ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 68, No. 1, October, pp. 28–43, 1996
ARTICLE NO. 0087

A Comparison of Two Process Tracing Methods for Choice Tasks

GERALD L. LOHSE

Operations and Information Management Department, The Wharton School of the University of Pennsylvania

AND

ERIC J. JOHNSON

Marketing Department, The Wharton School of the University of Pennsylvania

Process tracing methods, particularly those based on information acquisition, are becoming commonplace. Because of this, it is important to examine both the reactivity and the validity of these techniques. This research compares information acquisition behavior for choice tasks using Mouselab, a computerized process tracing tool, and Eyegaze, an eye tracking system. In an experiment using apartment selection tasks and gambles, we found significant differences contingent upon the process tracing method for 10 process tracing measures including subsequent choices. Computerized process tracing tools increase the amount of time needed to acquire information compared with eye tracking equipment. As a result, subjects using Mouselab tend to have more systematic information acquisition behavior than that observed with eye tracking equipment. Additional research is needed to explore the magnitude and consequences of these differences.

INTRODUCTION

The information acquisition processes underlying judgment and choice have seen greater interest over the past 20 years. Information acquisition includes an analysis of the content of the information sought, how long the subject examines this information, the sequence of acquisition, and the amount of information acquired (Einhorn & Hogarth, 1981). These data are important for research in behavioral decision making and decision support systems for at least three reasons. First, patterns of information processing suggest certain strategies for evaluating information (Payne, 1976). Understanding these prototypical patterns for evaluating information helps identify behavior that could constrain or alter decision processes. Second, the information acquisition patterns directly influence cognition and memory. Subtle changes in presentation format can change the frequency of preference reversals (Johnson, Payne, & Bettman, 1988), change decision making strategy (Bettman & Kakkar, 1977; Jarvenpaa, 1989; Todd & Benbasat, 1991) and alter decision performance (Ashton & Ashton, 1988). Third, because the way information is displayed can change decisions, understanding these influences is important in the design of interfaces for electronic commerce (Widing & Talarzyk, 1993).

Process tracing methods often examine the information individuals seek before making a choice and how that information produces a choice. Information acquisition behavior has been studied using eye tracking equipment, information boards, and computerized process tracing tools (CPT) (Abelson & Levi, 1985). Verbal protocols have been used to study information acquisition behavior and processing stages of choice concurrently (see Ford, Schmitt, Schechtman, Hults, & Doherty (1989) for a review of 45 experiments that used either verbal protocols or information display boards). Verbal protocols and information acquisition techniques have seen widespread application in the behavioral decision-making literature (Einhorn & Hogarth, 1981) and in the decision support system literature (Todd & Benbasat, 1991).

Each process tracing method requires different levels of information acquisition effort on the part of the decision maker. Information boards present information in an envelope on a poster board containing an index card with one piece of information (e.g., rent for apartment C, see Payne, 1976). Russo found that the overt action of selecting and reading one card requires 15–20 s per acquisition (1978b). An experiment by Van Raaij (1977) had subjects reach for an actual package. By turning

We thank Peter Todd, Carol A. E. Nickerson, and two anonymous reviewers for their helpful comments and suggestions. Address correspondence and reprint requests to Gerald L. Lohse, Department of Operations and Information Management, 1319 Steinberg Hall-Dietrich Hall, The Wharton School of the University of Pennsylvania, Philadelphia, PA 19104-6366. E-mail: lohse@wharton.upenn.edu.

0749-5978/96 $18.00

the package, subjects viewed attribute information from all four sides with an information acquisition rate of 4 s per acquisition. Computerized process tracing (CPT) environments replace manual acquisition with a computer-based pointer that requires less time per acquisition depending upon the pointing device, including a mouse, light pen, or keystrokes (e.g., *Mouselab*, Payne, Bettman, & Johnson, 1993; *IS Lab*, Cook & Swain, 1993; *Search Monitor*, Brucks, 1988; and other tools Jacoby *et al.*, 1985, 1987; Payne & Braunstein, 1978; Todd & Benbasat, 1991; Williams, 1990). CPT systems typically take 1–2 s per acquisition, whereas an eye fixation typically requires 0.2–0.4 s to acquire the same piece of information (Russo, 1978a; Card, Moran, & Newell, 1983).

These differences in effort raise an interesting and perhaps fundamental question. Do various process tracing methods change the decision process? In other words, are they reactive procedures? Of course, one could argue that all processing tracing methods affect the process. The real questions are how they affect the process, whether any "true" information acquisition process can be observed, what if any effect process tracing methods have upon what is chosen and finally how potential experimental manipulations interact with these changes in process. For example, Russo, Johnson, and Stephens (1989) found that concurrent verbalization could change the underlying process and the accuracy of the decision process as well as that the nature of these could change across tasks.

Verbal protocols measure information acquisition and internal processes directly. In contrast, eye tracking equipment, information display boards, and CPT tools observe only information acquisition behavior; internal cognitive processes are not directly observable. Investigators must infer the underlying cognitive strategy from the information acquisition data. Sometimes those inferences are incorrect. For example, Payne, Braunstein, and Carroll (1978, p. 37) note that just because information is acquired does not mean that it has been processed in a certain manner. Information acquisition data showed that one subject examined all information in an interdimensional fashion consistent with an additive decision rule. However, excerpts from the concurrent verbal protocol made it clear that some information was ignored and that the number of attributes actually considered for each alternative varied. Payne *et al.* (1978) assert that "The ability of verbal protocols to detect such activity represents one of the greatest advantages of protocol over eye movement and explicit information acquisition procedures." Thus, concurrent methods (verbal protocol and eye tracking equipment) seem to help avoid misinterpretation problems.

Surprisingly, very little research has examined the influence of different information acquisition procedures upon the underlying process itself. Russo (1978b) found significant differences in information acquisition behavior between information display boards and eye movements. But did not compare the same subjects working on the same task. Van Raaij (1977) made a within-subjects comparison of eye tracking and information boards and found significant differences for all five process measures examined. More recently, personal computers spurred rapid development of CPT systems for capturing information acquisition behavior. Despite a 10-year history of information acquisition research using CPT tools, we are unaware of comparisons between eye tracking and CPT tools.

We approach a comparison between eye tracking and CPT tools from a cost-benefit perspective. Consistent with prior research (Bettman, Johnson, & Payne, 1990; Newell & Simon, 1972), costs are measured as a function of the number of elementary information processes (EIPs). In the spirit of computational models of cognition (Card *et al.*, 1983), we use estimates of the times required by each EIP and sum over the number of EIPs to predict total task time. Thus, costs include information acquisition effort and mental processing while benefits reflect the utility or accuracy of the decision.

The goal of this paper is to examine differences between two particular methods of observing information acquisition: eye movement recording (the Eyegaze System) and one particular mouse-based CPT tool (Mouselab). We do this because it illustrates how one might understand, in general, how different ways of presenting and monitoring information might change decision processes. It is also interesting to know what differences exist in this specific case. We predict that differences in processing tracing measures will be explained by differences in time required to acquire information using each method. We test the generalizability of our predictions by using two choice tasks (gambles and apartment selection) and a range of information loads (2×2, 2×7, 7×2, and 7×7 alternatives and attributes).

The next section of this paper develops the conceptual framework and the set of hypotheses that follow. We then discuss the methods and results of an experiment that compare the two process tracing methods for two choice tasks using varying amounts of information. Finally, we examine implications for information acquisition tools as well as the application of the effort-accuracy model to understand the impact of informa-

TABLE 1

Elementary EIPs Used in Decision Strategies (from Bettman, Johnson, & Payne, 1990)

Read	Read an alternative's value on a given attribute into working memory
Compare	Compare two alternatives on an attribute
Difference	Calculate the size of the difference of two alternatives for an attribute
Add	Add the values of an attribute in working memory
Product	Weight one value on an attribute in working memory
Eliminate	Remove an alternative or attribute from consideration
Move	Go to the next element of external environment
Choose	Announce preferred alternative and stop process

tion acquisition methods and computer based decision aids.

COGNITIVE EFFORT FRAMEWORK

Johnson and Payne (1985) proposed a set of EIPs to describe and measure cognitive effort for different decision strategies (Table 1). This paradigm predicts cognitive effort as a function of the number of EIPs required to execute a particular decision strategy. Total task completion time is the sum of the subcomponent times for each EIP. We adopt this framework for calculating EIP counts and task completion times to predict differences between Mouselab and Eyegaze. Bettman *et al.* (1990) show that 75% of the variance in decision times can be accounted for by a count of these mental operations, and that individual differences in the cost of these operations can help predict individual differences in the selection of decision strategies.

We illustrate the EIP counts using the gamble task from Fig. 1 with two alternatives and seven attributes. The EIPs are Read a piece of information into WM, Multiply a probability times its payoff, and Add or Compare two values in WM. Assuming a weighted adding rule, expected value calculations for gambles have two steps. First, subjects evaluate each alternative by multiplying each probability times the attribute payoff and adding those products for all attributes. For the first alternative, there would be 14 Reads, seven Products, six Additions, and no Comparisons to determine the value of gamble A (436.25). Second, subjects retain the value of the best alternative and its label in working memory (WM) while computing the value of the remaining alternatives. After processing the second alternative, there would be a Comparison of the values for gamble A and gamble B. Hence, there is a total of 28 Reads, 14 Products, 12 Additions, and one Comparison. Table 2 shows the EIP counts for other gamble tasks.

We posit that much of the effect of a process tracing method is explained by differences in the time and effort required to do certain operations. While EIP counts are irrespective of the processing tracing method, differences between the two process tracing tools depend on the amount of time required to acquire information. Mental computations for multiplication, addition and comparison are identical regardless of the process tracing method. After Bettman *et al.* (1990), who review previous estimates and derived EIP values empirically, we use 0.84 s as an average value for all mental operations. The only difference between the two process tracing methods is the amount of time required to read a piece of information into WM.

Visual search for reading or viewing a picture involves foveal vision, peripheral vision, eye movements and head movements. Foveal vision is for viewing detail over a narrow region (about two degrees) called the fovea. Peripheral vision lacks the fine detail of foveal vision. Saccades are the most common form of eye movement from 10 min of visual angle to a 30° angle (Schiffman, 1976). Saccadic eye movements occur during reading and the viewing of stationary scenes. Saccades are a rapid, abrupt jump to a new point of regard taking about 0.03 s and dwelling there 0.160 to 0.400 s. Beyond 30°, head movements occur to reduce the visual angle.

Pointing at objects using a mouse requires eye movements *and* motor movements. Card *et al.* (1983) use Fitts Law (Fitts, 1954) to describe pointing times as a function of distance and size of target. They found that pointing with a mouse at objects whose distances and target sizes varied requires an average 1.1 s per selection. They also empirically derived the constants in

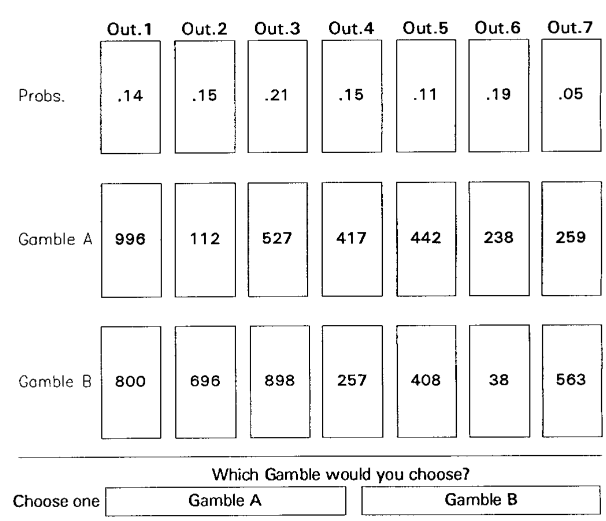

FIG. 1. Gamble with two alternatives and seven attributes.

TABLE 2

Weighted Additive Rule Used to Predict the Number of EIPs by Type to Complete the Gamble Task with a Comparison of Predicted Times and Actual Times for Gambles

					Eyegaze time in seconds	Mouselab time in seconds
Gamble alternatives	2	2	7	7		
Gamble attributes	2	7	2	7		
Gamble probabilities	2	7	2	7		
Reads (probabilities + attributes) × alternatives	8	28	28	98	0.23	1.19
Products (attributes × alternatives)	4	14	14	49	0.84	0.84
Additions (alternatives × (attributes − 1))	2	12	7	42	0.84	0.84
Comparisons (alternatives − 1)	1	1	6	6	0.84	0.84
Predicted total “ideal” time for Eyegaze	7.7	29.1	29.1	104.0		
Predicted total “ideal” time for Mouselab	15.4	56.0	56.0	198.1		
Predicted ratio Mouselab:Eyegaze	1.99	1.92	1.92	1.90		
Actual average time for Eyegaze	32.9	57.4	36.6	90.7		
Actual average time for Mouselab	46.5	103.3	71.5	178.2		
Actual ratio Mouselab:Eyegaze	1.41	1.80	1.95	1.96		

Fitts Law [time in seconds = 1.03 s + .096 $\log_2$(distance/size + 0.5)]. There is a large constant time of 1.03 seconds to begin a move, regardless of the distance or size of the target. The constant represents the time required to adjust the grasp on the mouse and begin movement. It assumes the hand is already on the mouse. Once the mouse is in motion, the constant is no longer applied. It requires an additional 0.36 s to move the hand from a keyboard to a pointing device or vice versa and a button press would require an additional 0.2 s (Card *et al.*, 1993). However, since subjects always hold the mouse and Mouselab does not require a button press to select information, we do not include either time.

Bettman *et al.* (1990) empirically derive a value of 1.19 s for the Read EIP that includes moving a mouse, making an eye fixation, and reading a value into WM. Independently, the time required for each Read EIP component is 1.1, 0.23, and 0.3 s, respectively, for a total of 1.63 s (Card *et al.*, 1983). Of course, these activities may well occur in parallel, therefore the 1.19 s Read EIP value seems appropriate for most situations. Thus, the Read EIP values of 1.19 and .23 s using Mouselab and Eyegaze are completely comparable values for information acquisition and reading a value into WM.

HYPOTHESES

Total time. The total time required to make a choice is one measure of cognitive effort. Total time to complete a task is a function of fixation duration and the number of fixations. Table 2 shows the predicted total time to complete the choice tasks using Mouselab and the Eyegaze System. Times range from 7.7 s for the two alternative, two outcome gamble to 198.1 seconds for the seven alternative, seven outcome gamble task.

Time predictions reported by Card *et al.* (1983) for cognitive models of other problems represent the performance of skilled, error-free, experts. Only an ideal subject with error-free behavior using a pure weighted adding rule would process the information as shown in Table 2. It is not uncommon to find longer times from empirical data for real users. Thus, we use the EIP predictive framework to estimate the relative time difference between the two methods. Specifically, we predict that Eyegaze will be nearly twice as fast as Mouselab. *Hypothesis 1: The time to complete a choice task with Mouselab will be about twice that of Eyegaze.*

Read EIP. The Read elementary information process involves reading an alternative’s value on an attribute into WM. The second hypothesis merely compares the empirically derived values of 0.23 s for Eyegaze and 1.19 s for Mouselab. *Hypothesis 2: The mean time per fixation for Eyegaze is less than for Mouselab.*

Total number of fixations. An ideal subject with error-free behavior would have an identical number of fixations using either Eyegaze or Mouselab. However, the number of fixations does not include fixations to support the intermediate mental calculations for “Products,” “Additions,” and “Comparisons.” Because of the low effort of acquiring information, it is likely that subjects using Eyegaze will perform arithmetic calculations using the values in the display as a visual aid for intermediate calculations and comparisons. Eye fixations for intermediate calculations would increase the total number of fixations for subjects using Eyegaze. Because it is impossible to keep two boxes open simultaneously using Mouselab, the additional time reacquiring a piece of information using a mouse suggests that subjects using Mouselab would not make as many additional fixations to support mental arithmetic and comparisons. *Hypothesis 3: The total number of fixations will be greater for Eyegaze than for Mouselab.*

Accuracy. The EIP predictions for total time assume errorless performance. A vast literature acknowledges WM as a bottleneck in human information processing and a potential locus for error (Miller, 1956; Ericsson, Chase, & Faloon, 1980; Ericsson & Kintsch, 1995; Simon, 1974). Mental arithmetic such as that used in calculating the expected value of a gamble involves well-learned procedures, problem solving skills, and reliance on WM. Most adults use associative memory to access simple products like 3×4 directly (Campbell & Graham, 1985; Dehaene, 1992; McCloskey, Harley, & Sokol, 1991). For more complex calculations (116×996), some mental arithmetic errors occur because subjects fail to retain an accurate record of carries and intermediate calculations in WM (Hitch, 1978).

The EIPs from Bettman *et al.* (1990) do not contain operators to model the detailed memory and bookkeeping operations. These would include operators to keep track of where one is in the process, remembering items of information, modeling the contents of WM and when, if at all, WM capacity is exceeded. There are two reasons we expect bookkeeping and mental operations to be higher for Mouselab than for Eyegaze. First, concurrent arm movement required by Mouselab disrupts the retention of spatial information while concurrent visual input does not (Logie, 1986; Quinn & Ralston, 1986; Smyth & Pendleton, 1989). Second, longer information acquisition times using Mouselab increase the total time information must be retained in WM. With sufficient levels of activation, WM retains information for about 7 s (Card *et al.*, 1983). Given both of these reasons, the mental effort involved in using a mouse is slightly greater than that of using Eyegaze.

While the apartment selection task does not have a known optimal choice, there is an optimal choice for gambles. Further, the expected value calculations for gambles place a very high load on WM. Because of that, we expect that Mouselab will cause more slips, forgetting, and errors in the arithmetic calculations. As a result, performance will be less accurate with Mouselab than with Eyegaze.

We have two measures of accuracy in gamble selection: (1) expected value (EV) and (2) a relative EV ratio (Johnson & Payne, 1985). Expected value is a special case of expected utility that combines values and beliefs. EV Ratio, compares the expected values of the selected, non-optimal gambles to those of the optimal gambles. EV ratio is the expected value of the chosen gamble divided by the optimum expected value for that gamble task. *Hypothesis 4: Subjects will be more accurate in gamble selection using Eyegaze than using Mouselab.*

Percent information searched. The mean proportion of information searched is the number of cells examined divided by the total number of cells. It is well established that the proportion of information searched in a choice process decreases as the number of pieces of information increases (alternatives $\times$ attributes) (Svenson, 1979; Ford *et al.*, 1989). However, since the effort associated with eye-tracking is less, we expect more information to be searched with Eyegaze than with Mouselab. *Hypothesis 5: The percent information searched will be higher with Eyegaze than with Mouselab.*

Search pattern. The search pattern index (Cook & Swain, 1993; Payne & Braunstein, 1978; Schkade & Johnson, 1989), compares the number of within-attribute transitions to the number of within-alternative transitions as follows: (total number of within-alternative transitions minus the total number of within-attribute transitions)/(total number of within-alternative transitions plus the total number of within-attribute transitions). Within-attribute transitions are instances in which the nth + 1 item searched is of the same attribute as the nth. Within-alternative transitions are instances in which the nth + 1 item searched is of the same alternative as the nth. Note that the index is independent of the number of transitions occurring with a change of both alternative and attribute (a diagonal move). A score of 1.0 represents a strict alternative-based search while a score of -1.0 represents a strict attribute-based search.

While the search index measure has raised some controversy (Böckenholt & Hynan, 1994; Payne & Bettman, 1994), it is relatively easy to understand and interpret. In general studies have found that as the number of alternatives increases, there are more intradimensional search processes (Cook & Swain, 1993; Payne & Braunstein, 1978; Schkade & Johnson, 1989), but one study (Stone & Schkade, 1994) found the opposite results using attributes that were scaled differently.

Assuming an English reading order of left to right and top to bottom, subjects viewing a matrix screen display tend to read row-wise (Tullis, 1988; Galitz, 1985) which is characteristic of an interdimensional search. While this may be a common search pattern for both methods, we expect subjects using Eyegaze to exhibit more intradimensional (within-attribute) search for two reasons. First, the predicted additional fixations to support mental arithmetic calculations will tend to be intradimensional (e.g., "Product" of gamble probability times its payoff or "Comparison" of multiple values for a particular apartment attribute). Second, we expect subjects using Eyegaze will adapt their information processing behavior more opportunistically to the demands of the data because subjects can reacquire the information at a "low cost." Thus, we expect Eyegaze should lead to more intradimensional search processes than Mouselab. *Hypothesis 6: Eyegaze should lead to more intradimensional search processes than Mouselab.*

Reacquisition rate. Classic studies of WM show a direct relationship between the amount of information retained in WM and rehearsal speed (Barsalou, 1992). Because the time required to acquire a piece of information using Eyegaze is less than that for Mouselab, we expect differences in forgetting. As reported by Russo (1978a, p. 101), we expect subjects to adopt strategies, such as memorization, to cope with increased effort per information acquisition. Memorization of needed information eliminates the need for information reacquisition but increases the WM information processing burden. Memorization is a strategy subjects tended to use with information boards (Russo, 1978b). Russo reported the reacquisition rate for information boards was only 2 to 7% of all fixations, whereas 75% of all eye fixations were reacquisitions. Van Raaij also found higher reacquisitions for eye tracking than for information boards (1977). Because of the high effort involved with reacquiring information, we expect that there will be fewer reacquisitions of information using Mouselab as compared with the Eyegaze System. We compute reacquisition rate as the number of cells viewed at least twice divided by the total number of cells viewed. Thus, the denominator reflects only actual cells viewed not the total number of cells in the problem. *Hypothesis 7: The reacquisition rate will be higher for Eyegaze than Mouselab.*

Variability in information search. In a typical recognize-act cognitive cycle, recognition is a fundamentally parallel task and the action phase is serial. Mouselab imposes strictly serial information acquisition. In contrast to a normal information processing environment, Mouselab eliminates the possibility of acquiring information from other cells in the matrix (nonlabel cells) using peripheral vision. It is more effortful for Mouselab users to scan data values to direct attention. Mouselab supports strictly serial processing. In contrast, Eyegaze involves parallel recognition and a serial action phase. Eyegaze facilitates data-driven, opportunistic, bottom-up processing. Thus, variability in percent information searched will be greater for Eyegaze than for Mouselab. *Hypothesis 8: Eyegaze will lead to higher variation in information acquisition processes than Mouselab.*

	Appear.	Distance	Rent	Safety	Kitchen	Landlord	Laundry
Apt. A	V.Clean	10 min.	$150	Average	Partial	Dffcult	Coin
Apt. B	Clean	24 min.	$200	Unsafe	None	Helpful	None
Apt. C	V.Dirty	24 min.	$550	Average	Partial	Dffcult	Free
Apt. D	Dirty	12 min.	$300	V. Safe	Full	Dffcult	None
Apt. E	Dirty	24 min.	$450	Safe	None	Helpful	Free
Apt. F	V.Clean	15 min.	$600	Safe	Partial	Dffcult	None
Apt. G	Dirty	30 min.	$300	Unsafe	None	Average	Coin

Which Apartment would you choose?

Choose	Apt A	Apt B	Apt C	Apt D	Apt E	Apt F	Apt G

FIG. 2. Apartment selection task with seven alternatives and seven attributes.

We adopt measures for variability in percent information searched from Cook and Swain (1993). Variability by alternatives measures the standard deviation of the percentage of information searched per alternative across the set of available alternatives. Variability by attributes measures the standard deviation of the percentage of search per attribute across the set of available attributes. In the standard deviation formula, n is the number of total attributes and x_i is the percentage of attribute (alternative) i searched. Searching the same proportion of information for each alternative (low variability of search) indicates the use of a compensatory strategy where a highly variable search pattern suggests a noncompensatory search strategy (Payne, 1976; Ford *et al.*, 1989).

METHODS

Stimuli. Gamble stimuli were modeled after those used by Payne, Bettman, and Johnson (1988) (Fig. 1). The first row of the display contained a vector of probabilities that sum to one. Each gamble was a set of payoffs in each subsequent row and one gamble maximized expected value. The order of the rows and columns was the same for all subjects.

The apartment choice task presented each apartment as a row in a matrix (Fig. 2) with two or seven attributes describing each apartment. The seven attributes selected by prior pretest were Appearance (5 levels from very dirty to very clean), Distance (in minutes from campus), Rent, Safety (5 levels from very unsafe to very safe), Kitchen (3 levels none, partial, full), Landlord (3 levels: difficult, average, helpful), and Laundry (3 levels: none, coin, free). The two attribute case used Distance and Appearance. Information about each attribute was a column in a matrix. The order of the rows and columns was the same for all subjects. No alternative dominated another on all attributes.

Subjects. Thirty-six undergraduate students with prior experience using a mouse participated in the study. Subjects also had some familiarity with expected value from their course work. All subjects were native speakers of English. In addition to a payment of $20.00 at the end of the second session, the one subject with the most correct gambles as defined by the expected value, received an additional award of $100. In the event of ties, the winner was selected on the basis of shortest cumulative time for the correct gambles.

Process tracing tools. The Eyegaze System (LC Technologies Fairfax, VA) uses the pupil-center/corneal reflection method to determine eye gaze (Young & Sheena, 1975). This method captures voluntary, saccadic eye movements that fixate a target object on the fovea, a region of high visual acuity on the retina. Saccadic eye movements have two parts: a movement phase ranging from 30 to 120 ms and a latency phase fixating from 100 to 300 ms. A typical saccade duration is 230 ms. A video camera, sensitive in the infrared range and positioned below the computer monitor, continually observes the subject's eye. Specialized image processing software generates *x*, *y* coordinates for the gaze point on the monitor screen. Other measures include: fixation duration, pupil diameter, and eye blinks. The observer's eye is about 20 inches from the screen of the computer monitor. The Eyegaze System collects data at 60 Hz or about every 16.7 milliseconds within an accuracy of 0.25 inches. The minimum fixation duration was 100 ms. The eye tracking equipment does not require attachments to the head (e.g., no bite bar or chin rest). The calibration procedure for each new subject takes less than one minute.

Mouselab monitors the information acquisition stages of decision behavior in many contexts (Payne *et al.*, 1993). Using a mouse, the user points to a box which then reveals the information behind the box. When the user moves the cursor out of the box, the information is no longer visible to the user (Fig. 3). Data include information about the time, sequence, and frequency of a user's information acquisition behavior. While these data approach the level obtained with eye movement recording, Mouselab does not capture data about scanning to row and column labels.

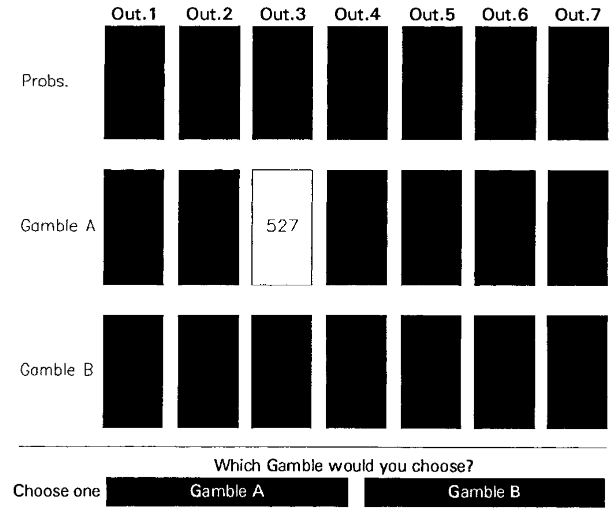

FIG. 3. Gamble with two alternatives and seven attributes using Mouselab.

Procedure. Subjects attended two sessions held 1 week apart. At one session, subjects used Mouselab; at the other, they used the Eyegaze system. Practice sessions helped subjects learn to make eye movements with minimal head movement and familiarize them with Mouselab. In the primary data collection, subjects evaluated four apartment selection problems and four gambles representing the factorial combinations of alternatives (2, 7) and attributes (2, 7). The choice tasks alternated (e.g., apartment selection, gamble, apartment selection, gamble, etc.) and the order of choice problems for each subject varied randomly. All other factors were counterbalanced across subjects.

Design. Thirty-six subjects participated in the study. There were four within-subjects factors: process tracing tool (Mouselab versus Eyegaze), two alternatives (2 or 7 items), two attributes (2 or 7 items), two tasks (apartment selection versus gamble), and one between-subjects factor, Order.

RESULTS AND DISCUSSION

Preliminary data analysis. The data collection resulted in 26,447 Eyegaze fixations and 16,992 Mouselab fixations over all 36 subjects. These fixation data exclude row and column labels since Mouselab does not capture this information. Eyegaze noted 93% of the total time as fixations on the display. Seven percent of the total time was lost due to blinks or looking off the computer screen. In contrast, Mouselab captured 71% of the total time as fixations on the display.

There are 576 observations (36 Subjects × 2 Methods × 2 Tasks × 2 Alternatives × 2 Attributes) for each of

TABLE 3a

ANOVA *F* Tests with Significance (*p* > *F*) for the Main Effects (Method, Task, Alternatives, Attributes, and Order) and One Interaction Term (Alternatives × Attributes) Using the Nonadditive Repeated Measures Model

H	Measure	Method (1, 34)	Task (1, 34)	Altern (1, 34)	Attrib (1, 34)	Alt × Att (1, 34)	Order (1, 34)
1	Total time (seconds)	36.78 (.0001)	62.79 (.0001)	101.55 (.0001)	125.27 (.0001)	30.54 (.0001)	5.18 (.0292)
2	Read EIP (seconds)	262.57 (.0001)	10.22 (.0030)	7.54 (.0096)	70.87 (.0001)	8.11 (.0074)	.46 (.5003)
3	Fixations	59.17 (.0001)	57.32 (.0001)	60.65 (.0001)	112.88 (.0001)	21.98 (.0001)	2.07 (.1597)
4a	EV	5.52 (.0248)	Gambles only	.50 (.4859)	762.29 (.0001)	.09 (.7687)	8.54 (.0061)
4b	EV ratio	4.51 (.0461)	Gambles only	4.95 (.0329)	6.82 (.0133)	5.94 (.0202)	8.37 (.0066)
5	Search index	10.36 (.0028)	9.32 (.0044)	7.04 (.0120)	21.48 (.0001)	4.35 (.0446)	.09 (.7653)
6	% Information	4.81 (.0352)	.56 (.4599)	129.83 (.0001)	80.85 (.0001)	50.02 (.0001)	.25 (.6178)
7	% Reacquisitions	46.59 (.0001)	46.09 (.0001)	81.66 (.0001)	6.94 (.0126)	.06 (.8016)	.48 (.4747)
8a	Column variation	5.07 (.0309)	1.51 (.2273)	201.03 (.0001)	81.41 (.0001)	24.34 (.0001)	1.07 (.3092)
8b	Row variation	3.66 (.0474)	2.88 (.0991)	122.43 (.0001)	84.23 (.0001)	34.16 (.0001)	2.26 (.1421)

Note. The mean square error for within-subjects factors is Subject*Order**Treatment* with 34 degrees of freedom [2(18-1)(2-1)] as an error term. The mean square error for Order, a between-subjects factor, has 34 degrees of freedom with Subject*Order [2(18-1)] as the error term.

10 dependent measures. The measures included total time, Read EIP time, number of fixations, accuracy, percentage of the total information searched, search pattern, reacquisition rate, variability by alternative, and variability by attribute. MANOVA examined these performance effects simultaneously while controlling for the multiple dependent measures. We found a significant effect for method (Wilk's λ = .081, $F(9,169) = 213.4$, $P > .0001$). Ten univariate ANOVA models examined the effects in more detail. Tables 3a and 3b show the ANOVA F test and significance levels for the main effects (Method, Task, Alternatives, Attributes, and Order) as well as for the five interaction effect (Alternatives × Attributes, Method × Task, Method × Alternatives, Method × Attributes and Method × Alternatives × Attributes). Tukey's test examined multiple comparisons among the reported means at the .05 level of significance.

Total time. Overall, subjects with Mouselab required 67% more time to complete the tasks (73.9 s versus 43.6 s). The gamble task was significantly longer than the apartment selection task (77.1 s versus 40.5 s). Tasks with seven alternatives required significantly more time than those with two alternatives (74.4 s versus 43.2 s). Tasks with seven attributes required significantly more time than those with two attributes (82.5 s versus 35.2 s). There was also a significant effect for Order (first 66.5 s versus second 51.1 s). Total task completion time decreased with practice. A significant two-way interaction for Alternative × Attribute shows that task completion time increased as task complexity increased. The apartment task with two alternatives and two attributes differed the least in the total time required to complete the task (Eyegaze 16.6 s and Mouselab 17.1 s). The gamble task with seven alternatives and seven attributes differed the most in the

TABLE 3b

ANOVA *F* Tests with Significance (*p* > *F*) for the Interaction Effects of Method Using the Nonadditive Repeated Measures Model

H_0	Measure	Method*Task	Method*Altern	Method*Attrib	Method*Alt*Att
1	Total time (seconds)	24.91 (.0001)	15.80 (.0001)	23.00 (.0001)	5.33 (.0267)
2	Read EIP (seconds)	6.85 (.0131)	6.16 (.0182)	31.51 (.0001)	8.10 (.0074)
3	Fixations	3.20 (.0825)	.21 (.6513)	12.45 (.0012)	.97 (.3314)
4a	EV	Gambles only	4.26 (.0467)	5.26 (.0281)	.05 (.8239)
4b	EV ratio	Gambles only	4.65 (.0382)	3.66 (.0642)	.01 (.9471)
5	Search index	.17 (.6825)	21.82 (.0001)	38.55 (.0001)	.44 (.5144)
6	% Information	.06 (.8048)	3.51 (.0696)	1.81 (.1876)	2.93 (.0959)
7	% Reacquisitions	21.14 (.0001)	4.45 (.0424)	7.50 (.0097)	.21 (.6503)
8a	Column variation	2.58 (.1175)	1.78 (.1916)	2.79 (.1043)	3.23 (.0808)
8b	Row variation	.17 (.6869)	.01 (.9146)	.35 (.5572)	4.53 (.0404)

Note. The mean square error term Subject*Order**Treatment* has 34 degrees of freedom [2(18-1)(2-1)].

amount of time required to complete the task (Eyegaze 88.6 s and Mouselab 183.1 s). Thus, the more complex the task, the greater the difference in total time required to complete the task using Eyegaze and Mouselab.

Our analysis of EIPs yields not just directional but ratio predictions for the effect on Method. The ratio of predicted times compares favorably to the actual average time ratio (Table 2). Except for the 2 × 2 gamble, Mouselab required nearly twice as much time to complete each task. The large and significant task completion time differences establish that Mouselab was more effortful than Eyegaze. While internal cognitive processes are unobservable, we attribute the total time differences to greater WM loads imposed by Mouselab.

Read EIP. Overall, Mouselab Read EIP values were 336% longer than those for Eyegaze (1.265 s versus .377 s). There were also main effects for Task, Alternatives and Attributes. The Read EIP value was slightly longer for the gamble task than for the apartment selection task (.881 versus .761 s [note that the means average across process tracing method]). There were also slight differences based on the number of alternatives and the number of attributes. Tasks with seven alternatives had lower Read EIP times than those with two alternatives (.781 s versus .860 s). Tasks with seven attributes also had lower Read EIP times than those with two attributes (.735 s versus .906 s). Interpretation of a significant two-way interaction for Alternative × Attribute shows that Read EIP values decrease as task complexity increased. There was no significant Order effect.

The Eyegaze Read EIP value (0.377) is higher than a comparable timing parameter (0.230) reported by Card *et al.* (1983) and Russo (1978a). The Mouselab Read EIP time (1.276 sec) also is longer than the value of 1.19 s reported by Bettman *et al.* (1990) as well as the value of 1.10 s reported by Card *et al.* The values reported by Card *et al.* are for skilled, error-free, ideal behavior. It is not uncommon to find longer times for empirical data from unskilled users. Thus, it is not surprising that these times are slightly longer than those reported previously in Card *et al.*

Number of fixations. Subjects averaged 54 more fixations per task using Eyegaze than with Mouselab to complete each task (120 versus 66). Subjects averaged more fixations for the gamble task than for the apartment selection task (115 versus 72). Tasks with seven alternatives had more fixations than those with two alternatives (117 versus 69). Tasks with seven attributes had more fixations than those with two attributes (136 versus 51). Interpretation of a significant two-way interaction for Alternative × Attribute shows that number of fixations increase as task complexity increases. There was no significant Order effect.

Since subjects using Mouselab required nearly twice as much time to complete the task, the rate of fixations per minute was nearly three times greater for Eyegaze than for Mouselab. Russo (1978b) found the number of fixations per minute was at least ten times greater for eye movements than for information boards. Mouselab is intermediate between eye tracking and information boards, but is closer to eye tracking in granularity.

Accuracy. If subjects are consistent in applying their preference criteria to the choice tasks, their choice should not change as a function of the method used to collect process tracing data. A paired difference comparison of the apartment selection and gamble data found almost one-third (93/288) of the choices changed as a function of the process tracing method. This does not rule out the possibility that the choice differences reflect a change in preferences during the one week interval between trials or indicate which are correct. The experiment may be capturing a shift in preferences rather than true differences attributable to the process tracing method. Given the subjective nature of the apartment selection task, it is difficult to determine which process tracing tool facilitated more accurate decisions. While this might be true for apartments, gambles only have one optimum value. Thus, the remaining analyses and discussion focuses solely on accuracy in gamble selection.

A separate univariate, nonadditive ANOVA analysis for EV found significant differences for method. Expected value was greater with Eyegaze than with Mouselab (624 vs 602). There were no differences between tasks with two or seven alternatives (616 vs 611), however, there were significant differences between tasks with two or seven attributes (724 vs 501). There was a significant Order effect (first 598 vs second 628). EV ratio also was greater with Eyegaze than with Mouselab (.96 vs .93). There were significant differences between tasks with two or seven alternatives (.95 vs .93) and between tasks with two or seven attributes (.96 vs .93). There was a significant Order effect (first .92 vs second .96). For both EV and EV ratio, the significant Method × Alternative and Method × Attribute interactions show that differences in performance between Eyegaze and Mouselab increased as the number of alternatives or attributes increased.

Because of a greater burden on a capacity constrained WM, we postulated that Mouselab would cause more slips, forgetting and errors in the arithmetic calculations. Post-hoc analysis found that accuracy in the gamble selection is contingent upon the process tracing method. While Mouselab was never more accurate than Eyegaze, subjects were more accurate using Eyegaze for tasks with seven alternatives. In summary, the accuracy differences are small but tend to increase as the information processing demand of the problem increases.

Percent information searched. Results for percentage of information searched are counter to our predictions. Subjects using Mouselab examined 3.3% more information than Eyegaze users (93.3% versus 90.0%). There was no main effect for Task (gamble versus apartment selection). Subjects examined less information for tasks with seven alternatives than for those with two alternatives (85% versus 98%). Subjects also examined less information for tasks with seven attributes than those with two attributes (87% versus 96%). Interpretation of a significant two-way interaction for Alternative × Attribute shows that percent information searched decreases as task complexity increases. There was no significant Order effect.

Considering the cognitive effort associated with information search, we expected that subjects using Eyegaze would search more information. One possible explanation is that subjects using Eyegaze were able to gather information from the periphery of their eye without making a saccadic eye movements (Haber & Hershenson, 1980). This type of information acquisition is not captured with the Eyegaze System. Peripheral vision might enable subjects to view information in adjacent cells of a matrix without additional saccadic eye movements. For Mouselab, this additional information is hidden until pointed to with a mouse. Thus, Mouselab provides an additional level of control in preventing subjects from viewing peripheral information.

Another possible explanation is that Mouselab may predispose people to use a more systematic search and process more information than they normally would because of the way subjects acquire information using Mouselab. This is similar to the ideas proposed by Todd and Benbasat (1991) to design decision aids that facilitate the use of particular decision strategies. Because Mouselab users can view the data only one cell at a time, subjects may complete a systematic scan of the data initially to see the range of values. A systematic search would increase the percentage of information searched.

It is also important to note that the analyses compared only data that Mouselab was capable of capturing. Neither the Eyegaze nor Mouselab data include scans to row and column labels. Thus, we cannot determine the effect this additional information would have on the amount of information searched.

Search pattern. Payne (1976) observed a general tendency for search pattern to become more intradimensional as information load increased. Payne and Braunstein (1978) reported that search pattern became more intradimensional as the number of alternatives increased. Cook and Swain (1993) report similar results. Our findings also show that search pattern became more intradimensional as the number of alternatives increased.

Subjects acquired information from Mouselab with a slightly more interdimensional search than with Eyegaze. The difference is striking considering the same subject used a different information search strategy contingent upon the process tracing tool. There was a significant main effect for method (Mouselab .075 versus Eyegaze −.046). There also was a main effect for Task (gamble −.061 versus apartment .090). Search became more intradimensional as the number of alternatives increased (two alternatives .066 versus seven alternatives −.037). Search became more interdimensional as the number of attributes increased (two attributes −.065 versus seven attributes .095). There was also a significant two-way interaction for Alternative × Attribute [$F(1,34) = 4.35$, $P < .0446$; 2 × 2 .012, 2 × 7 .121, 7 × 2 −.142, 7 × 7 .068]. There was no significant Order effect.

Figure 4 compares the Mouselab and Eyegaze information acquisition data for the apartment selection task (2 × 7) for Subject 11. The search index was .21 for Eyegaze and .87 for Mouselab. Although Fig. 4 is a more extreme example, there are visually more transitions from attribute to attribute for Mouselab than for Eyegaze. Figure 5 compares the Mouselab and Eyegaze information acquisition data for the gamble task (7 × 7) for Subject 3. The search index was −.37 for Eyegaze and .09 for Mouselab. These figures illustrate differences in the search pattern index between the two methods.

Reacquisition rate. Overall, subjects reexamined more information using Eyegaze, 69%, than using Mouselab, 47%. Subjects reacquired more information for gambles than for the apartment selection task (gamble 65% versus apartment selection 51%). Reacquisition rate was greater for tasks with two alternatives than for tasks with seven alternatives (two alternatives 67% versus seven alternatives 49%). Also, reacquisition rate was greater for tasks with two attributes than for tasks with seven attributes (two attributes 60% versus seven attributes 56%). There was no significant Alternative × Attribute interaction effect or Order effect.

Russo (1978a) reported a reacquisition rate of 7% for information boards and 75% for eye tracking equip-

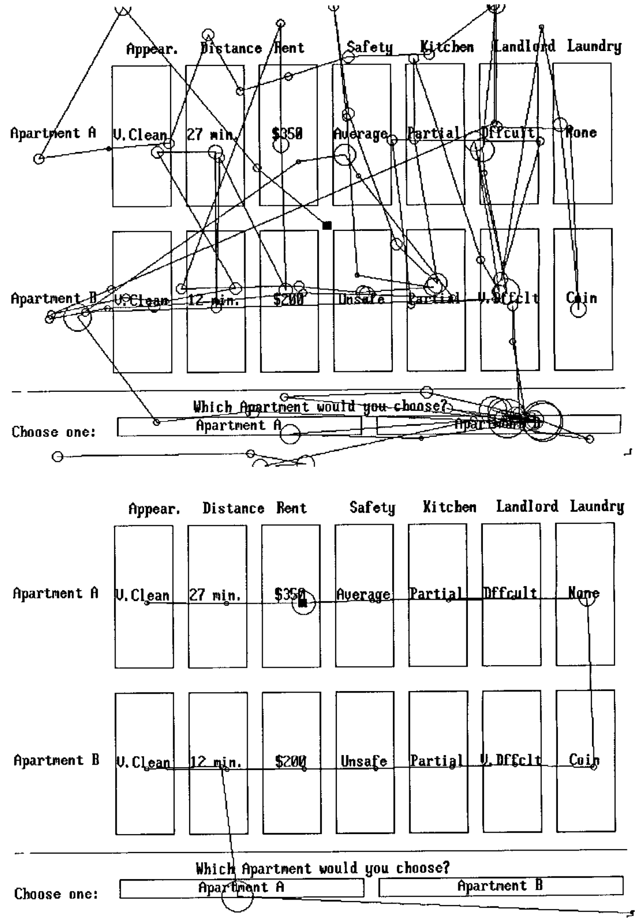

FIG. 4. Information acquisition data for the apartment selection task (2 × 7) for Subject 11. The top panel shows Eyegaze data; the lower panel shows Mouselab data. Mouselab data do not represent exact mouse locations. A jitter algorithm has been applied to a fixed pixel location for each box in the Mouselab display. A square shows the first information acquisition.

Method	Time (seconds)	Number	Choice	Search (%)	Search index
Eye	37.6	66 fixations	B	100%	.21
Mouse	23.0	18 fixations	A	100%	.87

ment. The reacquisition rate for Eyegaze, 69%, exceeds the value of 56% reported by Russo (1978b). These differences probably reflect differences in stimuli and eye tracking equipment. The reacquisition rate of 47% for Mouselab is much closer to the rate for eye tracking equipment than the rate for information boards. Mouselab has a much higher reacquisition rate than an information board reflecting the smaller time required to acquire one piece of information using Mouselab compared to an information board.

Variability of information search. Eyegaze had a greater variation than Mouselab in the proportion of information searched by alternative and by attribute. There was not a main effect for variation in the proportion of information searched for Task or for Order. Information search variability by alternative increased as the number of alternatives increased (two alternatives .016 versus seven alternatives .122) and as the number of attributes increased (two attributes .038 versus seven attributes .101). Variation in the proportion of information searched by attribute increased as the number of alternatives increased (two alternatives .021 versus seven alternatives .106) and as the number

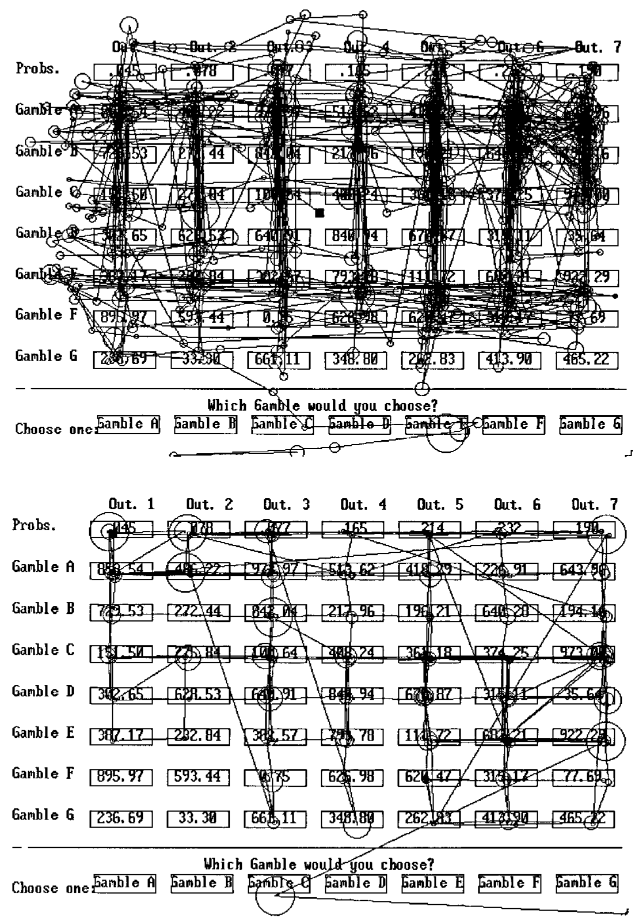

FIG. 5. Information acquisition data for the gamble task (7 × 7) for Subject 3. The top panel is Eyegaze data; the lower panel is Mouselab data. Mouselab data do not represent exact mouse locations. A jitter algorithm has been applied to a fixed pixel location for each box in the Mouselab display. A square shows the first information acquisition.

Method	Time (seconds)	Number	Choice	Search (%)	Search index
Eye	215.1	693 fixations	E-correct	100%	−.37
Mouse	300.2	227 fixations	C-incorrect	89.8%	.09

of attributes increased (two attributes .029 versus seven attributes .097). Subjects had less variation in their information acquisition strategies using Mouselab. Increased variability in search patterns suggests that Eyegaze subjects use more selective, data-driven (bottom-up) information processing.

Method interactions. The pattern of results reported in Table 3b reveals some interesting patterns for the method interactions. A significant Method × Task interaction shows a differential response of task contingent upon the process tracing method. There was a greater difference between Eyegaze and Mouselab for total time and Read EIP time for the gamble task than for the apartment selection task. The reverse was true for reacquisitions. Mouselab reacquisitions approached that of Eyegaze for gambles but were almost half that of Eyegaze for the apartment selection task. Table 3b also shows that problem size had a larger effect on method interactions than task. Of the ten dependent measures, 6/10 had significant Method × Alternative interactions and 6/10 had significant Method × Attribute interactions. There was a greater difference between process tracing method for tasks with seven Alternatives or seven Attributes than for tasks with only two Alternatives or two Attributes. In general, the more complex the task, in terms of amount of information to process, the greater the difference between Eyegaze and Mouselab.

IMPLICATIONS FOR RESEARCH

The method of recording information acquisition influenced the decision process. Eyegaze used less time, more fixations, and more reacquisitions, but resulted in less search of the total information and had a more variable pattern of information search. Eyegaze tended to have a more intradimensional search than Mouselab. Further, the interactions between Method, and the number of Alternatives and Attributes suggests that these differences between methods were greater for more complex problems. In contrast, the differences between the two tasks, selection of apartments and gambles were smaller than the effects attributable to problem size. For gambles, Eyegaze had more optimal choices. Thus, these results show that there are differences in decision processes as a result of the process tracing method used.

Are these differences important? One possibility is that the differences are relatively small and not of a sufficient magnitude to generate concern. However, the magnitude of the differences between Eyegaze and Mouselab for these process tracing measures is similar to the magnitude of some significant differences reported in recent experiments. Stone and Schkade (1994) compared speed–accuracy trade-offs of three attribute scales for an 8 × 6 choice task. The largest time difference was 23.4 s. In our study, Eyegaze was 40 s faster than Mouselab for the 7 × 7 apartment selection choice task. Similarly for search indices, Stone and Schkade (1991) noted in their abstract that ". . . relative to numbers, words led to more alternative-based information search and less compensatory processing." The main effect search index means were .407 for words and .320 for numbers for a difference of .087. Also, Payne *et al.* (1988, p. 547) reported that "Of greatest importance for the hypothesis of a hierarchy of time pressure effects was the finding of a significant effect of time pressure on pattern of processing in the first day for the 15-s condition ($M = -.11$ [no time pressure] vs $M = -.17$) [15 s time pressure], $F(1,2154) = 4.86$, $P < .05$, with more attribute-based processing under time pressure." The difference in search index between Mouselab and Eyegaze (−0.10) is as great as these effects. For variability of search, a similar story can be told. Using IS Lab, a keyboard-based CPT tool, Cook (1993) reported significant differences in the variability of information searched by alternatives (3Alt = .065 vs 10Alt = .115; $F = 25.53$, $P < .0001$) and the variability of information searched by attributes (3Alt = .155 vs 10Alt = .215; $F = 13.02$, $P < .0004$). Again, we found a similar size difference in these measures between Mouselab and Eyegaze (−.14 for alternatives and −.15 for attribute). Thus, the size of these effects are of the same size as other findings in the process tracing literature.

A more critical question is whether these differences are substantively important. It seems that this depends on the goal of the research, and the nature of the research question. Roughly speaking, process tracing research could be conducted for two reasons. The first is theory testing, examining how a manipulated independent variable affects choice behavior. The second is descriptive, examining the nature of the choice process itself. The implications of this research differs for these two goals.

For theory testing, the existence of differences between treatment conditions (e.g., time pressure, or display format) is the focus of the research. The process tracing method would affect theoretical findings only if there is an interaction between information acquisition effort and the manipulation. If the CPT tool is too effortful, subjects may adopt strategies that minimize information acquisition thus defeating the purpose of the method in providing a detailed nonverdical trace of the choice process. More importantly, the existence of a crossover or disordinal interaction in which the process tracing method would change the direction of the effect of a independent variable. For example, Russo, Johnson, and Stephens (1989) found that different methods of requesting verbal reports had very different effects for different tasks. Concurrent verbalization sometimes increased accuracy, sometimes resulting in lowered performance. In cases like this, the choice of a process tracing method can be quite problematic.

In other cases where information acquisition effort does not interact with the independent variables, it would be relatively harmless if a CPT tool slowed information acquisition and resulted in longer decision times. Such simple main effects are not likely to change the nature of the research conclusion. For example, if Mouselab slows acquisition of information, the results of Payne, Bettman, and Johnson that time pressure can lead to adaptive behavior would probably be larger, if eye movement recording had been used as the process tracing method, since its speed lends itself to more opportunistic behavior on the part of decision makers. In fact, our results here largely replicate some standard results in the process tracing literature (e.g., the proportion of information searched decreases as number of alternatives and number of attributes increases). Thus, when hypotheses being examined have directional predictions, and the manipulations are independent of the particular process tracing method that is used, the differences between process tracing methods are worth understanding.

The second reason for doing process tracing research is primarily descriptive, especially for a category of studies that describe the choice process itself and attempt to generalize the results to real world settings. For example, Johnson, Meyer, Hardie, and Anderson (1996) examined the role that various attributes play in the choice process for desktop computers using a Mouselab like system. If inferences about choice processes, such as the importance of price, were based on search patterns and looking time alone, then the process tracing tool would have been crucial. In fact our theoretical analysis suggests the direction of such effects. If an attribute is much easier to access than in the real world, for example data about the reliability of the computers, then it may appear to be used more frequently in the simulated choice environment than in reality. Johnson *et al.* acknowledged this problem and used multiple methods, such as discrete choice analysis, to supplement the process tracing analysis. Ultimately the usefulness of process tracing methods for purely descriptive work depends upon the closeness of the match between the simulated choice environment and the actual choice environment.

Our results suggest that generalized descriptions of choice processes across different kinds of process methods may have rather important limitations. For example, given that Eyegaze and Mouselab resulted in different search pattern values, it may be difficult to generalize search patterns found using a specific CPT tool beyond the scope of a specific process tracing method. The descriptive nature of process tracing measures like search patterns may not reflect the underlying real world processes.

However, with the advent of computer-mediated choice environments on the World Wide Web, the external validity issues become moot. Consumers navigating cybermalls and information services on the Internet provide vast quantities of clickstream data that are becoming an important source of intelligence for cybermarketing. Clickstream data capture and timestamp all mouse movements consumers use to navigate a Web site when a consumer visits a store on the Internet. These data also identify the previous site a consumer visited prior to entering the store. Clickstream data are analogous to Mouselab data in that the data provide information about the sequence and frequency of a user's selection of each item in a product display or catalog as well as total browsing time and choices.

Perhaps the most important methodological implication of this research is an initial understanding of how decision environments affect choice. The componential cognitive effort technique using EIPs from Bettman *et al.* (1990) helped estimate the magnitude of the differences in total time between the two process tracing methods. Our a priori analysis suggested, based on information acquisition time differences for basic operations, a set of hypotheses concerning the difference between these two methods. Similar analyses could be used to make predictions about the effect of other process tracing methods. Our task analysis which compared Mouselab to Eyegaze suggests that a comparison of those methods to eye movement recording will show greater differences that those between eye movements and a mouse, but that they will be substantially in the same direction.

Different CPT tools impose different levels of information acquisition effort. For example, CPT tools using keystrokes or a light pen would increase information acquisition costs relative to Mouselab. Some allow the user to select an entire row or column of data (e.g., Todd & Benbasat, 1991) and the information remains available to the user for the duration of the task. Others, like the keyboard-based, IS Lab (Cook & Swain, 1993), operate like Mouselab but indicate which information has and has not been viewed. Because each CPT tool is different, a description of the process tracing tool must explicitly state the details of these system functions. For example, a comparison of computer pointing devices found that a mouse is 5% slower than a theoretical optimal pointing device (Card *et al.*, 1983). In contrast, typing commands on a standard keyboard is 107% slower. CPT systems that use a keyboard like IS Lab (Cook & Swain, 1993) and the tool used by Todd & Benbasat (1991) will exhibit even greater demands on cognitive effort just for the mechanics of acquiring information. Thus, keystroke-based CPT tools would place even greater demands on WM than mouse-based CPT tools.

DIRECTIONS FOR FUTURE RESEARCH

Process tracing tools have largely been used for small choice tasks with 2–30 alternatives and 2–11 attributes (Johnson *et al.*, 1988; Schkade & Johnson, 1989; Todd & Benbasat, 1991, 1994). Often the CPT tool limits the size of the choice task because of pragmatic concerns such as font size in an 80 column by 40 row character-based display. Increasingly, the use of computer-mediated choice environments in real decisions as well as laboratory studies suggests the need for a richer set of stimuli, including non-matrix displays, multi-page sequential displays, displays with missing information, and tasks with a large number of alternatives and attributes. Each of these are important areas for future research exploring the effects of information acquisition patterns on consumer choice.

Non-matrix displays. Display format influences information acquisition and subsequent consumer behavior. For example, Russo (1977) induced supermarket shoppers to purchase products with lower unit prices by providing unit price information on a single list. Also, Bettman and Zins (1979) found that information processing strategy depends on the structure of the information. Alternative formats led to processing by alternatives. Attribute formats led to processing by attributes. However, contemporary information displays rarely present a matrix of alternatives and attributes in the row–column format used by Mouselab and other CPT tools. Eye movement equipment allows researchers to use test stimuli with very irregular arrangements of information such as retail catalogs, newspaper advertisements, Yellow Pages directories, retail kiosks, multimedia information services, and digital information products for electronic commerce. In some cases, such as the cybershopping on the World Wide Web, the analysis of information acquisition using clickstream data would actually be more realistic than eye movement recordings. Ultimately the selection of a process tracing method depends upon the fidelity of the system and the real world application.

Scanning sequence. The American Airlines airline reservation system is a classic example of the effects of scanning order on choice (Phillips & Thomas, 1988). The Sabre reservation system listed American Airline flights first. Being first resulted in more bookings. The government tried to block this unfair use of reservation systems by forcing the systems to begin giving competing flights equal display on computer screens. After a 12-year legal battle, Sabre ordered flights by time of departure with carrier arranged randomly if there were time ties.

The analyses of scanning patterns have implications for interface design of consumer information systems, especially for systems which force consumers to acquire information serially. The matrix-based displays is not likely to be as useful as eye movement recording for studying scanning sequence effects, especially for long multiple page listings.

Incomplete information. Research has explored the effects of incomplete or missing information on choice (Burke, 1988; Slovic & MacPhillamy, 1974). The matrix information display used by CPT tools heightens the awareness of missing information (an empty cell). However, it is generally much more difficult to discern what information is missing when the consideration set is large and the number of attributes is also large. If information were organized in a non-linear manner, missing information would be much less noticeable.

Large problem spaces. In many instances, one could argue that consumers make choices from dozens of alternatives each with numerous attributes. Large problem spaces are not well suited for study with current, matrix-based CPT tools. The common use of multiple page layouts for airline flight reservations, Yellow Pages ads, and newspaper classifieds exacerbates the difficulty of collecting information acquisition data for real world choice tasks. In contrast, eye tracking equipment is useful for large choice tasks with an irregular arrangement of information (e.g., Yellow Pages directories, Lohse, 1995; and retail catalogs, Janiszewski, 1992). However, since information acquisition via computer is becoming much more widespread, more sophisticated versions of computer-based tracking methods are viable in such environments.

CONCLUSION

In conclusion, the choice of process tracing methods is one which depends upon the goals of the research as well as pragmatic concerns. For example, the cost of Eye tracking equipment ($20,000–$100,000+) is prohibitive to many research institutions, whereas CPT tools are available for a nominal cost and are much easier for setting up and implementing a research study. It is also important to mention that the discussion of eye tracking equipment in the literature (Svenson, 1979; Cook & Swain, 1993) is based on equipment that is over 20 years old. Just as computers have rapidly increased in power and performance, eye tracking equipment has also improved significantly. Old eye tracking equipment had limited precision and accuracy for detecting small regions on a display. Current eye tracking systems are able to detect regions as small as 1.5 square centimeters. Contemporary eye tracking equipment no longer requires bite bars and other attachments to the head.[1] Thus, it seems time to reevaluate the importance of eye tracking equipment as a tool for process tracing studies. On the other hand, clickstream data from Internet transactions are extremely close to CPT data collected in the laboratory. Most computer-mediated choice environments, such as on-line services and World Wide Web applications, are mouse-based and the analysis of these data in this industry is becoming part of the ordinary conduct of business. Such data will become an important future extension of current CPT methods for studying consumer choice behavior. Our hope is that the EIP analysis presented here can provide a foundation for understanding when a method is a nonreactive and veridical representation of the underlying choice process.

REFERENCES

Abelson, R. P., & Levi, A. (1985). Decision making and decision theory. In G. Lindsay & E. Aronson (Eds.), *The handbook of social psychology,* (3rd ed., pp. 231–309).

Ashton, A. H., & Ashton, R. H. (1988). Sequential belief revision in auditing. *The Accounting Review,* **63,** 623–641.

Barsalou, L. W. (1992). *Cognitive psychology: An overview for cognitive scientists.* Hillsdale, NJ: Erlbaum.

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes,* **45,** 111–139.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research,* **3,** 233–240.

Bettman, J. R., & Zins, M. A. (1979). Information format and choice tasks in decision making. *Journal of Consumer Research,* **6,** 141–153.

Böckenholt, U., & Hynan, L. S. (1994). Caveats on a process-tracing measure and a remedy. *Journal of Behavioral Decision Making,* **7,** 103–117.

[1] Companies that produce eye tracking devices include LC Technologies, 4415 Glenn Rose Street, Fairfax VA 22032. (703) 424-7509; Iscan Inc, 125 Cambridgepark Drive, P.O. Box 2076, Cambridge MA 02238. (617) 868-5353; Applied Science Labs, 175 Middlesex Turnpike, Bedford, MA 01730-1428. (617) 275-4000.

Brucks, M. (1988). Search monitor: An approach for computer-controlled experiments involving consumer information search. *Journal of Consumer Research,* **15,** 117–121.

Burke, S. J. (1990). The effects of missing information on decision strategy selection. In M. E. Goldberg, G. Gorn, & R. W. Pollay (Eds.), *Advances in Consumer Research,* (Vol. 17, pp. 250–256). Provo, UT: Association of Consumer Research.

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology,* **39,** 338–366.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction.* Hillsdale, NJ: Erlbaum.

Cook, G. J. (1993). An empirical investigation of information search strategies with implications for decision support systems. *Decision Sciences,* **24,** 683–697.

Cook, G. J., & Swain, M. R. (1993). A computerized approach to decision process tracing for decision support system design. *Decision Sciences,* **24,** 931–952.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition,* **44,** 1–42.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology,* **32,** 52–88.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science,* **208,** 1181–1182.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review,* **102,** 211–245.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology,* **47,** 381–391.

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Processing tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Performance,* **43,** 75–117.

Galitz, W. O. (1985). *Handbook of Screen Format Design,* QED Information Sciences, Inc. Wellesley, MA.

Haber, R. N., & Hershenson, M. (1980). *The Psychology of Visual Perception.* New York: Holt, Rinehart and Winston.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology,* **17,** 417–444.

Jacoby, J., Kuss, A., Mazursky, D., & Troutman, T. (1985). Effectiveness of security analyst information accessing strategies: A computer interactive assessment. *Computers in Human Behavior,* **1,** 95–113.

Jacoby, J., Jaccard, J., Kuss, A., Troutman, T., & Mazursky, D. (1987). New directions in behavioral process research: Implications for social psychology. *Journal of Experimental Social Psychology,* **23,** 146–175.

Janiszewski, C. (1992). *The influence of page layout on directing and sustaining attention.* Working paper, College of Business Administration, University of Florida.

Jarvenpaa, S. L. (1989). The effect of task demands and graphic format on information processing strategies. *Management Science,* **35,** 285–303.

Johnson, E. J., Meyer, R. J., Hardie, B., & Anderson, P. (1996). *Process aided choice modeling.* Working paper, Department of Marketing, The Wharton School, University of Pennsylvania.

Johnson, E. J., & Payne, J. W. (1985). Effort and accuracy in choice. *Management Science,* **31,** 395–414.

Johnson, E. J., Payne, J. W., & Bettman, J. R. (1988). Information displays and preference reversals. *Organizational Behavior and Human Decision Processes,* **42,** 1–21.

Logie, R. H. (1986). Visio-spatial processing in working memory. *Quarterly Journal of Experimental Psychology,* **38A,** 229–247.

Lohse, G. L. (1995). An eye fixation analysis of consumer choice from phone directory advertisements. In (Ed.), *Winter Marketing Educator's Conference,* La Jolla, CA, February 11–14, Chicago, IL, AMA Press.

McCloskey, M., Harley, W., & Sokol, S. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain damaged subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **17,** 377–397.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review,* **63,** 81–97.

Newell, A., & Simon, H. (1972). *Human problem solving,* Englewood Cliffs, NJ: Prentice Hall.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance,* **16,** 366–387.

Payne, J. W., & Bettman, J. R. (1994). The costs and benefits of alternative measures of search behavior: Comments on Böckenholt and Hynan. *Journal of Behavioral Decision Making,* **7,** 119–122.

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,* Cambridge, England: Cambridge University Press.

Payne, J. W., & Braunstein, M. L. (1978). Risky choice: An examination of information acquisition behavior. *Memory and Cognition,* **5,** 554–561.

Payne, J. W., Braunstein, M. L., & Carroll, J. S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavioral and Human Performance,* **22,** 17–44.

Phillips, C., & Thomas, P. (1988, August 19). Texas Air asks transportation agency to examine rivals' reservation systems. *Wall Street Journal,* p. 13.

Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *Quarterly Journal of Experimental Psychology,* **38A,** 689–703.

Russo, J. E. (1977). The value of unit price information. *Journal of Marketing Research,* **14,** 193–201.

Russo, J. E. (1978a). Adaptation of cognitive processes to eye movement systems. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye Movements and Higher Psychological Functions,* (pp. 89–109), Hillsdale, NJ: Erlbaum.

Russo, J. E. (1978b). Eye fixations can save the world: A critical evaluation and a comparison between eye fixations and other information processing methodologies. In H. K. Hunt (Ed.), *Advances in consumer research (Volume 5),* (pp. 561–570), Ann Arbor, MI: Association for Consumer Research.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition,* **17,** 759–769.

Schiffman, H. R. (1976). *Sensation and perception an integrated approach.* New York: John Wiley and Sons.

Schkade, D. A., & Johnson, E. J. (1989). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes,* **44,** 203–231.

Simon, H. A. (1974). How big is a chunk? *Science,* **183,** 482–488.

Slovic, P., & MacPhillamy, D. (1974). Dimensional commensurability and cue utilization in comparative judgment. *Organizational Behavior and Human Performance,* **6,** 649–744.

Smyth, M. M., & Pendleton, L. R. (1989). Working memory for movements. *Quarterly Journal of Experimental Psychology,* **41A,** 235–250.

Stone, D. N., & Schkade, D. A. (1991). Numeric and linguistic information representation in multiattribute choice. *Organizational Behavior and Human Decision Processes,* **49,** 42–59.

Stone, D. N., & Schkade, D. A. (1994). Effects of attribute scale on process and performance in multiattribute choice. *Organizational Behavior and Human Decision Processes,* **59,** 261–287.

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance,* **23,** 86–112.

Todd, P., & Benbasat, I. (1991). An experimental investigation of the impact of computer based decision aids on decision making strategies. *Information Systems Research,* **2,** 87–115.

Todd, P., & Benbasat, I. (1994). The influence of decision aids on choice strategies under conditions of high cognitive load. *IEEE Systems, Man and Cybernetics,* **24,** 537–547.

Tullis, T. S. (1988). Screen Design. In M. Helander, *Handbook of Human-Computer Interaction.* New York, NY: Elsevier Science Publishers, Inc.

van Raaij, W. F. (1977). Consumer information processing for different information structures and formats. In *Advances in Consumer Research* (pp. 176–184). Provo, UT: Association for Consumer Research.

Widing, R. E., & Talarzyk, W. W. (1993). Electronic information systems for consumers: An evaluation of computer-assisted formats in multiple decision environments. *Journal of Marketing Research,* **30,** 125–141.

Williams, J. D. (1990). Risk assessment: An investigation of auditor's decision processes. *Proceedings of the Audit Judgment Symposium,* University of Southern California.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods and Instrumentation,* **7,** 397–429.

Received: October 30, 1995