 1: Comparison across paradigms: Average (standard deviation) aggregate AOI processing variables.

AOI Processing Variables	Moving window	Eye tracking	Mouse tracing	F	η^2_p
AOI fixations	12.98 (3.35)[ac]	18.25 (7.43)[a]	16.02 (6.38)[c]	5.09*	.12
Avg. AOI duration, in *ms*	474.56 (99.54)[ac]	271.47 (36.92)[ab]	573.65 (157.53)[bc]	51.29*	.58
Proportion of AOIs acquired	0.987 (.025)[a]	0.912 (.166)[ab]	.996 (.007)[b]	5.85*	.14
AOI reacquisition rate	0.312 (.139)[a]	0.492 (.171)[ab]	0.377 (.185)[b]	7.77*	.17
Search index	0.724 (.131)	0.775 (.143)	0799 (.106)	2.35	.06

Note: Within each row, indices denote: [a] statistically significant comparison between DMW and ET, [b] statistically significant comparison between ET and MT, and [c] statistically significant comparison between DMW & ET. F-ratios represent main effect of paradigm: df = (2, 75 and * denotes p-values < 0.05.

was present between the two eye-tracking-based methods, t (50) = 9.75, p <. 01, d = 2.71. The fact that we observed longer fixation durations in DMW compared to ET can shed light as to why there are differences in total number of fixations. Perhaps spending more time per cell allows users more time to consider the information in a given cell and therefore they are less apt to have to visit the cells again. We next turn to an analysis of acquisition and reacquisition rates to explore this possibility.

AOI Acquisition and Reacquisition Rate. Recall that there were a total of eight AOIs that pertain to gamble information. We calculated the proportion out of the total eight AOIs acquired on each trial, as well as the reacquisition of AOIs. Although, on average, subjects examined almost all the cells, the ET paradigm resulted in a slightly lower acquisition rate than the MT paradigm, t (50) = 2.57, p =.02, d = 0.71 and the DMW paradigm, t (50) = 2.27, p =.03, d = 0.63. This is most likely due to the fact that in the ET paradigm all information is visible to the subjects such that they can acquire additional information via peripheral vision. It could also be clearer in this condition, relative to when information is occluded, that the probabilities sum to one in each gamble, rendering acquisition of the complementary probability redundant.

Additionally, the ET paradigm produced more reacquisition of AOIs (resampled proportion of total fixations) than the MT paradigm, t (50) = 2.32, p =.02, d = 1.15 and the DMW paradigm, t (50) = 4.15, p < .01, d = 0.64. As suggested earlier, although people may spend more time on average on each AOI in the DMW (as indicated by the average fixation durations), once they gather the information, they are less apt to reacquire information. Furthermore, this may also qualify the result concerning total fixations, where the DMW showed significantly fewer fixations than ET—it seems this difference can be attributed to increased reacquisitions in the ET paradigm.

It is also interesting to consider the differences above relative to previous work employing all three methods in a probabilistic inference task (Franco-Watkins & Johnson, 2011). The number of fixations observed by Franco-Watkins and Johnson (2011) showed the DMW was more comparable to ET, with MT resulting in fewer fixations. Across both eye-tracking paradigms (ET and DMW), our risky pairwise choices resulted in far fewer (less than half) fixations than the previous study's probabilistic inference among three options with four binary attributes, while the number of fixations for MT was relatively constant across tasks. Franco-Watkins and Johnson (2011) also found much larger reacquisition rates for the eye-tracking methods (DMW and ET), whereas the overall pattern of fixation durations is consistent with the current work (although the absolute magnitudes were somewhat shorter in the previous work).

The lower number of fixations—primarily due to a decrease in reacquisitions—for the eye-tracking paradigms (ET and DMW) in our experiment might be simply due to task differences between the current work and Franco-Watkins and Johnson (2011). Namely, the lower demands associated with a preferential choice, especially given the redundancy in probability information of the current task due to complementarity, might not require as many cell fixations compared to an inferential task. Furthermore, perhaps searching for an objectively correct answer in an inferential task promotes more "double-checking" of information when search costs are low (i.e., when using eye-tracking methods). However, it appears that MT might not be as sensitive to attentional processing, and/or may not allow for an increase in low-cost reacquisitions for verification, given the consistency of fixations across different decision task demands. These potential explanations should be subject to more rigorous empirical verification in future work.

Attentional dynamics. Beyond the summary statistics presented above, sophisticated analyses of attentional dynamics are being explored in eye-tracking research (e.g., Hacisalihzade, Stark, &Allen, 1992; see Day, 2010, for

https://doi.org/10.1017/S1930297500004174 Published online by Cambridge University Press

Table 2: Transition matrices by paradigm.

Moving window

	p(A1)	A2	p(A2)	B1	p(B1)	B2	p(B2)
A1	12.8%	7.5%	2.7%	1.1%	0.3%	1.1%	1.1%
p(A1)		2.6%	0.3%	0.3%	1.4%	2.3%	0.3%
A2			15.5%	0.6%	1.8%	1.7%	0.8%
p(A2)				1.5%	0.2%	1.4%	0.2%
B1					14.9%	7.2%	1.8%
p(B1)						2.3%	1.1%
B2							15.5%

Eye tracking

	p(A1)	A2	p(A2)	B1	p(B1)	B2	p(B2)
A1	13.1%	9.6%	2.1%	0.7%	0.5%	0.7%	0.5%
p(A1)		2.0%	3.1%	0.8%	3.4%	0.7%	0.6%
A2			15.1%	0.5%	0.1%	1.2%	1.3%
p(A2)				0.5%	0.0%	0.7%	3.3%
B1					11.5%	8.2%	1.3%
p(B1)						2.0%	2.1%
B2							14.3%

Mouse tracing

	p(A1)	A2	p(A2)	B1	p(B1)	B2	p(B2)
A1	15.5%	14.8%	0.0%	1.2%	0.0%	0.2%	0.0%
p(A1)		0.5%	0.0%	0.2%	4.9%	0.2%	0.5%
A2			14.6%	0.7%	0.2%	0.9%	0.2%
p(A2)				0.0%	0.7%	0.2%	6.3%
B1					12.7%	10.9%	0.0%
p(B1)						0.7%	0.2%
B2							13.7%

Notes. Row and column headers give AOI labels. Values indicate relative frequency of total transitions, averaged across subjects between the associated row and column header AOIs. For example, in the "Moving window" condition, 7.5% of total transitions occurred between outcome A1 and outcome A2. Values are aggregated regardless of transition direction (i.e., across the main diagonal; transition from A1 to A2 is aggregated with transitions from A2 to A1). Greater depth of shading corresponds to greater proportions. Dashed outlines indicate cells contributing to "alternative-wise" shifts in attention among attributes within an option. Matrices sum to approximately 100% (due to rounding).

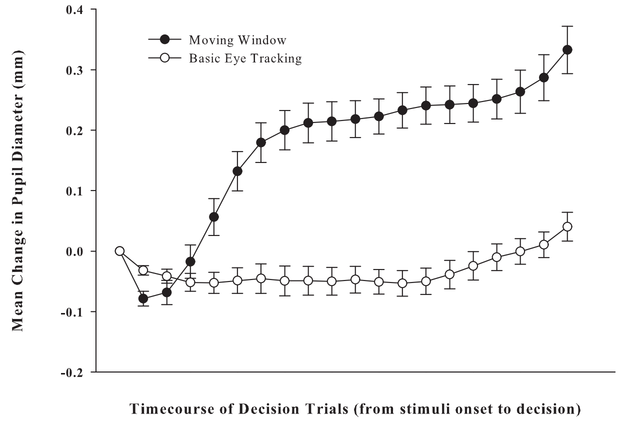
a related application to decision making). Many additional comparisons are possible such as the analysis of specific transition frequencies, although they are rarely considered in decision research (see Johnson, Schulte-Mecklenbeck, & Willemsen, 2008; and Ball, 1997 for notable exceptions). In particular, we advocate the novel use of the entire *transition matrix* (Table 2) in decision research to present the dynamic information lost in summary statistics. For each subject, we calculated the percentage of transitions between each AOI (row) and every other AOI (column); these percentages were averaged across subjects to produce Table 2. It is clear that the vast majority of transitions were between an outcome value and its associated probability, which is consistent with previous results and supports previous interpretations of an integrative process over a lexicographic heuristic (from Glöckner &Herbold, 2011). Moreover, the transition proportions were very similar across methods (in z-tests on proportions, all p-values > 0.50). If desired, the typical (but coarser) indices such as the search index developed by Payne et al. (1988) can easily be calculated using the information in Table 2—this crude search index is reported in Table 1 as well to ease comparison with previous research.

More importantly, with the complete transition matrix given in Table 2, one can also derive more precise, theoretically-driven metrics that are not possible with the search index. Researchers can develop exact predictions for the transition matrix based on decision making theories, and then compare these predictions with the empirical matrices (e.g., presented in Table 2). For example, many different lexicographic strategies might suggest "attribute-wise" processing and a search index value less than one—and perhaps even the same value—making them difficult to differentiate. However, these strategic variations could be captured in the transition matrix by highlighting *exactly* which information is predicted to receive attention, and in which order. Johnson and Koop (2010) have also used this approach by training subjects to use specific, popular strategies in order to develop empirical estimates of their application error. Furthermore, the raw frequencies used to create Table 2 could be used for more sophisticated (e.g., Markov, sequential) analyses which are not easily obtainable with summary mouse-tracing data, such as the order of search dependencies, lag, homogeneity, or stationarity (see Gottman & Roy, 1990; Stark & Ellis, 1981).

Pupil Diameter and Processing. Another measurement made uniquely possible by the eye-tracking methods ET and DMW—and underutilized in decision research—is the task-evoked pupillary responses (i.e., pupil dilations). These faithfully reflect variations in processing load between qualitatively different cognitive tasks (e.g., short-term memory, language processing, perception, and reasoning: Beatty, 1982). Changes in pupil dilation can reflect mental effort involved in a task (Beatty & Kahneman, 1966)—as the task increases in difficulty, so does the amplitude of the pupil dilation (Beatty & Wagoner, 1978; Kahneman, Beatty, Pollack, 1967). In Figure 4, we present an illustration of the richness of using pupillary measures for the DMW and ET paradigms by plotting changes in pupil dilation over the course of a trial as a finer-grain analysis to better capture pupillary changes across the decision process. Because each individual had different trial lengths depending on when they executed their decision, we standardized the time course of each trial of each subject by dividing the total trial into 20 approximately equally-sized bins. In each bin, we averaged all measurements from both pupils for each subject. By expressing these pupillary measures relative to the baseline established at trial onset (first bin), we see the increase/decrease in pupillary diameter over the course of each trial. These trajectories were averaged across trials and then subjects to produce Figure 4. Note that pupil diameter reliably increases throughout the trial and amplitudes towards the end of the trial when the subject is ready to execute a response for the DMW paradigm. The pattern for pupillary changes is more constant across trial for the ET paradigm. A 2 (paradigm: ET vs. DMW) x 20 (trial increments) mixed factorial ANOVA revealed that the differences between paradigms across the trial produced a significant interaction, F (19, 950) = 30.99, $p <$ .01, $\eta^2{}_p = .38$. Additionally the main effects of paradigm and trial increment were statistically significant, F (1, 50) = 43.62, $p < .01$, $\eta^2{}_p = .47$ and F (19, 950) = 39.89, $p <$ .01, $\eta^2{}_p = .44$, respectively. Although the changes in luminance (e.g., shifts from dark boxes to grey information cells) might slightly affect pupillary measures, we found reliable changes in pupil diameter across time even in a pilot condition using seven subjects, controlling for luminance (using grey boxes instead of the black boxes shown in Figure 1).

The richness of this pupil information can potentially aid in the development of new dynamic measures of attention—beyond simply location-based measures—that capture the processing sequence. In addition to pupillary measures, Horstmann et al., (2009) found reliable differences using fixation durations as an index of effort processing between automatic and deliberate decision making in a more complex task. Thus, eye-tracking variables provide additional mechanisms for exploring the level of cognitive effort associated with decision processes that cannot be observed by examining only outcome data. Furthermore, eye-tracking measures can be used to examine individual differences (e.g., working memory, selective attention) in dynamic processes during decision making.

Figure 4: Changes in pupil diameter during decision process for eye-tracking paradigms.



4 Discussion

We concur with the general arguments of others for the benefit, if not necessity, in using process measures for decision research, and believe our new DMW paradigm can contribute to such efforts. Researchers should be aware of how differences in process measures might arise, depending on the paradigm. The DMW produced fewer fixations and reacquisitions than basic ET; however, fixation duration was longer than basic ET. This pattern indicates differences at the process level between occlusion and non-occlusion methods in eye-tracking paradigms. The DMW also shows fewer fixations, of shorter duration, than MT, showing effects also of the user interface, holding constant occlusion of information. Moving beyond these summary statistics, however, our novel (to JDM research) analyses of the complete transition matrix suggest no differences across paradigms in terms of the specific information acquisition streams. Just as others have found (Glöckner & Betsch, 2008b), our results do suggest that different paradigms might even influence resulting choice behavior, regardless of similarities or differences in information search. For example, choice of the risker gamble in a pair seemed to increase for the DMW relative to MT, and to some degree also compared to basic ET. Thus, examining both choice level and process level measures are important to fully understand how the different methodologies might affect subsequent behavior. Moving forward, there may be a need to build new metrics that can begin to better elucidate the differences between paradigms.

An intriguing possibility for applying and extending our results is to develop metrics of covert attention that move beyond the "correspondence assumption" that equates visual attention with exclusive processing. We contend that the benefits of using eye-tracking methods allow for better tests of theoretical and/or model assumptions, especially when attention during the decision process is factored into theories and models. Given the established link between pupillary response and mental effort (Beatty & Kahneman, 1966), one could develop a graded attention metric that incorporates this important pupillary information. Furthermore, results such as the apparent increase in effort over the course of a trial in the DMW but not ET could reveal systematic effects of working memory on decision making. The DMW uniquely allows the researcher to manipulate working memory load (i.e., number of cells, items to hold in working memory, etc.) relative to when all information is available. We propose the DMW can have some additional benefits over ET. Specifically, the DMW paradigm allows greater internal validity by removing peripheral acquisition; the ability to guide information search and/or strategies by only revealing specific cells (e.g., to test assumptions regarding framing and/or order effects); and the ability to control the latency before the mask is removed to insure fixations are meaningful.

There are always potential concerns or challenges associated with any new method. For one, implementing the DMW requires additional programming than the typical MT and ET paradigms to employ an interactive gaze-contingent methodology. The DMW has been examined in two decision making tasks (risky choice and probabilistic inference) and has not been extended to other decision domains, thus, it is uncertain whether potential paradigm by task interactions will be observed in other types of decisions. The increase in internal validity comes at a cost of decreased external validity: namely, the inability to use peripheral information can be construed as a less natural accrual of information. However, the use of a template such as in the current task (Figure 1), with identified information locations, allows one to use peripheral vision to plan subsequent eye movements, while the mask allows one to retain a stronger inference regarding current visual attention and information processing. Lastly, the use of eye-tracking paradigms requires more resources in that the eye-tracking apparatuses are costly and limit data collection to one subject at a time in comparison to MT where several subjects can be collected simultaneously (limited to the number of computers). In terms of eye-tracking paradigms in comparison to MT, the most notable advantage is a more natural interface when using one's eyes to reveal and process information with DMW, which allows freedom from the psychological tether of the mouse. Others have shown that requiring additional physical exertion (over eye-tracking)—even simply head movements—can affect strategies and behavior (Ballard, Hayhoe, & Pelz, 1995), and that the requirement of moving the mouse to reveal information increases errors and slows learning (Gray & Fu, 2004). In prior work comparing all three methods in a probabilistic inference task, we noted that MT, but neither eye-tracking paradigm, was susceptible to significant variability over the course of an experiment, perhaps due to fatigue (Franco-Watkins & Johnson, 2011). Given that all the paradigms have advantages and disadvantages, and can even produce differences in choice and process measures, one should take into consideration all aspects associated with a paradigm when selecting a method to test a specific theory.

We presented the DMW paradigm as a complementary integration of two successful paradigms (basic eye-tracking and mouse-tracing methods) important for understanding decision processes. All three methods are valuable to decision researchers and which method is applicable will depend on the theoretical aspects of the decision theory being tested. The advances put forth herein add to the measurement and inferential tools we use to accompany the sophistication of modern process-based theories of decision making.

References

Ball, C. (1997). A comparison of single-step and multiple-step transition analyses of multiattribute decision strategies. *Organizational Behavior and Human Decision Processes, 69,* 195–204.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7,* 66–80.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91,* 276–292.

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5,* 371–372.

Beatty, J. & Wagoner, B. L. (1978). Pupillometric signs of brain activation vary with level of cognitive processing. *Science, 199,* 1216–1218.

Day, R-F. (2010). Examining the validity of the Needleman-Wunsch algorithm in identifying decision strategy with eye-movement data. *Decision Support Systems, 49,* 396–403.

Franco-Watkins, A. M., & Johnson, J. G. (2011). Decision moving window: Using eye tracking to examine decision processes. *Behavior Research Methods, 43,* 853–863.

Glöckner, A., & Betsch, T. (2008a). Do people make decisions under risk based on ignorance? An empirical test of the priority heuristic against cumulative prospect theory. *Organizational Behavior and Human Decision Processes, 107,* 75–95.

Glöckner, A., & Betsch, T. (2008b). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1055–1075.

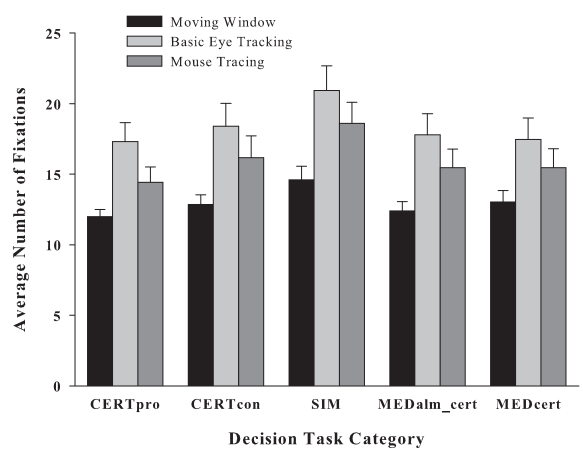
Glöckner, A., & Herbold, A. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making, 24,* 71–98.

Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis: A guide for behavioral researchers.* Cambridge:

Cambridge University Press.
Gray, W. D., & Fu, W. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science, 28,* 359–382.
Hacisalihzade, S. S., Stark, L. W., & Allen, J. S. (1992). Visual perception and sequences of eye movement fixations: A stochastic modeling approach. *IEEE Transactions on Systems, Man, and Cybernetics,* 22, 474–481.
Hoffman, J. E. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention* (pp. 119–153). East Sussex, United Kingdom: Psychology Press.
Horstmann, N., Ahlgrimm, A., & Glöckner, A. (2009). How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgment and Decision Making, 4, 335–354.*
Johnson, J. G., & Koop, G. J. (2010). Testing critical assumptions of mouse-based process-tracing. Unpublished manuscript.
Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. J. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review, 115,* 263–272.
Kahneman, D., Beatty, J. & Pollack, I. (1967). Perceptual deficit during mental task. *Science, 157,* 218–219.
Koop, G. J., & Johnson, J. G. (2011). Continuous process tracing and the Iowa Gambling Task: Extending response dynamics to multialternative choice. *Judgment and Decision Making, 6*, 750–758.
Lohse, G. E., & Johnson. E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes, 68,* 28–43.
Norman, E., & Schulte-Mecklenbeck, M. (2010). Take a quick click at that! Mouselab and eye-tracking as tools to measure intuition. In A. Glöckner and C. Witteman's (Eds.), *Foundations for tracing intuition: Challenges and methods* (pp. 24 - 44). New York: Psychology Press.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534–552.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.
Ranyard, R., & Svenson, O. (2011). Verbal data and decision process analysis. In M. Schulte-Mecklenbeck, A. Kuehberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research* (pp. 115–137). New York, NY: Psychology Press.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Review, 124,* 372–422.
Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7,* 219–226.
Russo, J. E., & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition, 3,* 267–276.
Schulte-Mecklenbeck, M., Kuehberger, A., & Ranyard, R. (2010). *A handbook of process tracing methods for decision research.* New York, NY: Psychology Press.
Stark, L., & Ellis, S. R. (1981). Scanpaths revisited: Cognitive hypothesis direct active looking. In J. W. Senders, D. Fisher, and R. A. Monty (Eds.), *Eye Movements and Psychological Processes* (pp. 192–226). Hillsdale, NJ, Lawrence Earlbaum Associates.
Wedel, M., & Pieters, R., (2000). Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science, 19,* 297–312.

Appendix

Figure A1. Average (standard error) fixations per decision category across paradigms.



Fixations. Figure A1 presents the fixation data by paradigm and decision task. In addition to the main effect of condition presented in Table 1 of the manuscript, The mixed factorial ANOVA also revealed a main effect of decision task category, F (4, 300) = 19.48, $p < .01$, $\eta^2_{p} = .21$, however, the interaction effect was not statistically significant (p =.71). The SIM category overall resulted in a greater number of fixations than the other categories, replicating Glöckner and Herbold (2011). Post-hoc analyses revealed differences between the two eye-tracking paradigms for all categories. None of the other post-hoc comparisons were statistically significant.

https://doi.org/10.1017/S1930297500004174 Published online by Cambridge University Press

Figure A2. Average (standard error) fixation duration per decision category across paradigms.

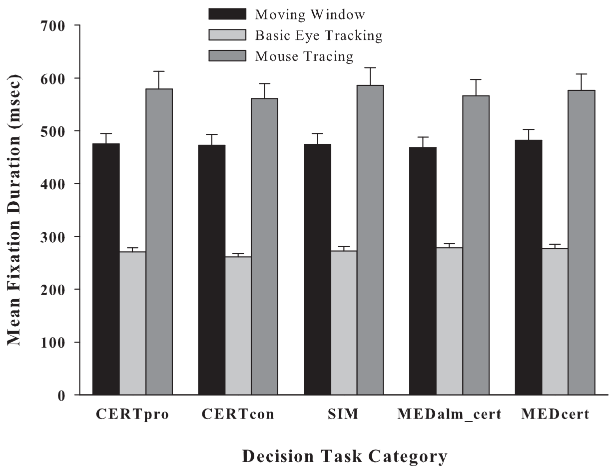
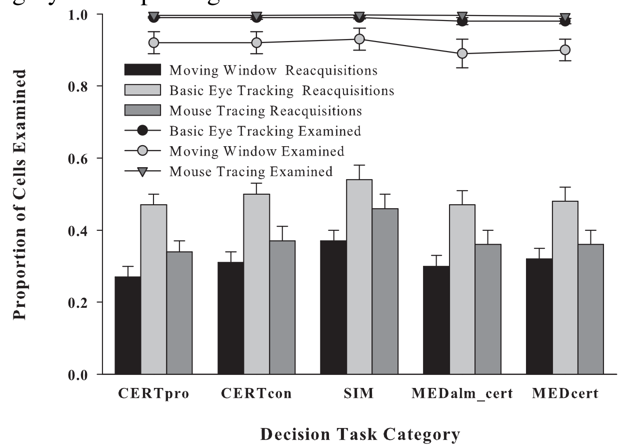


Figure A3. Proportion (standard error) of AOIs examined (lines) and reacquisitions (bars) per decision category across paradigms



Duration. Figure A2 presents the average fixation duration by paradigm and decision task. The mixed factorial ANOVA revealed a main effect of decision category, F (4, 300) = 2.85, p =.02, $\eta^2{}_{p\,=}$ .04. The interaction of paradigm by category was not significant (p = .27). Post-hoc analyses revealed differences between all three paradigms for all categories.

AOIs acquired and reacquisition rate. Figure A3 presents the average proportion of cells examined (lines) and reacquisitions (bars) per paradigm and decision task. Two mixed factorial ANOVAs were conducted, one per dependent measure. In terms of AOIs examined, a main effect of decision category was present, F (4, 300) = 5.66, p <.01, $\eta^2{}_{p\,=}$ .07 and an interaction of paradigm by category, F (8, 300) = 2.34, p =.02, $\eta^2{}_{p\,=}$ .06. Post-hoc analyses revealed differences between ET compared to MT and DMW for all categories, except for the SIM category between the two eye-tracking paradigms. These differences are most likely due to the fact that the ET paradigm can rely on peripheral vision; direct AOI examination is not necessary unlike the DMW and MT paradigms.

In terms of AOI reacquisitions, the main effect of decision category was significant, F (8, 300) = 20.90, p <.01, $\eta^2{}_{p\,=}$ .22. Consistent with the other process measures, the SIM category has more reacquisitions which are indicative of the difficulty of making choices between two similar gambles. Post-hoc analyses indicated that the ET differed from DMW (all categories) and MT ($\mathrm{CERT_{PRO}}$,$\mathrm{CERT_{CON}}$, and $\mathrm{MED_{CERT}}$). Coupled with the fixation duration data above, this pattern suggests that subjects with occluded information required more time per AOI and therefore might not need to return to the AOI information as often as the basic ET, where subjects might initially (quickly) scan the AOIs and then revisit to potentially solidify information gathering.

In sum, the decision categories did affect the process variables such that the categories that appeared to be more difficult (e.g., SIM) resulted in more fixations, fixation durations, and reacquisitions than decision categories that favored one outcome with greater certainty. The global differences between paradigms observed in the manuscript were replicated at the decision category level.