THE JOURNAL OF BEHAVIORAL FINANCE, 11: 135–149, 2010

ISSN: 1542-7560 print / 1542-7579 online
DOI: 10.1080/15427560.2010.507155

ARTICLES

An Experimental Examination of Heuristic-Based Decision Making in a Financial Setting

Lucy F. Ackert
Kennesaw State University

Bryan K. Church
Georgia Tech

Paula A. Tkac
Federal Reserve Bank of Atlanta

This paper reports the results of an experiment designed to examine information acquisition and evaluation in a financial setting, predicting mutual fund performance. We compare behavior across four distinct subject pools to provide insight into how training, knowledge, and experience affect decision making. We manipulate the decision environment by first increasing the time constraint and then increasing the decision cost. Although we find differences in behavior across subject pools, subjects' performance is similar. Outcomes are similar across the distinct subject pools, despite significant differences in financial education.

Keywords: Heuristics, Financial decisions, Mutual funds, Financial education

INTRODUCTION

Over the last few decades, economists, psychologists, and other researchers have devoted considerable attention to how people make choices. Increasingly, financial researchers are delving deeper into the decision making and behavior of individual market participants.[1] Typically, the behavior of investors is compared to that of a model of optimal decision making, which assumes that behavior is rational in a full-information, full-optimization environment without decision costs. Not surprisingly, much empirical evidence suggests that investor actions deviate from the predictions of such a model. Some authors have argued that this is indicative of irrationality or behavioral biases, while others, most often in the psychology literature, argue that we are observing the effect of decision costs.[2] To make progress in understanding and interpreting investor behavior along these lines, there is a need for deeper understanding of investor decision making in financial contexts. This paper reports the results of an experiment designed to address the gap in our knowledge by studying information acquisition and evaluation in a controlled financial environment.

Specifically, the task facing our experimental subjects is to predict mutual fund performance using information on a variety of historical performance measures. By asking subjects to predict only (as opposed to asking them to choose a preferred fund), the influence of risk aversion on subject decisions and actions is not a concern. Similarly, by presenting the same information and task to subjects who vary along demographic dimensions, we can analyze any effect of characteristics such as financial education, experience, and gender on the decision-making process. Finally, the software we utilize in our experiment allows us to observe not only subject decisions (i.e., predictions of future fund performance) but also to track each subject's pattern of information acquisition (e.g., what pieces of information they access, in

Address correspondence to Lucy F. Ackert, Department of Economics and Finance, Michael J. Coles College of Business, Kennesaw State University, 1000 Chastain Road, Kennesaw, GA 30144. E-mail: lackert@kennesaw.edu

what order and for how long). Our experimental method does not include any feedback to subjects during the experiment regarding the accuracy of their decisions. Therefore, we do not need to incorporate learning into our analysis and instead provide a static description of the decision-making process.

We compare behavior across three distinct subject pools, which may be thought to vary in expertise and sophistication: (1) students in the Georgia Tech Master's Program in Quantitative Computational Finance (QCF), (2) students in the Georgia Tech Master's Program in Business Administration (MBA), and (3) a group of employees at the Federal Reserve Bank of Atlanta (FRB). As a group, the FRB pool is fairly similar demographically to the universe of mutual fund investors in the marketplace. However, there are some potentially interesting and important differences across employees as well. In particular, we separate out the FRB subjects who have had no formal financial education (e.g., in high school, college) to study the role of financial training.

Prior instruction and training have been shown to play important roles in the speed and accuracy of decisions. Experts organize knowledge in long-term memory in a way that allows them to effectively perform certain tasks (e.g., Chase and Simon [1973], Anderson [1982]). However, expertise is not necessarily an immediate consequence of education or experience but also results from practice (Ericsson, Krampe, and Tesch-Romer [1993]). Theory argues that knowledge structures are enriched through practice, which allows decision making to become automatic. Hence, one conclusion is that experts are able to make superior decisions in less time than others (Rikers, Winkel, Loyens, and Schmidt [2003]).

The influence of prior financial training and experience on decision time is just one example of the various aspects of decision making that we investigate. In addition, we present evidence on both the amount and characteristics of the information accessed by each subject pool, a comparison of the accuracy of their performance, and some characterizations of the decision strategies used.

By varying the computerized environment, we also investigate how time constraints and decision costs effect information acquisition and evaluation. Each experimental session consists of three phases across which the decision environment becomes more time constrained. In the initial phase, participants are given two minutes to evaluate each mutual fund pair and make a decision (predict which will perform better in the future). We then impose a 30-second time constraint on each comparison in phase two in which a subject's payoff equals 0 if he or she fails to make a decision within the 30 seconds. Finally, in the third set of paired comparisons, time is costly; subjects' compensation for a correct prediction is scaled down by the amount of time used to make the decision. A constant opportunity cost of time, which can be translated into a constant opportunity cost of effort. Under costly effort, we effectively impose a tradeoff on the subjects' decision-making environment: effort v. accuracy. Hypothetically subjects may respond to this tradeoff by merely decreasing the time to make a decision, but they may also alter their strategy of information collection (i.e., collect less) or even the way in which they process information to make a decision.

This situation resembles a real problem faced by investors. When making decisions of a financial nature, people are often inundated with information. Taking proper account of their decision-making environment is crucial because there is a trade-off between the benefit and cost of information search and processing. The ability to collect and assimilate all the information required to make a decision that *ex post* can be evaluated as optimal is beyond mere mortals. Simon (1990) conceptualized decision making as involving the limitations of the human mind and the structure of the environment in which the person operates. Recent research has documented the value of useful decision tools known as heuristics. Heuristics are simplified strategies for processing information and making decisions that allow reasonable judgments to be made when faced with time and knowledge constraints.[3]

The remainder of this paper is structured as follows. In the following section we develop a framework to provide a basis for our experimental investigation. In the third section we describe our research method, including participants, procedures, and tools. The fourth section reports the results of the experiment, and the fifth section summaries our findings and offers concluding remarks.

FRAMEWORK

Clearly people have large differences in financial knowledge. In cognitive psychology, declarative knowledge is what is known about a particular domain (Anderson [1982, 1987]). Declarative knowledge is based on facts and rules. In contrast, training and experience provide procedural knowledge, which relates to how to do something in a particular environment. In our setting, declarative knowledge includes information on mutual funds and the measurement of their performance, and procedural knowledge pertains to how a person will choose a superior-performing fund. Psychologists recognize that both declarative and procedural knowledge are required for learning. In fact, they are complementary; a larger store of declarative knowledge facilitates gains in procedural knowledge and experience enhances declarative knowledge.

Our goal is to provide insight into how knowledge facilitates financial decision making. What does it take to make better decisions? A large body of evidence has accumulated on expertise in problem solving across a variety of domains. For example, Chi, Fletovich, and Glaser [1981] show that experts use more abstract, richer mental representations of physics problems, as compared to novices. Rikers, Winkel, Loyens, and Schmidt [2003] document that medical experts process case information faster and more accurately than nonspecialists or students. Ste-Marie [1999] shows that expert gymnastics judges are better at anticipating information, judging accurately, and exhibit more depth and breadth in their declarative knowledge. In the business domain, Collins, Gong, and Hribar [2003] conclude that investor sophistication mitigates mispricing.[4]

The literature thus closely links expertise with faster and more accurate decisions. It has also shown that more experienced decision makers use a directed information search strategy, whereas those with less experience are more likely to use a sequential search strategy.[5] With a sequential search, a person considers all information in an ordered manner. More experienced decision makers use more sophisticated information search and evaluation strategies, including heuristics adapted to the particular environment.

We use an experimental method to examine individual decision making because it affords control over extraneous factors and information availability.[6] Of course, we cannot control the information subjects bring to the experiment in their declarative knowledge base. However, this is precisely what we manipulate by including four subject pools. Ball and Cech [1996] argue that if significant differences are observed across subject pools, more investigation of individual characteristics is necessary. It is these individual differences in knowledge and experience that we seek to study in a financial environment.

Often in the finance literature, knowledge and expertise are labeled "investor sophistication," and differences across investors are limited to studying the behavior of large groups of investors who vary along this dimension (i.e., institutions vs. individuals).[7] Individual investors, however, are not a homogeneous group. Research is beginning to take this into account. For example, Ackert and Church [2001] document differences in market behavior across subject pools with different levels of knowledge of financial markets and experience with a particular trading institution. Certainly initiatives for increased investor education by the Securities and Exchange Commission (SEC) and the Federal Reserve, among others, assume that such education will make a positive difference in decision making. We view our examination as an initial foray into the elementary aspects of individual decision making in financial settings. Our results will shed light on the impact of financial education, establish some stylized facts concerning investor decision making, and provide a basis on which to assess our conclusions about investor rationality and market function.

RESEARCH METHOD

Overview

We conduct an experiment to examine information acquisition in a financial setting. Subjects' earnings are determined by their ability to choose the best performing mutual fund among pairs of funds. In all sessions, subjects are given four cues, including alpha, the Sharpe Measure, average annual return for the preceding five years, and a fund rating. After considering the cues, subjects choose which fund is expected to perform better (in terms of raw return) than the subsequent period. For each correct choice, earnings increase \$1.00.[8] Within each session, across three sets of fund pairs, we increase the decision cost by first decreasing the time permitted to evaluate each fund pair and then imposing a payoff function that is contingent on the time elapsed. At the conclusion of each session, subjects' earnings are computed; no feedback is provided until then.

Participants

We vary the subject pool across treatments. In the first set of sessions, students were recruited from the Master of Science degree program in Quantitative and Computational Finance (QCF) at Georgia Tech and PhD programs in economics decision analysis. The QCF program is an intensive graduate program in finance with a quantitative emphasis, and its students have proven mathematical ability. The 20 QCF and 7 PhD students were recruited through e-mail and participated in staggered 60-minute sessions on the same day. The average age of these subjects was 26.82 years. These sessions were conducted in a student computer lab on the university campus that was used only by QCF students.

In the second set of sessions, we recruited students in the Masters in Business Administration (MBA) program, also through e-mail at Georgia Tech. The average age of these 37 subjects was 27.47 years. The sessions were conducted in a student computer lab in the management building at Georgia Tech.

In the final set of experiments, we recruited employees of the Federal Reserve Bank of Atlanta (FRB), located approximately one-half mile from the Georgia Tech campus. These subjects were recruited using advertisements carried on the internal computer network, as well as ads on television screens that are located throughout the building. The average age of the 67 subjects was 42.28. The sessions were conducted in groups of 12 or less in a computer training lab in the FRB building.

Procedures

At the beginning of each session, subjects are seated in front of a computer terminal and given consent forms. After the consent forms are collected, instructions for the initial section of the experiment are distributed and read aloud by an experimenter.[9] The purpose of this exercise is to elicit a measure of risk tolerance. Subjects are asked to make a single choice. They are endowed with \$10 and instructed that they can invest any portion of the endowment in a risky asset that has a 50% chance of success. The amount not invested is theirs to keep. At the conclusion of all sections of the session, an experimenter individually tossed a coin for each subject to determine if the investment was successful.

After all complete the risk measure, another set of written instructions is distributed. An experimenter reads the instructions aloud, while subjects read along. They are asked to evaluate the relative performance of paired mutual funds. The funds are labeled generically, and names are not provided. The set includes 7 funds so that subjects make 21 paired evaluations and are given 2 minutes to make each decision.[10] Decisions are made on the computer using the MouseLab software, as discussed subsequently. We refer to the first group of paired comparisons as the unconstrained set (*UNCONSTRAINED*). The cue and performance information for all funds in this study were constructed based on extensive analysis of a sample of equity mutual funds from the *CRSP Survivorship Bias-Free Mutual Fund Database*, supplemented with data from Morningstar, Inc. With the universe of domestic, equity mutual funds for 1997–1998, we estimated the predictability of performance given the 4 cues. We then generated a representative set of data for 21 funds. Within the simulated data, the predictive power of the four information cues is 85%.

After completion of the first set of paired comparisons, subjects are asked to raise their hands. They are then given another set of instructions, which they read before evaluating the second set of 21 fund pairs.[11] In the second set of paired comparisons, subjects are limited to 30 seconds. Otherwise, the procedures are identical to those used in the first set of paired comparisons. The second set of paired comparisons is referred to as the constrained set (*CONSTRAINED*). As before, upon completion subjects are asked to raise their hands so that an experimenter may give them instructions for the third and final set of comparisons. Again, 21 comparisons are made; however, subjects' payoff for a correct choice decreases as time elapses. The payoff is computed as follows: payoff = (1-t)*$1 where t is the fraction of time taken to make the final decision. The group of paired comparisons is referred to as the clock set because of the increased importance of "beating the clock" (*CLOCK*).

Questionnaires are distributed at the end of the experiment. Subjects characterize their confidence in their understanding of the four fund information cues provided on a set of 11-point bipolar adjective scales, with scale endpoints (bipolar adjectives) not confident/very confident. Subjects are also asked to evaluate the predictability of mutual fund returns (unpredictable/predictable), the skill of fund managers (unskilled/skilled), the attractiveness of funds as investments (unattractive/very attractive), and the ease of identifying superior performing funds (very difficult/very easy) using the same type of 11-point scales. Next, subjects complete a demographic questionnaire while the experimenters toss a coin for each subject to determine the value of the investment used to measure risk tolerance and calculate their earnings. Last, they are paid and dismissed.

All QCF and MBA students are compensated immediately in cash, and FRB subjects receive "Fed dollars." Prior to the experiment the FRB subjects are encouraged to examine the prizes they could purchase with the "Fed dollars" earned during the experiment. A wide variety of items is offered including backpacks, clocks, calculators, pens, coffee tumblers, and snack coupons for the cafeteria. Feedback from the participants indicated that they valued the prizes and found them motivating.

For each fund pair, a subject's performance is evaluated by comparing the fund selected to a model fit to actual mutual fund returns. Based on the sample of equity funds described earlier, we regressed future fund return on the 4 information cues (prior performance measures). We then added a random error term to each fund's predicted return so that the in-sample accuracy of the regression was 85%. These simulated returns were used to identify the superior performing fund. This method removes extraneous noise in fund performance while retaining the realistic relation between each past performance cue and future fund return.[12]

The MouseLab Tool

The MouseLab program is a computer-based system developed to study decision making and information acquisition.[13] The system, developed by Johnson, Payne, Schkade, and Bettman [1991], has been used in numerous research studies.[14] MouseLab allows an experimenter to present a multiattribute decision problem in a matrix schema. The program records the information acquired, the order, the duration of time spent on each information cue acquired, and the final choice. Response time accuracy is within 1/60th of a second.[15]

Prior to the computerized portion of the experiment, an experimenter reads written instructions aloud. These written instructions include definitions of the information cues, which subjects keep for later reference. To ensure complete understanding of the task, the experimenter encourages subjects to ask questions at any time, and the instructions are repeated on text screens on the computer. Participants are then given the opportunity to familiarize themselves with the MouseLab program.

Figure 1 presents a sample screen. In our sessions, subjects were given a sample screen to consider prior to beginning the computerized portion of the experiment. The screen shows 8 information boxes that can be opened, one at a time, by a click of the mouse. A box must be closed before another box can be opened. For each fund in the pair, the boxes contain information on alpha, the Sharpe Measure, average annual return, and a fund rating. When the subject is ready to choose between the two funds, he clicks on either asset 1 or asset 2 on the bottom of the screen. A clock in the top left-hand corner of the screen displays the time elapsed and time remaining. The program stores all actions of the subject, including the time boxes are opened and closed, and the final decision.

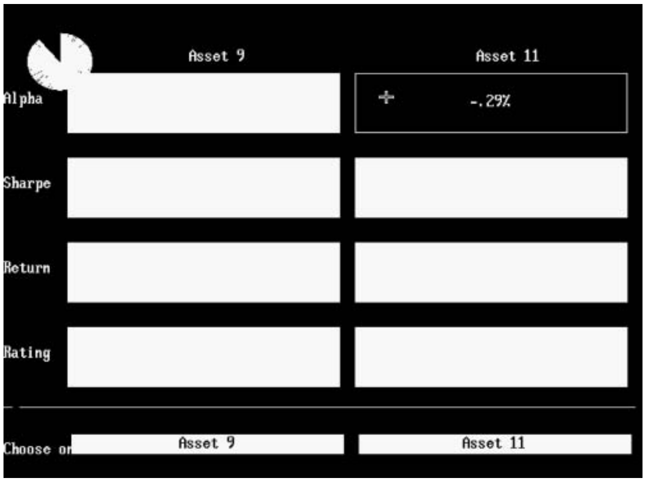

FIGURE 1 The MouseLab screen.

RESULTS

Subject Pool Demographics, Experience and Expertise

Table 1 reports comparisons of demographic information and self-reported measures of investment experience and expertise for 3 subject pools, MBA, QCF, and FRB employees. The FRB pool is fairly diverse with respect to financial education; the range of courses taken is 0–15. To isolate the role of prior financial training, we have further divided this pool into employees with some finance courses in their academic training (FRB Finance) and those with none (FRB No Finance).

The subject pools differ significantly across the three demographic characteristics we measured: gender, age, and years employed. The FRB pool has the largest percentage of women (70%) compared to the MBA pool (35%) and the QCF pool (11%). Not surprisingly, the FRB pool is on average older (42.5 years of age) and has more years of work experience (9.5 years of experience) compared to the MBA (28 years of age and 4.7 years of experience) or QCF (27 years of age and 3 years of experience) pools. These differences are not sample specific but rather reflect differences in the larger universe of corporate employees, MBA students, and those pursuing Master's degrees in a program such as QCF.[16] Not surprisingly, due to their student status, MBA and QCF subjects have lower levels of household income than FRB subjects.

In terms of general education and financial training, the pools again differ significantly. All subjects in the MBA and QCF pools have a four-year college degree versus 73% of the FRB subjects. The exposure to formal financial training echoes this pattern, with MBAs reporting the highest average number of finance courses taken (7.8), while QCF reports 5.7 and the FRB Finance pool 4.5. The FRB No-finance pool has no formal financial education, by definition.

TABLE 1
Subject Group Characteristics.

Characteristic	MBAs	Quant – Finance	FRB Finance	FRB No Finance
Total number of subjects	37	26	44	20
Men (percentage)***	65	89	36	15
Age (mean)***	27.95	26.82	42.32	42.89
Years employed*** (mean)	4.66	3.04	9.78	9.00
Four-Year college degree (percentage yes)***	100	100	91	45
Household income (median)***	$0–$25,000	$0–$25,000	$75,001–$100,000	$50,001–$75,000
Traded securities (percentage yes)*	36	37	18	11
Invested in mutual funds (percentage yes)	44	33	59	53
Risk tolerance (average)	6.67	6.67	6.30	5.90
Earnings Based on Choices (mean)*	$38.89	$40.22	$37.59	$37.64
Financial expertise*** (little/very knowledgeable)	5.76	6.59	6.14	5.44
Finance courses taken***	7.77	5.67	4.48	0.00
Alpha (not/very confident)**	6.69	6.73	5.53	4.95
Average return*** (not/very confident)	9.40	8.85	8.09	7.35
Sharpe** (not/very confident)	6.42	7.62	6.80	5.05
Star rating* (not/very confident)	8.06	7.27	7.42	6.30
Fund returns (unpredictable/predictable)	6.19	6.88	6.55	6.30
Managers' skill (unskilled/highly skilled)	6.36	7.32	7.00	7.05
Funds for investment (unattractive/attractive)***	7.33	7.44	8.36	6.80
Identification of superior performer (difficult/easy)	6.57	7.04	6.91	6.15

Note. ***,**,*denotes significance at the 1%, 5%, and 10% level, respectively.

For continuous variables, such as earnings and the confidence (likert) scales, we conduct a Kruskal-Wallis test to assess the null hypothesis that the median does not differ across subject groups. For the binary variables (yes/no variables), we conduct a chi-square test to assess the null hypothesis that responses are independent of subject group.

In terms of investment experience, the pools differ dramatically depending on the type of instrument considered. The FRB pools have less experience trading securities, with only 18% and 11% of subjects reporting this activity in the Finance and No-finance pools, respectively. The MBA and QCF pools both report a participation percentage of around 36% in securities trading. In contrast, the MBA and, more so, QCF pools have less experience investing in mutual funds than either FRB pool (44% and 33% vs. 59% and 53%, respectively).

As a whole, the FRB pool is demographically similar to the universe of mutual fund shareholders in the United States. According to the 2001 Mutual Fund Shareholder Survey conducted by the Investment Company Institute, the mean age of mutual fund investors is 47, 52% have a four-year college degree, 78% are employed, and household income averages \$95,700. One exception to this similarity is the high percentage of female subjects in the FRB pool; women are the sole decision makers in only 23% of mutual fund-owning households. However, the percentage of co-decision-making households is 53%, suggesting that women potentially affect decision making in 76% of mutual fund households.

Table 1 also reports the subjects' evaluations of their own financial expertise and confidence in understanding the four information cues provided. In terms of overall financial expertise, the QCF subjects had the highest median self-rating of 6.59 on an 11-point scale. The No-finance FRB subjects report a median expertise rating of 5.44, the lowest of the 4 pools. This is not surprising and correlates positively with the average number of finance courses taken. Interestingly, risk tolerance, as measured by our instrument, is positively correlated to self-reported levels of financial expertise, though there are no statistically significant differences in this measure across the groups. Whether the increased risk-tolerance of the QCF pool versus the FRB No-finance pool reflects their higher level of financial education and view of their financial expertise cannot be determined from this questionnaire, but is an interesting subject for future research. Does more financial education lead to greater risk tolerance and risk taking?

With respect to the four performance information cues used in the experiment, alpha, Sharpe measure, average return, and Morningstar rating, FRB No-finance reports significantly lower levels of confidence across all measures than the three other pools that have had financial training. Confidence appears to be related to prior financial education on these measures rather than the experience of investing in mutual funds per se. All 4 subject pools report being more confident using the average return measure than any of the other 3 performance measures. Except for the QCF subjects, who report Sharpe as their second most confidently used measure, all pools place Morningstar rating in this place. Irrespective of financial expertise and education, a simpler measure such as return or Morningstar rating appears to be easier to interpret and utilize than more quantitative measures such as alpha and Sharpe. Later we study subjects' information acquisition and decisions to see if these relative confidence levels correlate to decision-making actions.

Subjects were also asked to evaluate 4 aspects of the mutual fund investment environment: the predictability of mutual fund returns (unpredictable/predictable), the skill of fund managers (unskilled/skilled), the attractiveness of funds as investments (unattractive/very attractive), and the ease of identifying superior performing funds (very difficult/very easy). There are no noticeable differences across subject pools in the responses to these questions, though all subjects likely rate predictability and managerial skill higher than a reading of the academic literature on mutual funds would advocate.[17]

Our analysis of subject pool behavior during the experiment takes on two different perspectives. First we present an analysis focusing on subjects' actions within each fund comparison screen to characterize how they acquire and process information in this environment. Second, we analyze the subjects' actual decisions in an attempt to uncover their decision-making rules and assess performance and accuracy. In each analysis, a central question is whether there are any significant differences across the subject pools and to characterize the subject pools according to their similarities and differences.

Process Analysis

In studying the decision-making process of subjects, we take advantage of the ability of the MouseLab software to track every information cue viewed by subjects and the time spent accessing this information. We present data on the total decision-making time per screen, the number of cues accessed, and the distribution of cues accessed.

Decision-Making Time

The first dimension along which we will compare subject behavior is the time used to make a decision for each pair of mutual funds (i.e., each screen). Figure 2 illustrates the average time per screen with the subjects divided into the 4 pools and the screens grouped into 6 groups: the first and second 10 screens under each of the three time treatments (*UNCONSTRAINED, CONSTRAINED, and CLOCK*). Recall that after screen 20 the subjects shift into a constrained environment with only 30 seconds per screen; after 40 screens they begin the final phase in which they experience a constant opportunity cost of decision time.[18]

The drop in average decision time between the first and second half of *UNCONSTRAINED* is strong evidence that all subject pools were learning about the decision environment at the beginning of the experiment even without feedback. In contrast, once this experience was gained, the imposition of time constraints did not seem to precipitate further learning (i.e., the average decision time for halves 1 and 2 in *CONSTRAINED* and *CLOCK* are not different within each pool).

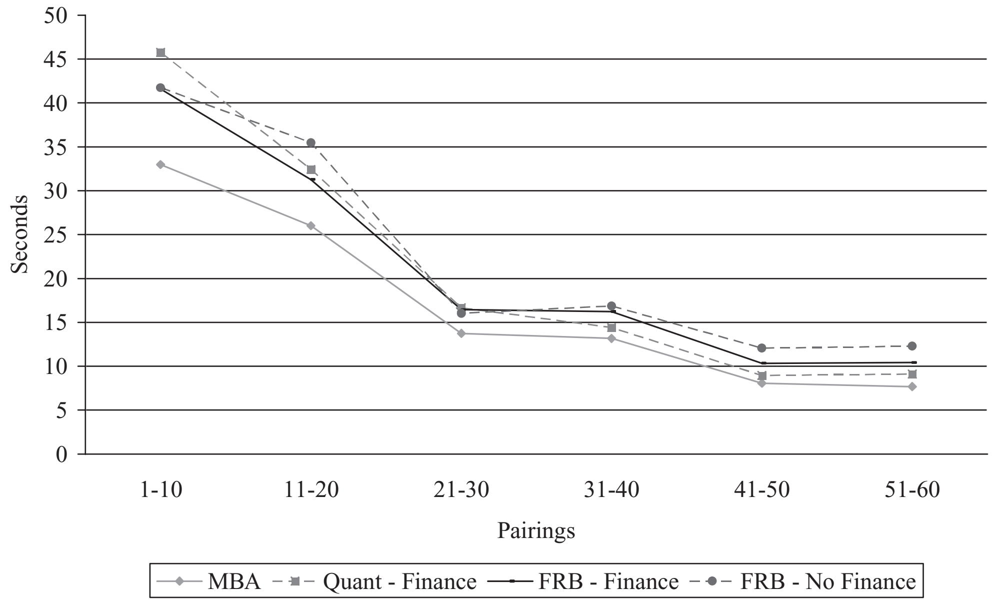

FIGURE 2 Seconds Per Choice Pair x Subject Group.

This is not to say that the time constraints were unimportant; indeed decision times did fall as constraints were imposed but within the two latter funds set subjects did not appear to need experience with the constraint to refine their decision strategy.

In every fund set the MBA students made their decisions faster than either the QCF students or the FRB employees though the difference is significant only in *UNCONSTRAINED*. Perhaps this reflects a superior ability or willingness to make quick decisions, characteristics that may be beneficial to a businessperson but not necessarily for the careers that attract QCF students. The differences across pools are largest in the initial set of 20 decisions. As time constraints are imposed, average decision times converge somewhat. For example, in the second half of the first set, once task-related learning had been achieved, the range in average decision time across the 4 pools is 9 seconds, but by the final 10 decisions (51–60), under a continuous time constraint, the range is only 5 seconds.

Number of Cues

Table 2 reports the average of the number of unique information cues accessed across subject pools and fund sets. The number of unique cues is counted by including only the first time a cue is accessed, so that multiple clicks on the same box within the same decision screen are not counted.[19] With 4 information cues and 2 funds, the upper limit on unique cues is 8. Notice that all pools access a large number of cues on average, about 7 in the first 2 sets of fund comparisons. Moreover there are no statistically significant differences in the number of cues accessed by the various subject pools within a given set of fund comparisons (e.g., *UNCONSTRAINED*). This indicates that, no matter what their level of financial training, subjects tend to look at almost all of the information available when making their decisions.

Such a finding may be surprising given that subjects with less financial training and lower confidence in their ability to use various cues (i.e., the FRB No-finance pool) still viewed these cues. A potential explanation may be that subjects generally inferred that all information given was important precisely *because* it was presented to them (i.e., it would not have been presented if it did not aid in decision making). A similar dynamic may well be at work when investors view mutual fund information in a prospectus or other disclosure—why is the reporting of such information mandatory if it is not helpful or valuable in some way? Such a view makes an 'information overload' scenario quite plausible. It is important to note, however, that viewing cues and weighting them

TABLE 2
Unigue Cues Accessed.

	Subject Pool				
Set	QCF	MBA	FRB Finance	FRB No-Fin	F-Test
UNCONSTRAINED	7.06	7.70	7.62	7.55	1.26
CONSTRAINED	6.76	6.98	7.29	7.05	.50
CLOCK	6.18	5.62	6.14	6.46	.84
F-Test	1.24	16.88**	11.96**	1.87	

Note. *,** denotes significance at the 5%, 1% level.

The table reports the average of the number of unique information cues accessed across subject pools and fund sets. The final column (row) reports an F-test of the null hypothesis of no difference across subject pools (fund sets).

heavily in decision making are two very different things. We will return to the issue of cue importance later.

As subjects move from the *UNCONSTRAINED* to the *CONSTRAINED* set of comparisons, they face a 30-second constraint on their decision time. If no decision is made within this time, the payoff for that screen is equal to 0. Such a discrete constraint increases the speed at which decisions are made, as seen in Figure 2, but does not significantly alter the average amount of information accessed.[20] Thus we cannot reject the hypothesis that subjects do not alter their decision strategy under this constraint; rather they seem to execute the strategy more quickly. Subjects do display, however, a tendency to conserve on information when faced with a continuous opportunity cost of time though the decline in the average number of cues accessed is not significant for the QCF or FRB No-finance students. This finding is consistent with a switch to more efficient decision strategies or heuristics (i.e., those that economize on information) under the *CLOCK* environment.

Overall, subject responses to the time constraints are consistent with an effort-accuracy tradeoff model of decision making. When a discrete time constraint is imposed, as in *CONSTRAINED,* the price of taking an additional second to acquire and process more information is 0 as long as the subject is within the 30-second boundary. In this time span, there is no external tradeoff imposed between effort and accuracy and, under this model, subjects would not be expected to change their behavior unless they were in danger of violating the hard constraint. On the other hand, in *CLOCK*, subjects face a constant opportunity cost of their time ($.033 per second), since more information-intensive strategies require more time, they are more expensive, and we would expect to see a decrease in the amount of information accessed. Both scenarios are broadly consistent with this predicted behavior.

Distribution of Information Cues Accessed

Next we consider which information cues were accessed across subject pools and fund sets. In general terms, we are looking to see if there is any evidence that the various subject pools viewed cues preferentially and if their information acquisition varies with the imposition of the time constraints. Recall that all subject pools report having the most confidence in their use of average return, followed by Morningstar rating (except for QCF, which ranked Sharpe measure in this position); and that QCF reports the highest level of confidence in the use of the more quantitatively sophisticated measures (alpha and Sharpe). Specifically, then, we investigate whether the QCF subjects weight their information acquisition more

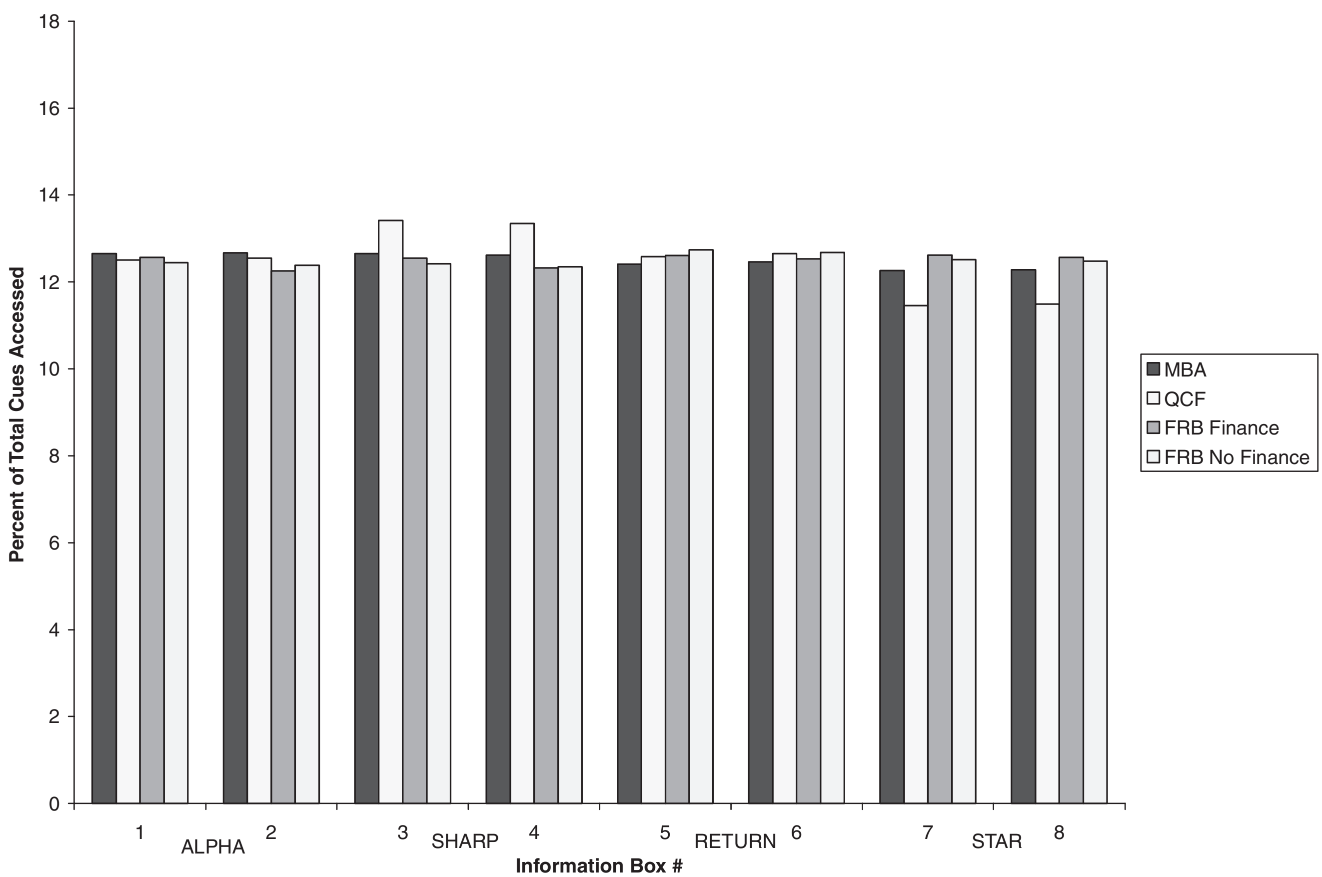

FIGURE 3 Distribution of information accessed: *UNCONSTRAINED* set.

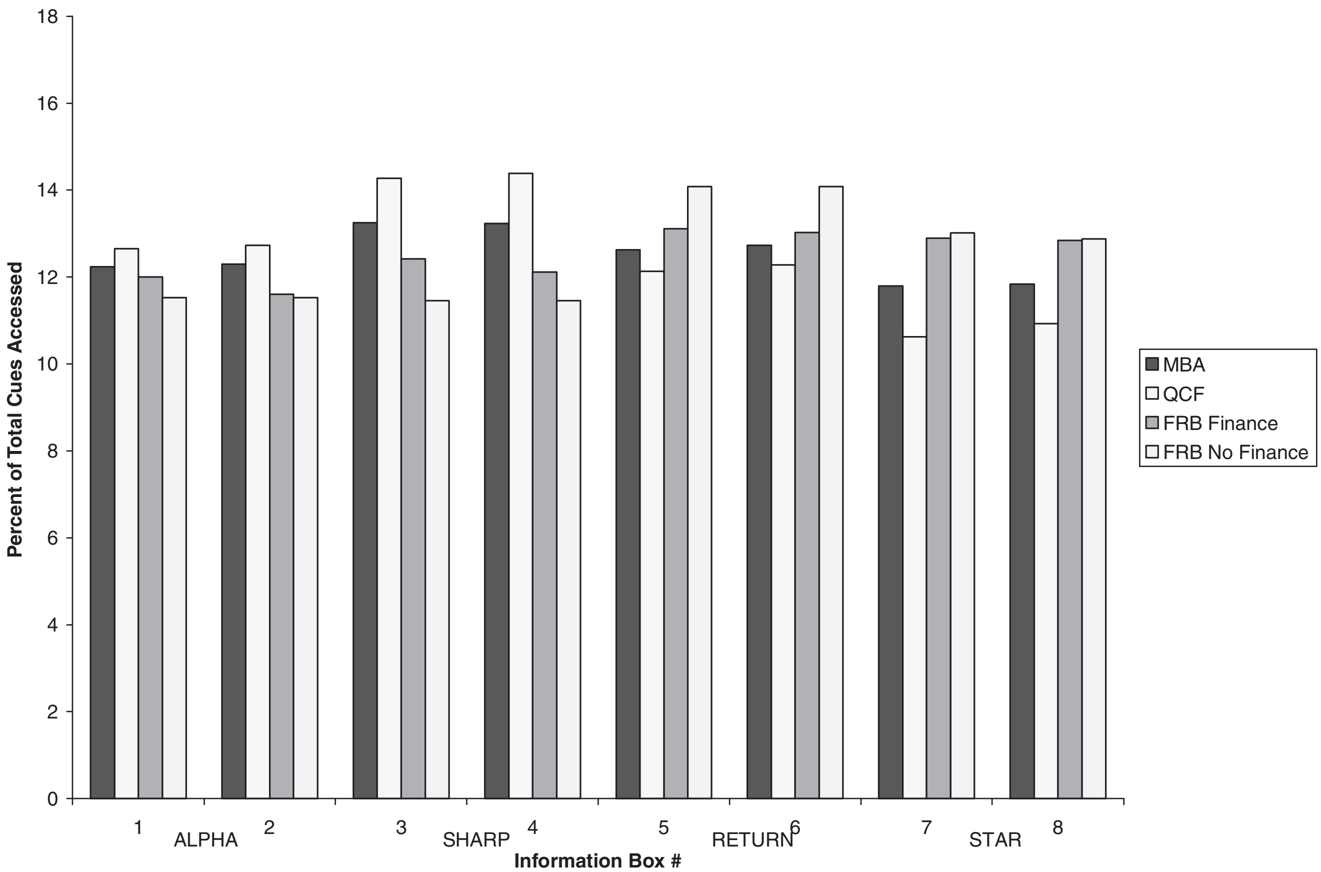

FIGURE 4 Distribution of information accessed: *CONSTRAINED* set.

heavily toward these quantitative measures, versus the other pools, and especially the FRB No-finance pool.

Recall from Table 2 that the average number of cues viewed by each pool in the *UNCONSTRAINED* fund set is over 7, suggesting that all cues are viewed. Figure 3 reinforces this finding by illustrating the frequency distribution of all cues viewed by each pool. These frequencies are computed by aggregating the data for all individuals in each subject pool. Figure 3 suggests that the cues are weighed equally, at least in terms of acquisition, with each cue viewed roughly 12.5% of the time by each pool. The QCF subjects display slightly more acquisition of Sharpe measures, and less acquisition of Morningstar ratings relative to the other cues but these deviations from equal viewing frequency of 12.5% are statistically insignificant.

Figures 4 and 5 show the same frequency calculations for the time constrained set of fund comparisons (*CONSTRAINED and CLOCK*). With less time, do subjects shift their pattern of information acquisition and, if so, does it change in favor of cues in which they have more confidence? In other words, do subjects respond to the time constraints by economizing on effort and time via a shift toward acquiring particular pieces of information? Indeed, we observe some shifting, even using these aggregate frequencies across all subjects in a pool (i.e., measures that obscure the heterogeneity within each pool). As noted earlier, in the *CONSTRAINED* set of comparisons, subjects in all pools continue to view most of the cues on average. Figure 4 reinforces this finding but shows that the QCF students show a slight tendency to view the Sharpe measure more and the Morningstar rating less. Again the switch to a constant opportunity-cost time constraint in *CLOCK* results in larger behavioral changes. The FRB and MBA subjects clearly shift their attention to average returns and Morningstar ratings and away from alpha and Sharpe while the QCF subjects increase their viewing frequency of the Sharpe measure even more.

Table 3 reports χ^2 tests of the null hypothesis that each information cue is accessed with equal frequency. Consistent with the insight provided by Figures 3–5, we cannot reject the null hypothesis that information acquisition is uniform across the 8 cues in the *UNCONSTRAINED* set. With the imposition of the 30-second constraint, significant differences across cue frequencies are observed for the QCF and FRB No-finance subjects. When every second counts, as in *CLOCK,* we see significant differences in information acquisition across cues for all the subject pools. The QCF, MBA, and FRB subjects respond similarly by modifying their information acquisition strategy to acquire fewer cues under the constant time

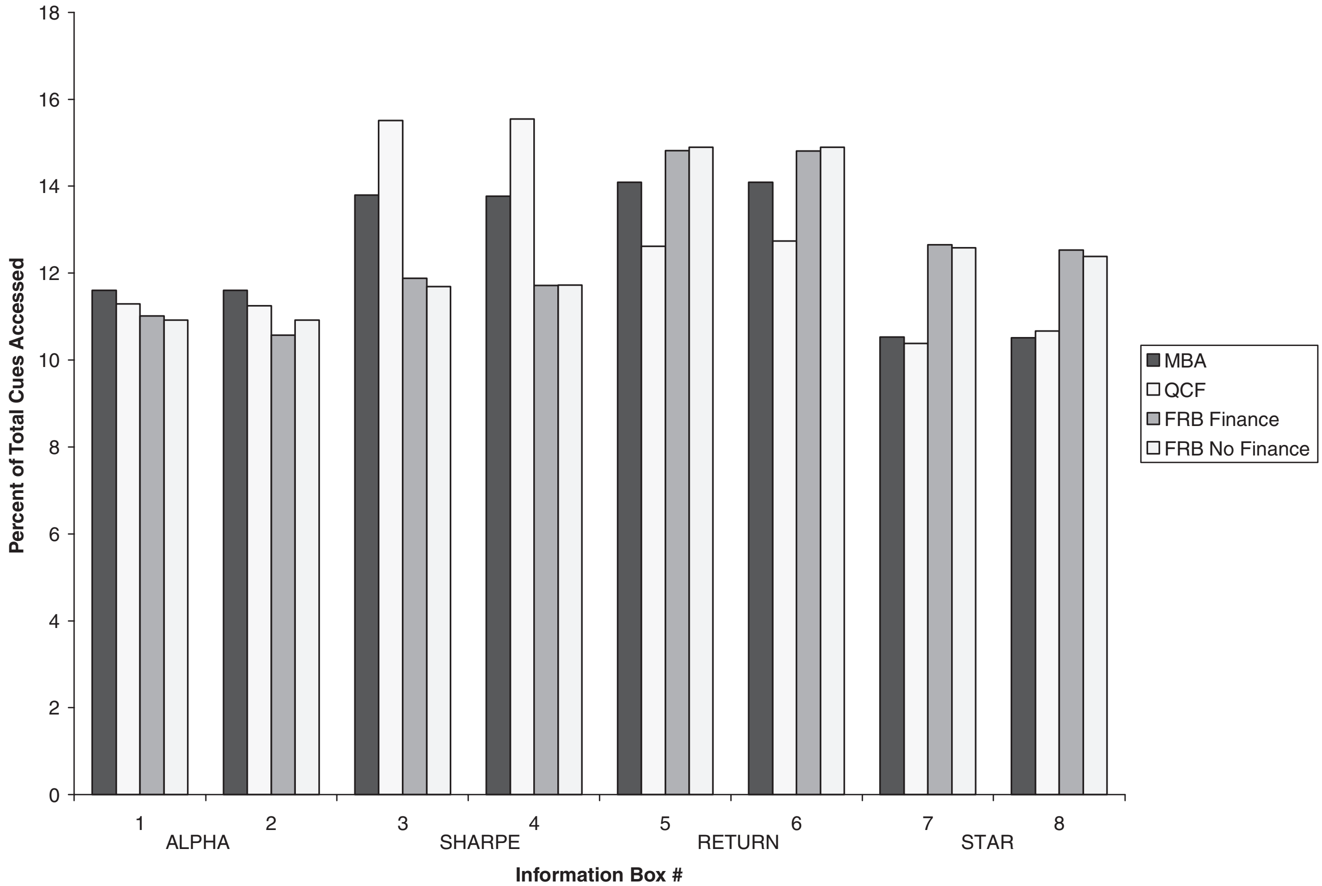

FIGURE 5 Distribution of information accessed: *CLOCK* set.

pressure but differ in the cues they use more often. Recall that all 4 subject pools reported their confidence in their ability to use average return as the highest. Interestingly, the QCF pool actually viewed the Sharpe measure more often. One interpretation of these results is that while this pool was very confident in its use of average return, pool participants did not view it as the most valuable information cue for the task at hand. Whether the MBA and FRB pools share this perception is not clear, but some combination of lower value and lower confidence led to their heavier reliance on average returns.

TABLE 3
Information Accessed.

	Subject Pool			
Set	QCF	MBA	FRB Finance	FRB No Finance
UNCONSTRAINED	6.45	0.88	.72	.34
CONSTRAINED	21.00**	8.70	11.61	21.37**
CLOCK	57.63**	57.23*	81.40**	36.85**
Most accessed information	Sharpe	Return	Return	Return

Note. *,** denotes significance at the 5%, 1% level.

The table reports χ^2 tests of the null hypothesis that each information cue is accessed with equal frequency. We compute the number of times each cue is accessed and compared the percentage of information acquisition within each subject pool and fund set. Only unique acquisitions are included. The final row reports the most frequently accessed cue for each subject pool.

Outcome Analysis

We now turn our attention to a comparison of the subject pools based on the decisions each subject made in the experiment. Each subject viewed the mutual fund pairs in the same order and saw the same underlying data on the performance cues. Therefore, we can analyze and compare subjects' fund choices (i.e., the fund in the pair that they predicted would go on to perform better) and the values of the information cues to attempt to identify and characterize their decision-making rule or strategy. The previous section provided evidence on what information subjects accessed, while this section will address how the subjects use this information to formulate their decisions.

Note that a subject's decision-making strategy may or may not be closely related to his or her information acquisition behavior. For example, the simplest decision strategy is to choose the fund in the pair based on the relative values of only one information cue, say, higher average return. The most efficient (i.e., least costly) method of following this strategy is to view only the two return cues and make a decision.

In the previous section we reported that most participants chose to acquire more than 2 pieces of information, which suggests that they did not follow a pure strategy in a cost-minimizing way. However, it may still be the case that subjects' decisions may have been heavily influenced by the average return cue.

Weighting Informational Cues

The average return cue may have more influence on decisions if subjects use a weighting scheme to combine the information that they access. In fact, typical notions of "optimal" decision making would assume that subjects do just that, using all available information in a way that best fits the data (e.g., an OLS model fit to the underlying data). This is another example of a decision-making strategy, one that can be estimated using conditional logit analysis. Using each fund comparison as a matched-pair observation, we estimate the relation between the probability that a particular fund is chosen and the relative values of its performance measures versus the comparison fund. For each subject we have 60 choices to feed into our model; however, we eliminate the final 20 choices due to the imposition of the continuous time constraint. If subjects respond to this constraint in their adoption of a decision-making strategy (which we documented in the previous section), then it is possible their entire decision strategy changes and this would weaken the power of the estimation.[21] Specifically, the conditional likelihood that fund A is chosen by a subject is:

$$\prod_i \left(1 + \exp\left(-\beta\left(x_i^A - x_i^B\right)\right)\right)^{-1}$$

where i = 1, ..., 20, indexes the fund comparisons and x is a vector of the 4 information cues for funds A and B.[22] This can be estimated by running a logistic regression on a data set with all observations of the response variable set equal to 1, no intercept, and explanatory variables equal to the differences in the mutual fund performance cues. We use maximum likelihood to estimate the β for each information cue. A significantly positive β estimate for, say, alpha, would indicate that the probability that the subject chooses fund A increases if fund A has a higher alpha value than fund B.

Table 4 reports the results of this estimation for 95 subjects in the 4 subject pools. Six subjects did not achieve convergence of the ML estimator while 27 subjects did achieve convergence but had no significant β estimates. In this table we present only the frequency of a significant estimate for each of the four β's and not the relative magnitudes, which would also help in characterizing decision-making strategies. For each of the four subject pools, the most frequently significant cue was average return. This is not surprising given that each pool reported that this is the measure in which they had the most confidence. In contrast, the significance of the

TABLE 4
Conditional Logit.

			FRB	
Cue	MBAs	Quant - Finance	Finance	No Finance
Average Return	64%	37%	51%	64%
Alpha	20	16	8	29
Sharpe	12	16	8	0
Star	8	5	19	7

The following shows the percentage of subjects for which a particular cue was significant in the individual logit out of the number of subjects who achieved convergence in the logits, with at least one significant estimate.

Sharpe measure is a point of distinction among the subject pools. MBA and QCF subjects use Sharpe at a much higher rate than both FRB pools. Indeed, no subject in the FRB No-finance pool exhibited a significant relation between the relative Sharpe measure of the funds and the probability of a fund being selected. This is perhaps not surprising for the QCF subjects who report a high degree of confidence in their use of this measure, but the MBA subjects rank Sharpe measure, on average, as the one in which they have the least confidence. Another surprising result, from this perspective, is the large percentage (20%) of FRB No-finance subjects who exhibit a statistically significant relation between their choices and the relative value of a fund's alpha. These subjects rank alpha as the measure they are least confident in using.

A possible explanation for our results is that confidence in using a measure is not equivalent to the value put on the measure in terms of decision making. If the use of a performance cue in decision making is indicative of a subject's perceived value of the information contained in the cue, then subjects in our experiment exhibit a willingness to use cues they perceive as valuable even if they have little confidence in their ability to use the cue.

The Good Features Heuristic

Rather than weighting all information, and comparing the relative magnitudes of each performance measure, subjects may use a decision heuristic, a rule of thumb that more easily and discretely combines the information to predict which fund will perform better. A plausible example of such a strategy is the "good features" heuristic (Alba and Marmorstein [1987]). In this strategy, each information cue is compared across the two funds, with fund A receiving a point for every dimension on which it has outperformed fund B. Once all cues have been compared, the points for each fund are added up and the fund with more points is the one predicted to perform better in the future, the fund with more "good features." Each information cue is weighted or valued equally and it is the collective weight of the "evidence" that determines the choice.

TABLE 5
Use of the Good Features Heuristic.

	Subject Pool				
Percentage	QCF	MBA	FRB Finance	FRB No Finance	Test
Dominant funds selected	90.45	90.37	89.67	92.44	1.18
Participants always using the heuristic	14.81	13.51	21.28	10.00	.56
Participants using heuristic often	62.96	56.76	72.34	60.00	.80

Note. *,** denotes significance at the 5%, 1% level.

The table reports information on the choice of the dominant fund. At least three information cues rank the dominant fund higher. For each subject pool, the table first reports the percentage of times the dominant fund is selected. Next the table reports the percentage of participants who always choose the dominant fund. Finally, the table reports the percentage of participants who often choose the dominant asset. These participants choose the nondominant fund in three or fewer paired comparisons. The final column reports an F-test of the null hypothesis of no difference across subject pools.

In our setting, the fund with a good feature is the one with higher alpha, Sharpe measure, previous return, or rating (i.e., the fund is better than its competitor on this dimension).[23] In 45 of our 60 fund comparisons presented to subjects, one fund of the pair is dominant; that is, one fund has higher values of 3 or more of the 4 cues. For these decisions, the good features heuristic has a feasible implementation, the dominant fund will be chosen, no matter which cue runs contrary to the other 3. Table 5 examines the use of the good features heuristic by the 4 subject pools. For each subject pool, the percentage of times the dominant fund is selected is close to 90%. This illustrates the relatively rare use of a contrarian strategy in which better past performance is assumed to lead to inferior future performance. There is no significant difference across subject pools but of course this could be because subjects are using different cues to make their decisions and many of the cue comparisons are, by definition, positively correlated.

So, as a first attempt to characterize the use of specific decision strategies, we identified the subjects whose decisions were consistent with the good features heuristic. Across the 4 subject pools, 17% of subjects displayed this behavior and there are no significant differences in this proportion across the pools. It may be the case that over the 45 decisions, some participants were using the good features heuristic but made mistakes. For this reason we also include in Table 5 the percentage of subjects who chose the nondominant fund in three or fewer paired comparisons, that is, the percentage of participants who often use the heuristic. Under this definition, more than 60% of subjects can be said to have used the heuristic, and again the F-test reported in the final column cannot reject the null hypothesis of no difference across subject pools.[24] The fact that a similar percentage of subjects in each pool exhibit behavior that follows the good features heuristic suggests that such decision strategies are used somewhat independently of the relative amounts of prior financial training and self-reported expertise. This is surprising in that we might expect to see subjects with less financial training, and less overall confidence in their ability to use the performance cues, to adopt an equal weighting on all cues rather than any one particular piece of information. One might also conjecture that following such a strategy might be related to risk aversion, such that subjects who are more risk averse might choose to equally weight the informational cues rather than to take a chance by relying more heavily on one cue than another. However, there is no statistically significant relation between the self-reported risk aversion measures of our subjects and their adherence to the good features heuristic.

Accuracy

Another question we ask of the data is whether the accuracy of subjects' performance in the prediction of comparative mutual fund performance differs across subject pools. In other words, does more financial training translate into higher accuracy in this task? The short answer is no. Figure 6 shows the average number of correct choices by for each subject pool for each 10 mutual fund pairings. Recall that pairings 1–20 represent the *UNCONSTRAINED* environment, 21–40 the *CONSTRAINED*, and 41–60 the continuous cost *CLOCK*. The accuracies of the 4 subject pools are remarkably close, almost identical. The variation in the number of correct choices varies significantly with the set of pairings (F-stat=234.08), reflecting largely much lower accuracy in pairs 31–40. All subject groups display a marked decrease in the number of correct predictions for these pairs, 44% versus rates of 63% and higher for the other sets of pairings.

As would be expected given our findings regarding accuracy, earnings across the subject pools are also rather similar. Referring back to summary statistics reported in Table 1, subjects' earnings varied by less than $3 across the subject pools. Though performance as measured by earnings is statistically different at the 10% significance level, the variation is not striking. Interestingly, experimental earnings on the fund choice task are almost identical for the two FRB groups.

In prior sections, we documented that the subject pools show some tendency to focus on different information cues in their decision making. One explanation for the similar performance might be a high correlation between the 4 performance measures. In our fund sample, however, this is not the case. The performance measures are indeed positively correlated but not to a degree that would induce such similar outcomes in the face of differing strategies. For example, the correlation between the choices of a subject using a simple average return strategy (i.e., pick the fund with the higher average return) has a correlation of .25 with a simple Morningstar strategy, .19 with a simple alpha strategy, and .20 with a simple Sharpe measure strategy. Other correlations are of

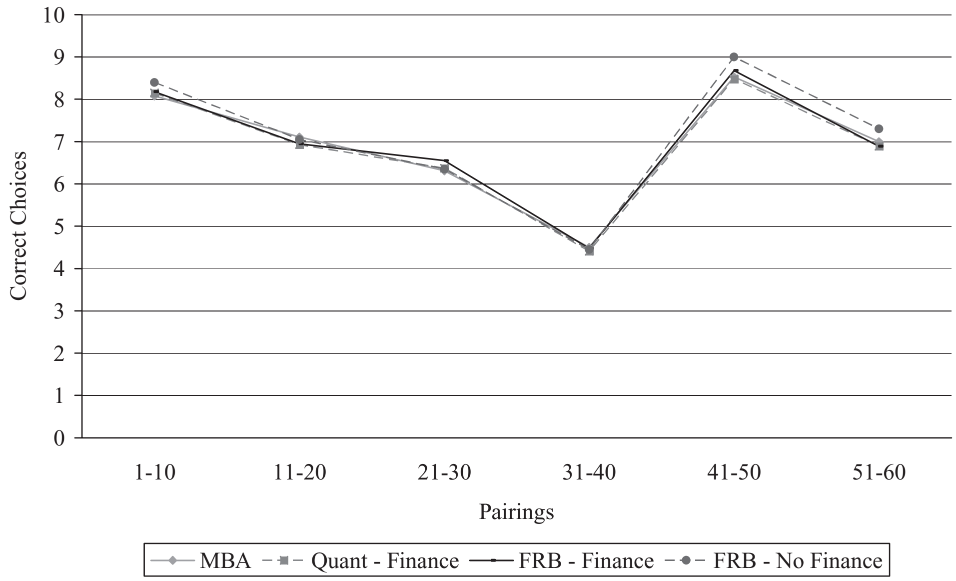

FIGURE 6 Correct Choices x Subject Group.

similar magnitude. Therefore, the similarity in accuracy does not seem to be an artifact of the particular performance measure values used in the experiment but rather is indicative of more general similarities in decision making across the pools.

The relatively high rates of accuracy overall do not necessarily translate into ability to predict mutual fund performance in the real world. Recall that our metric for an accurate choice is based on a regression model fit to the universe of funds underlying the specific sample chosen for use in the experiment. Thus there is much less noise in our model of future performance than exists among actual funds in practice. The important aspect of the accuracy issue for our study is the comparison across groups. MBAs, though deciding more quickly, made choices that were just as accurate in this environment as those of the other groups. FRB employees with no finance training also did not suffer in their performance even though they had less expertise in using and understanding the performance measures.

CONCLUDING REMARKS

In this paper we report the results of an examination of individuals' decision-making processes when evaluating mutual funds to predict future performance. Our results indicate that while there are important differences across the 3 subject pools considered, there are also important similarities. Although the MBA students made their decisions faster than the QCF students or the FRB employees, there was no significant difference in accuracy and experimental earnings across the 4 pools. In addition, although there was no significant difference in the number of information cues accessed: the QCF students appeared to rely most on the Sharpe measure while the MBA students and the FRB employees relied more heavily on the historical return. Our goal was to provide insight into how people make financial decisions. While our subjects were very different in terms of training, financial education, and experience, the outcomes in terms of performance and decision making were generally similar across subject pools. Particularly in light of the current financial crisis, which has resulted in massive losses of wealth, more work investigating how people make financial decisions is required. There remains a great deal to be learned about how education and training can be used to promote prudent financial decision making.

ACKNOWLEDGMENTS

The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. The authors appreciate the helpful comments of Jerry Dwyer, Ann Gillette, Ihsuan Li, and participants in the Wesleyan College workshop, Eastern Finance Association Annual Meetings, and Financial Management Association Annual Meetings. The authors also gratefully acknowledge the research assistance of Maria Alejandra Gonzalez, Lee Cohen, Budina Naydenova, Shalini Patel, and Pam Frisbee, and the financial support of the Federal Reserve Bank of Atlanta.

NOTES

1. For representative examples, see Grinblatt and Keloharju [2000] and Odean [1998].
2. An example of the behavioral bias perspective is Barber and Odean [2001]. Examples of the

decision-cost perspective can be found in Payne, Bettman, and Johnson [1993].

3. Some research has associated the term heuristic with biased decision-making (Tversky and Kahneman [1974]). Our use of the term heuristic as a useful decision tool for solving problems in uncertain environments is consistent with the historical meaning of the word and recent research in decision making (Gigerenzer, Todd, and ABC Research Group [1999]).
4. Collins, Gong, and Hribar [2003] also note that their results could be driven by an informational advantage.
5. See, for example, Hunton and McEwen [1997], who show that more accurate financial analysts use directive information search, whereas less accurate analysts use sequential search.
6. We do not attempt to document or study dynamic learning in this experiment. The goal is not for subjects to learn the most accurate method of predicting returns. Indeed, prior research shows that learning a linear function (such as that involved in predicting asset returns) is more difficult than it would seem to be, particularly if there is a random error component (Klayman [1988]).
7. See Del Guercio and Tkac [2002] for just one example.
8. Due to internal policy, we did not compensate the FRB employees with cash. As described subsequently, they earned "Fed dollars" that they could use to purchase a variety of prizes.
9. All instructions are included in the Appendix to this paper.
10. Several pretests indicated that 2 minutes was not a binding time constraint. Data for these pretests are not reported here as the subjects were not compensated.
11. Instructions for the second and third sets of comparisons are short (approximately half a page) and are identical to the first set except for the change in time constraint. To allow participants to proceed at their own pace, we did not require that all participants start a subsequent set of comparisons at the same time. When distributing additional instructions, an experimenter circled to room to answer any questions raised by participants and to ensure that they understood how the procedures were changing.
12. As no feedback was provided until the conclusion of the experiment, the method of evaluation is irrelevant to subjects' behavior in the experiment. Furthermore, subjects' earnings were actually higher with the simulated data because of the reduction in noise.
13. The complete documentation for the software is available at http://www.cebiz.org/mouselab.htm.
14. See, for example, the collection of papers edited by Payne, Bettman, and Johnson [1993].
15. Researchers have documented the benefits of the mouse as a pointing device (Card, Moran, and Newell [1983]). It is easy to learn to use, can be moved quickly, and is associated with low error rates.
16. For example, the percentage of women enrolled in U.S. MBA programs in 2004 was 30%, and 57% of MBA hopefuls taking the GMAT exam were between the ages of 24 and 27. Similarly, the current QCF cohort at Georgia Tech is 81% male with an average age of 24.
17. There is a statistically significant difference in the median pool response regarding the attractiveness of mutual funds as an investment. This likely reflects the difference between FRB pools with the no-finance group finding funds to be slightly less attractive, though they are still rated a 6.8.
18. All analysis reported subsequently excludes the last pair in each set of funds. Data for these three observations were not saved inadvertently to the data output files during the experimental sessions. We have no reason to believe that these three observations would impact our results in any way.
19. Since MouseLab records all mouse clicks, data on repeated accessing of various cues are available and could be analyzed if necessary.
20. In Bonferroni pairwise t-tests (not reported), the average number of cues accessed is not significantly lower across these two fund sets.
21. Unfortunately, 20 choices is too small a data set on which to get convergence of the Maximum Likelihood estimator that we use. Therefore, we do not analyze the choices made under *CLOCK* with this method.
22. For a derivation of this conditional likelihood based on a latent variable model of paired choice, see Maddala [1983].
23. This implicitly assumes that mutual fund performance exhibits momentum that "good" performance will continue.
24. We investigated whether deviations from the good features heuristic for those who used it often were suggestive of any kind of pattern, such as always going with the fund with higher past return. We could find no evidence consistent with consistent departures from the decision tool.

REFERENCES

Ackert, L. F. and B. K. Church. "The Effects of Subject Pool and Design Experience in Experimental Asset Markets." *Journal of Psychology and Financial Markets,* 2, (2001), pp. 6–28.

Alba, J. W. and H. Marmorstein. "The Effects of Frequency Knowledge on Consumer Decision Making." *Journal of Consumer Research,* 14, (1987), pp. 14–26.

Anderson, J. R. “Acquisition of Cognitive Skill.” *Psychological Review,* 89, (1982), pp. 369–406.

Anderson, J. R. “Skill Acquisition: Compilation of Weak-method Problem Solutions.” *Psychological Review,* (1987), 94, 192–210.

Ball, S. B. and P.-A. Cech. “Subject Pool Choice and Treatment Effects in Economic Laboratory Research.” In *Research in Experimental Economics* (pp. 239–292), Isaac, R. M. (ed.), (1996), Greenwich, CT: JAI Press.

Barber, B. and T. Odean. “Boys Will Be Boys: Gender, Overconfidence and Common Stock Investment.” *Quarterly Journal of Economics*, 116, (2001), pp. 261–292.

Card, S. K., T. P. Moran and A. Newell. *The Psychology of Human-Computer Interaction,* (1983), Hillsdale, NJ: Lawrence Erlbaum.

Chase, W. G. and H. A. Simon. “Perception in Chess.” *Cognitive Psychology,* 4, (1973), pp. 55–81.

Chi, M. T., P. J. Fletovich and R. Glaser. “Categorization and Representation of Physics Problems by Experts and Novices.” *Cognitive Science,* 5, (1981), pp. 121–152.

Collins, D. W., G. Gong and P. Hribar. “Investor Sophistication and the Mispricing of Accruals.” *Review of Accounting Studies,* 8, (2003), pp. 251–276.

Del Guercio, D. and P. Tkac. “The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds Versus Pension Funds.” *Journal of Financial and Quantitative Analysis*, 37, (2002), pp. 523–558.

Ericsson, K. A., R. Th. Krampe and C. Tesch-Romer. “The Role of Deliberate Practice in the Acquisition of Expert Performance.” *Psychological Review,* 100, (1993), pp. 363–406.

Gigerenzer, G., P. M. Todd and ABC Research Group. *Simple Heuristics that Make Us Smart*, (1999), New York: Oxford University Press.

Grinblatt, M. and M. Keloharju. “Investment Behavior and Performance of Various Investor Types: A Study of Finland’s Unique Data Set.” *Journal of Financial Economics*, 55, (2000), pp. 43–67.

Hunton, J. E. and R. A. McEwen. “An Assessment of the Relation Between Analysts’ Earnings Forecast Accuracy, Motivational Incentives and Cognitive Information Search Strategy.” *Accounting Review,* 72, (1997), pp. 497–515.

Johnson, E. J., J. W. Payne, D. A. Schkade and J. R. Bettman. “Monitoring Information Processing and Decisions: The MouseLab System.” Unpublished manuscript, (1991), Center for Decision Studies, Fuqua School of Business, Duke University, Durham, NC.

Klayman, J. “On the How and Why (Not) of Learning from Outcomes.” In *Human Judgment: The Social Judgment Theory Approach*, Brehmer, B. and C. Joyce (eds.), (1988), Amsterdam: North Holland Publishers.

Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics*, (1983), New York: Cambridge University Press.

Odean, T. “Are Investors Reluctant to Realize Their Losses.” *Journal of Finance*, 53, (1998), pp. 1775–1798.

Payne, J. W., J. R. Bettman and E. J. Johnson. *The Adaptive Decision Maker*, (1993), New York: Cambridge University Press.

Rikers, R., W. Te Winkel, S. Loyens and H. Schmidt. “Clinical Case Processing by Medical Experts and Subexperts.” *Journal of Psychology,* 137(3), (2003), pp. 213–223.

Ste-Marie, D. M. “Expert-novice Differences in Gymnastic Judging: An Information-processing Perspective.” *Applied Cognitive Psychology,* 13, (1999), pp. 269–281.

Simon, H. A. “Invariants of Human Behavior.” *Annual Review of Psychology,* 41, (1990), pp. 1–19.

Tversky, A. and D. Kahneman. “Judgment Under Uncertainty: Heuristics and Biases.” *Science,* 185, (1974), pp. 1124–1131.