The Learning (and not) of Effort and Accuracy Tradeoffs

JEFFREY S. LARSON*

* Jeffrey S. Larson is an assistant professor of Business Management, Marriott School of Management, Brigham Young University, 681 TNRB, Provo, UT 84606 (Phone: 801-422-2266; email: jeff_larson@byu.edu). This research was funded in part by the Russell Ackoff grant. This paper is based on the author's dissertation at the Wharton School of Business, University of Pennsylvania. The author wishes to thank Robert Meyer for serving as dissertation chair, Lisa Bolton, Deborah Small, and Uri Simonsohn for serving as committee members, and Joseph Redden for experimental design ideas.

When making decisions, consumers balance their desire for accuracy with their desire to conserve mental effort. This assumption pervades the consumer behavior literature. Despite its ubiquity, no research has examined the manner in which consumers come to learn the optimal balance of effort and accuracy for a given decision. By examining this dynamic process, the current research demonstrates that consumers insufficiently adjust original effort level when explicit changes to the task are made. This bias is significant because it implies that consumers are not as adaptive as previously thought. Three studies demonstrate the bias and show the unique contribution of three causal mechanisms.

For obvious reasons, deciding on a house is likely to inspire more deliberation than deciding on a brand of peanut butter. But precisely how much or how little cognitive effort should an individual invest in either of these decisions? Answering that question requires consideration of multiple factors, including the relative utilities of the available products, the cost of cognitive effort (both in negative utility of thinking and in opportunity cost terms), etc. It is a complex optimization problem that is also fundamental—decision makers have to solve it hundreds of times a day.

Despite its fundamental nature, we still know relatively little about how this metacognitive decision input develops. Several frameworks exist to explain the mechanics of the tradeoff, for example explaining how a decision maker can, given a specified balance of desired effort and accuracy, use various decision strategies to balance these competing demands (Beach and Mitchell 1978; Payne, Bettman, and Johnson 1988). These frameworks come with no formalized knowledge about how decision makers determine the precise amount of cognitive effort merited by a decision task, much less how they learn to map that given level of desired effort and accuracy onto the correct decision strategy. Instead, they seem to imply that, where this learning is necessary, it takes place quickly and without detriment to the decision process.

This research proposes an alternate hypothesis. Adapting to a change in task constraints requires metacognitive learning on multiple fronts. First, the decision maker may not immediately know the extent to which a constraint change affects the desired balance of effort and accuracy. Second, the mapping of a new effort-accuracy level to an available decision strategy may also require learning. Third, the fuzzy nature of the feedback in this domain might serve to reinforce rather than correct poor adaptation. Thus, I hypothesize that decision makers will adapt poorly to task constraint changes. The core contribution of this research is the discovery of the nature of this poor adaptation as well as its multiple contributing mechanisms.

The knowledge here exposited serves functions both theoretical and practical. The dynamics of consumer decisions have increasingly interested researchers (Erdem and Keane 1996; Meyer, Erdem, Feinberg, Gilboa, Hutchinson, Krishna, Lippman, Mela, Pazgal, Prelec, and Steckel 1997; Novemsky and Dhar 2005). As the tradeoff of effort and accuracy forms a primary input to almost any decision, understanding the dynamic workings of this balance fills a major gap in our theoretical knowledge of decision making. Of more practical importance, the way consumers adapt to changes in decision environments, whether caused by sales promotions, the introduction of a new product, or a structural change in the competitive landscape, can affect the results of such changes.

This research is presented in three sections. The first section relates prior research in the domain to the current research and presents justification in terms of three mechanisms for the hypothesis that decision makers adapt poorly to task constraint changes. The second section presents the results of four studies that demonstrate evidence of the central hypothesis and for each of the three hypothesized mechanisms. Finally, in section three, I conclude with suggestions for future research.

A MODEL OF CONSTRAINED DECISION MAKING

Figure 1 depicts a stylized model of decision making for the consideration of effort and accuracy tradeoffs. No previous research has made an integrated examination of the tradeoff at all the stages depicted in this organizing model. The majority of the research in this area has focused on the second stage depicted in the model—individual decision strategies and their use in the balance of the two conflicting goals. Very little research has systematically examined how decision makers determine the optimal balance of the two goals, and even less has addressed the role of feedback and learning in this domain. Because of this gap, the mechanisms surrounding the dynamic evaluation of effort and accuracy goals, application of decision strategies, and outcome evaluation are largely unknown.

Insert figure 1 about here

Task Constraints and Goals

Simon's work on bounded rationality introduced to the literature the consideration of decision goals other than accuracy (Simon 1957). Since then, many other situational decision goals have been examined, including the reduction of cognitive effort (Tetlock 1979), accountability to others (for a review, see Lerner and Tetlock 1999), minimization of negative emotion (Bettman, Luce and Payne 1998), as well as goals originating from individual differences, such as dogmatism (Tetlock, Stitka and Boettger 1989), need for cognition (Cacioppo, Petty, Feinstein and Jarvis 1996), need for precision (Viswanathan 1997), and need for cognitive closure (Webster and Kruglanski 1994).

As a result of the wide variety of decision goals, researchers have pushed the position that decisions should be evaluated relative to the baseline decision goals, and not some externally imposed evaluative criteria (Mellers, Schwartz and Cooke 1998). But with the ever-growing stream of literature demonstrating the spurious nature of many decision goals (Zajonc 1965; Bargh and Chartrand 1999; Bargh, Gollwitzer, Lee-Chai, Barndollar, and Troetschel 2001; Fitzsimons and Bargh 2003), the wisdom of the decision goals themselves ought also to be evaluated. This research explicitly examines the quality of the effort and accuracy goals set by decision makers and examines the mechanism underlying them.

In past research, the efficiency goal (or the goal to minimize effort) is often increased through the application of time pressure, often by imposing a hard and fast deadline by which time a decision must be made (Ben Zur and Breznitz 1981; Edland and Svenson 1993; Payne et al. 1988). While this manipulation has proved useful to the study of decision strategies, it removes the goal formation stage of decision making. In natural decision settings, decision makers set goals according to their interpretation of (mostly) self-imposed constraints. For example, a grocery shopper may want to finish a trip within a half hour, and thus imposes constraints on each purchase decision within that trip. Experimental decision settings that apply deadlines may not necessarily affect decision goals at all, but instead only cause decision makers to perform a constrained search for decision strategies that will be completed before the deadline. The current research includes goal formation in the decision process by using reward-based constraints, which requires participants to interpret the constraints and set effort and accuracy goals according to that interpretation.

The literature on goal setting provides some insights into how these goals might be set. Results by Mano (1990) indicate that when individuals are asked to give explicit anticipatory goals for accomplishing tasks in a given time frame, they set the goals using an anchoring and adjustment mechanism (though he finds no evidence of the underadjustment bias that typically accompanies the mechanism). Effort and accuracy goals typically function at a lower level of consciousness, but the same mechanism is likely to apply in the case of changing task constraints. The effort and accuracy goals from the decision maker’s first encounter(s) with the decision task will form an anchor from which goals will be adjusted to account for the task constraint change. The use of the anchoring and adjustment mechanism should result in the expected underadjustment of effort and accuracy goals, which will result in poor adaptation to the constraint changes (Epley and Gilovich 2001; Tversky and Kahneman 1974). I call this contributor to poor adaptation *goal persistence*. When a new task constraint changes the optimal balance of effort and accuracy, the new goals are anchored on the previous goal strengths, causing the older goal strengths to persist.

Decision Strategies

Once decision makers set effort and accuracy goals for a particular decision, they must find the decision strategy that best fits the set goals. Two frameworks deal explicitly with strategy selection. Beach and Mitchell (1978) focus on the use of either intuitive, unaided analytic, or aided analytic strategies. The current research examines the effort exerted within the unaided analytic decision strategy domain, thus limiting the usefulness of this framework. Of more direct application is the framework proposed by Payne, Bettman and Johnson (1993). According to their framework, decision makers must map each strategy in their repertoire on a two-dimensional grid formed by relative effort and relative accuracy at the axes. Given that the position of each decision strategy on the grid varies across decision environments, every new decision environment requires some learning. This places large demands on decision makers' metacognitive capacity. Busemeyer and Myung (1992) find that decision makers do have some ability to predict the payoffs of various decision strategies, but require further learning to fine-tune this knowledge. This fine-tuning is subject to some biasing influences.

Studies of decision strategy selection in the multi-cue probability learning literature show that changes to a cue environment often cause perseveration—that is, the continued use of an ill-suited strategy, even when the participant has received explicit feedback that the cue is invalid

(Restle 1962). This perseveration is likely also to occur in the strategy selection phase when a change in the decision environment requires the selection of a new decision strategy. Instead of selecting a new strategy to balance the new effort-accuracy goals, the decision maker is likely to continue using the same decision strategy that was used previously, which will cause poor adaptation to the change in the decision environment. This is called *biased selection*, and further contributes to poor adaptation.

Outcome Evaluation and Feedback

Research on the dynamic impact of decision outcomes on future decisions is sparse. Recently, Rieskamp and Otto (2006) showed that decision makers can indeed incorporate feedback to learn to apply a more accurate decision strategy, but they do not focus on the tradeoff of effort and accuracy. Novemsky and Dhar (2005) show how decision goals can be impacted by the result of a previous decision, but they do not examine effort and accuracy goals. Some previous work has examined the effect of feedback on effort-accuracy tradeoffs, but does so in a static context (Creyer, Bettman and Payne 1990). They find that explicit accuracy feedback affects the tradeoff, though effort feedback does not, presumably because effort feedback is already salient and accurately assessed by the decision maker.

The dynamic focus of the current research updates those previous findings. Effort, as measured by duration of time (Johnson and Payne 1985; Russo and Carlson 2006), is a perceptual continuum known to display a number of distorting biases documented in the psychophysical literature (Stevens 1957). Chief among these biases, at least for the current research, is perceptual distortion (Sherif, Taub, and Hovland 1958). Perceptions are reference-dependent, so experience under a given effort level establishes a reference point that form the basis of subsequent evaluations (Helson 1964). For example, two people who make a decision in one minute will differ in their evaluations of the efficiency of that decision if one typically spends two minutes on the decision, while the other typically spends 30 seconds.

Evaluation of accuracy is not perceptual, so is not subject to perceptual distortion, but another mechanism causes the same type of bias in the evaluation of decision accuracy. The cognitive evaluation scale is determined by the range and frequency of past experiences (Parducci 1963, 1995). If a decision maker is accustomed to perfect accuracy, 80% accuracy will be evaluated as poor, while a decision maker who usually makes random choices will be elated with 80% accuracy.

Because of these influences, the evaluation scales of both effort and accuracy will be distorted around their previous levels. This *scale distortion* causes poor adaptation to a decision task constraint change because insufficient adjustments to the accuracy and effort levels will be evaluated as sufficient due to skewing of the evaluation scale by prior experience.

HYPOTHESES

The three proposed mechanisms to poor adaptation—goal persistence, biased selection, and scale distortion—all result in the same bias: an insufficient adjustment to the effort and accuracy levels brought on by a task constraint change. The following hypotheses result. (In the following hypotheses and throughout the rest of the paper, I use the term ‘efficiency’ to mean the reciprocal of effort, to avoid the difficulties of describing the goals as being set in the opposite direction. Thus, the conflicting goals become the maximization of accuracy and the maximization of efficiency.)

H$_{1a}$: If a new constraint increases desired accuracy or decreases desired efficiency, decision makers will be less accurate and more efficient than a control.

H$_{1b}$: If a new constraint decreases desired accuracy or increases desired efficiency, decision makers will be more accurate and less efficient than a control.

This is the effort-accuracy underadjustment bias. Though underadjustment is not by itself a new phenomenon, this bias is especially important because of the ubiquity of effort-accuracy tradeoffs. Since these tradeoffs are a crucial part of every decision, this bias has important implications for myriad decision contexts.

The choice of a control group for this hypothesis is not straightforward due to practice effects—experience with a decision environment improves decision performance. Since hypothesis 1 seeks to examine decision making after some amount of practice, the control group should have an equal amount of practice. But to function as a control, this practice must be under a different constraint. For example, if the test group makes initial decisions under constraint A, we examine later decisions under constraint B. The control group should have an equal amount of practice, so we cause the control group to make initial decisions under constraint B, then examine later decisions under the same constraint. This procedure would be objectionable if practice effects were constraint-specific—the control group would have an unfair advantage due to extra practice under constraint B. In fact, empirical results will show that practice effects tend not to be constraint-specific. In addition, the test group decisions under constraint B will be compared with initial decisions from constraint B. Results will show that the effort-accuracy underadjustment bias is often strong enough to overcome any practice effects.

Because the studies presented here use an objective measure of performance (usually payment for correct answers), I can further hypothesize about the implications of this underadjustment bias. Despite the fact that practice effects are not constraint-specific, performance for constraint changers will be significantly worse than non-changers. The underlying cause of this poor performance, poor choice of effort levels rather than practice effects, is important because it demonstrates the importance of this bias in everyday choice situations.

H_2: Constraint changers will perform worse than the appropriate control group.

In addition to demonstrating the effort-accuracy underadjustment bias, this research demonstrates the independent contribution of three separate mechanisms to this bias.

H_3: Goal persistence causes underadjustment of relative goal strengths.

H_4: Biased selection causes use of previous (and poorly suited) decision strategies.

H_5: Scale distortion exaggerates the perceived size of effort and accuracy adjustments.

The studies are laid out as follows. All three studies show evidence of hypothesis 1. Study 1 will further address hypotheses 2 and 5. Study 2 will address hypothesis 3. Finally, study 4 will give evidence for hypothesis 4.

STUDY 1

Stimuli

The purpose of this study is to demonstrate the effort-accuracy underadjustment bias and show that scale distortion contributes to the bias. Because this research seeks to understand effort and accuracy tradeoffs, any experimental task needs to enable measurement of both dimensions. Consistent with prior literature, response times were taken as a proxy for cognitive effort in this and all subsequent studies (Bettman, Johnson, and Payne 1990; Johnson and Payne 1985). Because of the need to measure accuracy, all of the experimental tasks used here are problems in which the correct answer is knowable, but whose solution is cognitively taxing and thus amenable to the use of shortcut heuristics. In this study, the task was a two-tier pricing problem, where participants were asked to determine which of three cell phone pricing plans would yield the lowest price. They were given the month's usage, the base price (for 500 or fewer minutes), and the overage rate for three different plans. These values were randomly generated in every set (usage between 1 and 1500 minutes; base rate between $29 and $61; overage rate between $.01 and $.15 per minute). Figure 2 shows an example stimulus from the experiment. Participants performed the task in four different two-minute phases, preceded by a 45 second training phase to accustom participants to the controls. For each phase, participants were instructed as to how bonus pay would be administered, and answered as many or as few questions as they wished in the two-minute phase.

Insert figure 2 about here

In order to test participants' abilities to adapt to changing task constraints, two different payment schemes were created. In the accuracy condition, participants were told they would be paid $1 times their accuracy rate, plus $.01 per correct answer. In the efficiency condition, they were paid $.20 times their accuracy rate, plus $.04 per correct answer. Half of the participants started in the efficiency condition, the other half in the accuracy condition. After two test phases, half of the participants changed to the opposite condition, while half remained in their original condition. This yielded four conditions: 1) efficiency-efficiency (E-E), 2) efficiency-accuracy (E-A), 3) accuracy-accuracy (A-A), and 4) accuracy-efficiency (A-E).

After every phase, before they were told the results of their performance, participants were asked their perceptions of both their speed and their accuracy, on a 1 to 7 scale. According to the scale distortion hypothesis, participants' perceptions of their speed and accuracy would be affected by their initial constraints.

In total, 84 participants took part in the study, for which they were paid $10 for participation in this and two other studies, plus an average of $5.15 in bonus money earned during this task.

Results

The hypothesized underadjustment bias should result in decisions by constraint changers that are slower and more accurate, or faster and less accurate, than participants remaining in the original condition. Across the four phases, response times tend to decrease slightly (test of linear trend, $t(8719) = 11.9$, $p < .0001$), while accuracy rates remain relatively constant (test of linear trend, $t(8719) = 1.3$, $p = .20$). This is caused by a general (not a constraint-specific) practice effect, as will be shown in subsequent analyses.

This study was designed to test hypotheses 1, 2 and 5, so we begin by looking for evidence of the effort-accuracy underadjustment bias. An ANOVA model on the response time data was fitted using the mixed procedure in SAS. Because participants determine their own response speed, the number of responses is not balanced. The mixed procedure allows the model to account for the correlated nature of within-participant responses, so that data from participants making more responses than average are not overweighted. The accuracy data was analyzed using the nlmixed procedure, which allows for subject-level intercepts in the logistic regression.[1] Additionally, it should be noted that response times were transformed by a natural logarithm, consistent with recommended procedures for timing data (Allison, 1984; Kalbfleisch and Prentice, 1980). Also, p-values reported here are generally one-tailed, as the direction of differences was hypothesized a priori.

Conditions 2 and 4 changed constraints in phase 3, so we first examine phase 3 decisions. Figure 3 plots the average response times and average accuracy rates by condition for phase 3 (the first post-constraint-change phase) of the study. As hypothesized, participants in the E-A condition made decisions more quickly (M = 3.9 s) than A-A condition participants (M = 4.6 s), $F(1, 240) = 16.1$, $p < .0001$. As a result of this quicker response time, E-A participants made decisions with lower accuracy (M = 79.5%) than A-A participants (M = 90.4%), $t(80) = 3.0$, $p < .01$. Also as hypothesized, participants in the A-E condition made decisions more slowly (M = 3.0 s) than E-E condition participants (M = 2.4 s), $F(1, 240) = 55.6$, $p < .0001$. As a result of this slower response time, A-E participants made decisions with higher accuracy (M = 77.8% vs. M = 72.8%), though this difference was not significant, $t(80) = 1.1$, $p = .13$. As a result of this underadjustment bias, E-A participants on average earned marginally less money (M = \$0.96) than A-A participants (M = \$1.04), $F(1, 240) = 2.39$, $p = .06$, and A-E participants (M = \$1.03) on average earned significantly less money than E-E participants (M = \$1.14), $F(1, 240) = 3.83$, $p = .03$.

Insert figure 3 about here

----

[1] ANOVA’s weighted such that each participant’s data receive equal weighting yielded similar but not identical results due to decreased statistical sensitivity.

The test group phase 3 performance is also compared with control group phase 1 performance. The money earned in phase 3 by E-A participants (*M* = \$.96) is smaller than the money earned by accuracy condition participants (*M* = \$1.02) in phase 1, $F(1, 240) = 1.95$, $p = .08$. This was due to E-A participants' faster pace (*M* = 3.9 s vs. *M* = 6.0 s), $F(1, 240) = 126.9$, $p < .0001$, and lower accuracy (*M* = 79.6% vs. *M* = 91.2%), $t(80) = 3.8$, $p < .001$. In this comparison, the effort-accuracy underadjustment bias was strong enough to cause poor performance by E-A participants despite an extra two phases of practice. In the opposite direction, significant differences in performance were not found, due to the general tendency of respondents to increase response speed across the four phases.

While the results from phase 3 demonstrate the underadjustment bias, results from phase 4 demonstrate evidence of the enduring nature of the bias. A-E participants (*M* = 2.9 s) continue going slower in phase 4 than E-E participants (*M* = 2.2 s), $F(1, 240) = 69.0$, $p < .0001$. This again resulted in a higher accuracy level (*M* = 76.5% vs. *M* = 73.3%), though again this difference was not significant $t(80) = 0.7$, $p = .24$. E-A participants (*M* = 3.7 s) continue going faster than A-A participants (*M* = 4.6 s) in phase 4, $F(1, 240) = 126.9$, $p < .0001$. This higher speed should have resulted in a lower accuracy, but accuracy for these participants was anomalously high. This anomalous high accuracy was the result of random error, as attested by the fact that, adjusted for response time, accuracy levels are not significantly different than A-A phase 4 accuracy $t(80) = .8$, $p = 0.43$ (note that this also provides evidence that practice effects are not constraint-specific). Additional evidence that this result is anomalous comes from study 2, where this high accuracy in phase 4 of the E-A condition does not replicate. Of course, due to this anomaly, payment differences were not significantly different between E-A and A-A participants in phase 4, but the A-E participants (*M* = \$1.06) earn significantly less than E-E participants (*M* = \$1.24) in phase 4, $F(1, 240) = 11.04$, $p < .001$.

One of the hypothesized causes of the underadjustment bias is scale distortion. The perceived speed and accuracy responses will help determine if scale distortion indeed occurs. After each phase of the study, before told their actual accuracy and number of responses, participants were asked about their perceptions of their speed and accuracy in the previous phase. They responded on a 7-point semantic scale, anchored on the endpoints by ‘very slow’ and ‘very fast’, or by ‘very inaccurate’ and ‘very accurate’. The survey data from one participant was excluded for marking a ‘1’ on every scale in every phase (despite being one of the faster responders). Direct comparisons of perceived speed and accuracy are not meaningful, since actual speeds and accuracy levels differ among the key comparison groups. To test for differences in scale use on the two variables, perceived speed and perceived accuracy measures were subjected to a regression on actual speed (number of questions answered in that phase, log transformed due to right-skew) and actual accuracy (percent correct), respectively. (Quadratic and other functional forms in the regression were checked. All results were similar). Figure 4 presents these adjusted speed and accuracy reports from participants in phase 3. Note that in phase 3, E-A participants ($M = 2.8$) report a significantly slower perceived speed than A-A participants ($M = 4.8$), $t(79) = 6.3$, $p < .0001$. This perception continues to the fourth phase ($M = 2.4$ vs. $M = 4.5$), $t(79) = 6.5$, $p < .0001$. In the opposite direction, A-E participants ($M = 5.8$) report a significantly faster perceived speed than E-E participants ($M = 5.2$), $t(79) = 2.1$, $p = .02$. This effect also continues into the fourth phase ($M = 5.8$ vs. $M = 5.0$), $t(79) = 2.8$, $p < .01$.

Insert figure 4 about here

On the accuracy side, results follow the same pattern. Adjusted for true accuracy, A-E condition participants (*M* = 4.2) perceive themselves to have a lower accuracy than E-E condition participants (*M* = 5.0), $t(79) = 3.0$, $p < .01$. In the opposite direction, E-A participants (*M* = 4.7) report a marginally higher accuracy than A-A participants (*M* = 4.2), $t(79) = 1.4$, $p = .09$.

Discussion

Study 1 showed multiple layers of evidence for hypotheses 1a and 1b. The effort-accuracy underadjustment bias was found for those facing constraint changes in both directions. The underadjustment caused poor performance, as evidenced by the amount of money earned, in support of hypothesis 2. In one direction, constraint changers performed worse than even comparable phase 1 participants, despite an extra two phases of practice, further attesting to the strength of the bias. The effort-accuracy underadjustment bias is caused by the development of metacognitive knowledge in one condition of the decision environment that is misapplied to a different condition of the same environment. Participants in the E-A condition learned a balance of effort and accuracy that emphasized efficiency. These metacognitions caused them to answer too quickly once they moved into the accuracy condition and to interpret these overly quick responses as being sufficiently slow.

Further evidence of the spoiling influence of the original metacognitions comes from the perceived effort and accuracy levels reported by participants. As predicted, their interpretations of both their speed and accuracy were distorted by their original condition. Those in the E-A condition felt their responses were slow and accurate once they changed to the accuracy phases, when in reality they answered more quickly and less accurately than A-A participants. The strength of this distorting influence is evidenced most by speed responses from E-A participants. Even without adjusting for actual speed (which is significantly faster than A-A participants' speed), their perceived speed is slower (M = 3.3) than A-A participants' perceptions (M = 4.7), $t(79) = 3.63$, $p < .001$. Scale distortion was one contributor to the effort-accuracy underadjustment bias. Study 2 shows the contribution of goal persistence to the bias.

STUDY 2

Stimuli

Participants were told they would be shown the prices of four products from four different stores. Their task was to determine which of the four stores provided the lowest overall price for the four products. The four products were never specified; they were simply labeled, "Product 1", "Product 2", etc. The prices varied between $10 and $115, and were constructed such that the price range for any given product was at most $15, to reflect a realistic price range of a single product across several stores. As in study 1, participants performed the task in four two-minute phases. This study used two conditions, E-A and A-A, the transition from efficiency to accuracy occurring in phase 3.

Instead of playing for money, participants were rewarded points, and were instructed that the two top point-getters would receive $100, which was determined after all 169 participants had completed the study. In the efficiency condition, participants were informed they would receive 10 points for every correct guess and a 50 point bonus multiplied by their percentage correct. In the accuracy condition, they were informed they would receive 2 points for every correct guess and a 250 point bonus times their percentage correct.

To provide evidence of goal persistence, a seemingly unrelated task was added at the end of all phases of the study. This task showed a 10 by 10 grid of white squares, many of which were filled with red circles. Participants were asked to ascertain if the grid contained greater or fewer than 50 red circles (the grid never contained exactly 50 circles). They were instructed that they would be required to make 25 such guesses. Because of goal persistence, participants in the E-A condition were expected to respond more quickly on this task than those in the A-A condition.

Results

Figure 5 plots the average response time and accuracy for each condition by phase. As predicted, those in the E-A condition answered more quickly than participants in the A-A condition, in both phase 3 (M = 6.6 s vs. M = 9.5 s) and phase 4 (M = 6.5 s vs. M = 9.5 s), $F(1, 501) = 151.7$, $p < .0001$, and $F(1, 501) = 151.8$, $p < .0001$, respectively. This was also accompanied by significantly lower accuracy in both phases (Phase 3: M = 63.1% vs. 77.0%, and Phase 4: M = 68.3% vs. 77.0%), $t(167) = 4.1$, $p < .0001$, and $t(167) = 2.7$, $p < .01$, respectively.

Insert figure 5 about here

The more conservative contrasts, comparing E-A condition performance in phases 3 and 4 to accuracy condition performance in phases 1 and 2, respectively, yield the same pattern. Response times are faster for E-A participants in both phase 3 (M = 6.6 s) and phase 4 (M = 6.5 s) than comparable phase 1 (M = 9.1 s) and phase 2 (M = 9.2 s) participant decisions, $F(1,501) =$

117.7, $p < .0001$, and $F(1,501) = 130.4$, $p < .0001$. The accuracy rates also follow the expected pattern in both comparisons ($M = 63.1\%$ vs. $M = 70.9\%$) and ($M = 68.3\%$ vs. $M = 76.1\%$), $t(167) = 2.3$, $p = .01$, and $t(167) = 2.5$, $p < .01$.

It was predicted that E-A participants would respond to the red dot grid task more quickly than A-A participants as a result of higher efficiency goals from the store price task. Indeed, E-A participants ($M = 1.8$ s) answered significantly faster than A-A participants ($M = 2.1$ s), $F(1, 167) = 6.0$, $p = .02$. This also resulted in a significantly different accuracy rate ($M = 72\%$ vs. $M = 75\%$), $z = 2.1$, $p = .02$. Note that this average response time is over three times faster than the fastest average response times from the store price task. This makes it unlikely that scale distortion caused the difference in response times.

Discussion

In addition to providing another demonstration of the underadjustment bias, this study gave evidence of goal persistence (hypothesis 3) as a contributing mechanism to this bias. The red dot grid task is far removed in nature from the store price task, so any difference in decision strategies between the E-A and A-A conditions could not have carried over into the grid to task to cause the differences that occurred. The average response time in the E-A condition was over six seconds in the E-A condition and over nine seconds in the A-A condition. Responses in the grid task were around two seconds, less than a third of the speed of the store price task. It is unlikely that the perceptual distortion from the store price task would have extended to speeds so far removed. Additionally, it *is* likely that goals primed in one context might extend to unrelated tasks, as occurs often in the nonconscious goal pursuit literature. These facts support the assertion that goal persistence is an additional cause of the effort-accuracy underadjustment bias.

STUDY 3

Stimuli

Study 3 will demonstrate hypothesis 4, the contribution of biased selection to the effort-accuracy underadjustment bias. The stimuli for study 3 were identical to those of study 2. Participants were again asked to identify the store providing the lowest overall price on four products. Participants performed the task in two phases lasting four minutes each phase. The same four conditions were used as in study 1 (E-E, E-A, A-A, A-E), with the constraint change occurring between the two phases. Instead of playing for money, participants were instructed to treat the task like a game, and attempt to maximize the number of points earned. Efficiency condition participants were told that they would earn 5 points for every correct answer, plus a bonus of 100 points times their percentage correct. Accuracy condition participants were told that each correct answer would give them 2 points, plus they would receive a bonus of 200 points times their percentage correct.

One additional complication was added to the task. The price of each product was hidden from view until the participants moved the mouse cursor over the box containing that pricing information. This is identical to the Mouselab process tracing methodology (Payne et al., 1993). This allows for examination of processing strategies, which will be important for providing evidence of biased selection of decision strategies.

In total, study 3 collected data from 181 participants, for which they were paid $10 for their participation in the battery of four studies. Data from 26 participants were removed because they made multiple selections without revealing any information from the boxes.

Results

Figure 6 shows the phase 2 mean response times and accuracy levels across the four conditions. Consistent with hypothesis 1, A-E participants (M = 7.3 s) made decisions more slowly than E-E participants (M = 5.5 s), $F(1, 151) = 15.0$, $p = .0001$. As a result, A-E participants (M = 66.0%) were more accurate in their decisions than E-E participants (M = 56.2%), $t(151) = 2.1$, $p = .02$. In the opposite direction, response times also gave support for the bias, with E-A participants (M = 11.4 s) answering more quickly than A-A participants (M = 13.9 s), $F(1,151) = 5.3$, $p = .01$. E-A participants also answer with a lower average accuracy (M = 67.6% vs. M = 74.6%), though this difference was only marginally significant, $t(151) = 1.6$, p = .06.

Insert figure 6 about here

The more conservative test of the bias yields similar results. A-E participants answer more slowly in phase 2 (M = 7.3 s) than efficiency condition participants in phase 1 (M = 5.6 s), $F(1, 151) = 16.2$, $p < .0001$. As a result, accuracy for A-E participants is higher (M = 66.0% vs. M = 58.6%), $t(151) = 1.8$, $p = .04$. In the opposite direction, the bias was not observed on this stronger test, either for response speed (M = 11.4 s vs. M = 11.8 s), $F(1, 151) = .2$, $p = .35$, or accuracy (M = 67.6% vs. M = 70.4%), $t(151) = .6$, $p = .27$. In summary, the bias was significant in both directions, but only strong enough to overcome practice effects in one direction.

The processing data will allow for examination of decision strategies. Several measures were constructed to serve as proxies for decision strategies. Every time a mouse moved over a box containing price information, this was counted as an acquisition. NACQ is a measure of the number of acquisitions made for every decision. TPER tells the average amount of time each acquisition required. Proper analysis of these numbers required a natural logarithm transformation, due to the right-skewed distribution. It was expected that efficiency condition participants would have a lower NACQ and TPER than accuracy condition participants. This would reflect a greater use of simplified decision strategies. Biased selection hypothesizes that E-A participants will continue with lower NACQ and TPER than A-A participants, indicative of continued use of simplified decision strategies. It hypothesizes the same effect in the opposite direction.

Several other measures were constructed from the processing data, consistent with prior literature (Payne et al. 1988). These included a measure for repetition, for attribute-based versus holistic processing, and variance in processing time spent across attributes, across stores and across all boxes. The measure for repetition was dropped because it was highly correlated with NACQ. Previous work often finds that more compensatory strategies lead to more holistic processing and lower variance in time spent across the attributes. Unfortunately, the current study did not yield consistent differences across the conditions. This lack of difference is consistent with prior literature finding that decision makers heavily favor attribute-based processing under all conditions (Russo and Dosher 1983). For this reason, analyses focus on NACQ and TPER.

Consistent with expectations, accuracy condition participants have higher TPER (M = .28 s vs. M = .26 s) and NACQ (M = 33.0 vs. M = 24.9) than efficiency condition participants, $F(2, 5040) = 11.1$, $p < .001$, and $F(2, 5040) = 153.7$, $p < .0001$, respectively. It was hoped that E-A participants would show a lower NACQ and/or TPER than A-A participants in phase 2. Because a difference on one or both of these variables would equally reflect a different decision strategy, a MANOVA was run with these two dependent variables. A significant difference was found, $F(2, 5040) = 8.7$, $p < .01$. Individual follow-up tests indicated that the difference was caused by a difference in NACQ ($M_{NACQ} = 29.8$ vs. 34.5), and not TPER ($M_{TPER} = .33$ s vs. .32 s), $F(1, 5040) = 15.4$, $p <.0001$, and $F(1, 5040) = 4.2$, $p = .98$, respectively. Strangely, TPER goes in the opposite direction from the one predicted. E-A participants take slightly longer per acquisition, but on average make fewer acquisitions. This is still evidence of processing differences between the two conditions brought on by previous task constraints. In the opposite direction, the same MANOVA also yielded a significant difference between A-E and E-E participants, $F(2, 5040) = 21.5$, $p < .0001$. The fact that this test yields a higher level of significance is not surprising, given that the effort-accuracy underadjustment bias was stronger in this direction. Individual follow-up tests indicated that the difference was driven by NACQ ($M = 25.7$ vs. $M = 21.2$) rather than TPER ($M = .27$ s vs. $M = .27$ s), $F(1, 5040) = 41.4$, $p < .0001$, and $F(1, 5040) = .04$, $p = .85$.

While these results show strong evidence of the use of different decision strategies, the direction of causality can be questioned. Perhaps the different response times among the conditions caused the difference in the processing variables, and not vice versa. To test for this direction of causation, a MANCOVA was run with response times included as a covariate (the same analysis was also run with both linear and quadratic terms included, and the same pattern emerged). The difference in decision strategies remains significant both for the E-A to A-A comparison, $F(2, 5040) = 11.6$, $p < .0001$ and the A-E to E-E comparison, $F(2, 5040) = 13.9$, $p <$ 0001.

Discussion

Study 4 provided evidence in support of hypothesis 4—that biased selection causes participants to continue to use strategies similar to the previously used strategy. This evidence came from process-tracing data, which showed that decision strategies for participants who changed constraints remained similar to their previous strategies. This difference in decision strategy could have simply reflected the difference in response speed. After all, slower responses should result in higher NACQ and TPER. Instead, NACQ and TPER did not vary directly with response speed, but differed systematically depending on the starting condition. These systematic processing differences demonstrate that biased selection of decision strategies is another contributor to the effort-accuracy underadjustment bias.

GENERAL DISCUSSION

The metacognition that develops as a result of decision making in a given context will be limited in scope to the properties of the original context. When an environmental change in that decision task creates a need for the decision maker to adjust his/her previous levels of cognitive effort, that metacognition is often misapplied. This misapplication results in an anchoring effect of the previously determined effort levels, which I term the effort-accuracy underadjustment bias. The studies presented here demonstrate the bias in situations related to consumer decision situations, and thus focused on the bias's implications in that domain. However, the bias will also likely affect a number of other decision contexts where metacognition can be misapplied.

The goals activated by a decision context can be the result of metacognition from a source other than previous experience in that decision context. Wood and Lynch (2002) found that expertise in a consumer topic (not necessarily developed by choosing in that particular environment) causes deactivation of learning goals relative to a control group, which eventually led to poorer decision quality. The experienced consumers' metacognition in the domain was misapplied, just as in the studies above, though the metacognition had a different source. The precise mechanism by which consumers misapply their metacognition, and the extent to which this overgeneralization occurs, is an interesting topic for future research.

The term 'decision strategy' has been applied to myriad contexts, all of which could potentially be subject to biased selection. The studies presented here limited the examination of decision strategies to what Beach and Mitchell (1978) refer to as 'unaided analytic' strategies. The selection of intuitive versus analytic strategies is also potential source of biasing. Intuitive, System I decision strategies are considered the default decision strategy, and some property of the decision task needs to 'flag' the decision if more analytic, System II processing is going to be applied (Kahneman and Frederick 2002). The development (and non-development) of metacognition surrounding these flags is subject to a number of spurious influences, such as perceptual fluency (Oppenheimer 2006) and established social norms (Greene and Haidt 2002; Haidt 2001). Previous experience with a given decision topic will establish a metacognitive signal that directs processing to System I or System II. A change in the decision environment which should alter the processing System might not inspire action in the metacognitive system.

The decision model used in the current research assumed that the selection of a decision strategy mediated the pathway between decision goals and decisions. A number of frameworks forego the need for multiple decision strategies and instead apply a universal processing strategy (Shugan 1980; Busemeyer and Townsend 1993; Russo and Carlson 2006). Adopting such a framework highlights the importance of scale evaluation. All of these frameworks commonly suggest that the act of choosing among options is a process of sampling pieces of information to update one's preferential leanings. When one's leanings cross an internal threshold, the decision is made. Setting a higher threshold yields higher accuracy and higher effort, whereas a lower threshold guarantees the opposite.

Despite the centrality of this threshold to decisions, no previous work has explicitly examined how decision makers set these thresholds. Busemeyer and Diederich (2002) presume that more important decisions receive a higher threshold, but make no explicit examination of this assumption. The current research is the first to do so. It finds that once decision makers set their threshold in a given decision environment, they underadjust the threshold for changes to that environment. In such a framework, scale distortion alters the perceived pace at which the threshold is met.

This theoretical framework leads to important extensions of the current work to other research streams. For example, work in predecision distortion finds that decision makers distort information to favor a currently preferred alternative (Russo, Meloy and Medvec 1998). One of the hypothesized reasons for this distortion is to increase the pace of decision making. If this hypothesis is true, then the amount of predecision distortion will likely be moderated by the expected pace of the decision. If a decision maker expects to make a decision quickly, he/she will likely increase the amount of predecision distortion to enable a more rapid decision.

The effort-accuracy underadjustment bias affects the way decision makers approach, make, and evaluate their decisions. Decision makers approach a decision with the same goals as with previous decisions in the context, even when the environment recommends different goals. They make decisions the same way as before, even if a different strategy is optimal. They evaluate their decisions relative to previous decisions, even if those previous decisions are not as related as the decision maker believes. Because the bias affects decisions in general, its application to decisions varying in scope and type is a promising area of future research.

References

Allison, Paul D. (1984), *Event history analysis: Regression for longitudinal event data*, Beverly Hills, CA: Sage Publications.

Bargh, J. A., and Chartrand, T. L. (1999), "The Unbearable Automaticity of Being," *American Psychologist*, 54, 462-479.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A. Y., Barndollar, K., and Troetschel, R. (2001), "The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals," *Journal of Personality and Social Psychology*, 81, 1014-1027.

Beach, L. R., and Mitchell, T. R. (1978), "A Contingency Model for the Selection of Decision Strategies," *Academy of Management Review*, 3, 439-449.

Ben Zur, H., and Breznitz, S. J. (1981), "The Effect of Time Pressure on Risky Choice Behavior," *Acta Psychologica*, 47, 89-104.

Bettman, J. R., Johnson, E. J., and Payne, J. W. (1990), "A Componential Analysis of Cognitive Effort in Choice," *Organizational Behavior and Human Decision Processes*, 45, 111-139.

Bettman, James R., Luce, Mary Frances, and Payne, John W. (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (Dec), 187-217.

Busemeyer, Jerome R., and Diederich, Adele (2002), "Survey of Decision Field Theory," *Mathematical Social Sciences*, 43 (3), 345-370.

Busemeyer, Jerome R., and Myung, Injae M. (1992), "An Adaptive Approach to Human Decision Making: Learning Theory, Decision Theory, and Human Performance," *Journal of Experimental Psychology: General*, 121(2), 171-194.

Busemeyer, Jerome R., and Townsend, James T. (1993), "Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment," *Psychological Review*, 100(3), 432-459.

Cacioppo, John T., Petty, Richard E., Feinstein, Jeffrey A., and Jarvis, Blair G. (1996), "Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition," *Psychological Bulletin*, 119, 197-253.

Creyer, Elizabeth H., Bettman, James R., and Payne, John W. (1990), "The Impact of Accuracy and Effort Feedback and Goals on Adaptive Decision Behavior," *Journal of Behavioral Decision Making*, 3, 1-16.

Edland, Anne, and Svenson, Ola (1993), "Judgment and Decision Making Under Time Pressure: Studies and Findings," In *Time Pressure and Stress in Human Judgment and Decision Making*, ed. O. Svenson and A. J. Maule, New York: Plenum Press.

Epley, Nicholas, and Gilovich, Thomas (2001), "Putting Adjustment Back in the Anchoring and Adjustment Heuristic: Differential Processing of Self-Generated and Experimenter-Provided Anchors," *Psychological Science*, 12, 391-396.

Erdem, Tulin, and Keane, Michael P. (1996), "Decision Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15(1), 1-20.

Fitzsimons, Grainne M., and Bargh, John A. (2003), "Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated with Relationship Partners," *Journal of Personality and Social Psychology*, 84, 148-164.

Greene, Joshua and Haidt, Jonathan (2002), "How (and Where) Does Moral Judgment Work?" *Trends in Cognitive Sciences*, 6, 517-523.

Haidt, Jonathan (2001), "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review*, 108 (4), 814-834.

Helson, Harry (1964), *Adaptation-Level Theory*, New York: Harper and Row Publishers.

Johnson, Eric J., and Payne, John W. (1985), "Effort and Accuracy in Choice," *Management Science*, 31, 394-414.

Kahneman, Daniel, and Frederick, Shane (2002), "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," In *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. Thomas Gilovich, Dale Griffin, and Daniel Kahneman, New York: Cambridge University Press.

Kalbfleisch, John D., and Prentice, Ross L. (1980), "The Statistical Analysis of Failure Time Data," New York: Wiley.

Lerner, Jennifer S., and Tetlock, Philip E. (1999), "Accounting for the Effects of Accountability," *Psychological Bulletin*, 125 (2), 255-275.

Mano, Haim (1990), "Anticipated Deadline Penalties: Effects on Goal Levels and Task Performance," In *Insights in Decision Making*, ed. Robin Hogarth, Chicago: University of Chicago Press, 154-172.

Mellers, Barbara A., Schwartz, Alan, and Cooke, Alan (1998), "Judgment and Decision Making," *Annual Review of Psychology*, 49, 447-477.

Meyer, Robert, Erdem, Tulin, Feinberg, Fred, Gilboa, Itzhak, Hutchinson, Wesley, Krishna, Aradhna, Lippman, Steven, Mela, Carl, Pazgal, Amit, Prelec, Drazen, and Steckel, Joel (1997), "Dynamic Influences on Individual Choice Behavior," *Marketing Letters*, 8 (3), 349-360.

Novemsky, Nathan, and Dhar, Ravi (2005), "Goal Fulfillment and Goal Targets in Sequential Choice," *Journal of Consumer Research*, 32(3), 396-404.

Oppenheimer, Daniel (2006), "The Secret Life of Fluency," paper presented at the Wharton Decision Processes Seminar, Philadelphia, PA.

Parducci, Allen (1963), "The Range-Frequency Compromise in Judgment," *Psychological Monographs*, 77 (2, Whole No. 565).

________ (1995), *Happiness, Pleasure, and Judgment: The Contextual Theory and its Applications*, New Jersey: Lawrence Erlbaum Associates.

Payne, John W., Bettman, James R., and Johnson, Eric J. (1988), "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 534-552.

________ (1993), *The Adaptive Decision Maker*, New York: Cambridge University Press.

Restle, Frank (1962), "The Selection of Strategies in Cue Learning," *Psychological Review*, 69 (4), 329-343.

Rieskamp, Jorg, and Otto, Philipp E. (2006), "Strategy Selection Learning Theory," *Journal of Experimental Psychology: General*, 135 (2).

Russo, J. Edward, and Carlson, Kurt A. (2006), "Individual Decision Making," In *Handbook of Marketing*, ed. Barton Weitz and Robin Wensley, London: Sage Publications, 371-408.

Russo, J. Edward, and Dosher, Barbara (1983), "Strategies for Multiattribute Binary Choice," *Journal of Experimental Psychology: Learning, Memory, and Cogntion*, 9 (4), 976-696.

Russo, J. Edward, Meloy, Margaret G., and Medvec, Victoria Husted (1998), "Predecisional Distortion of Product Information," *Journal of Marketing Research*, 35 (Nov), 438-452.

Sherif, Muzafer, Taub, Daniel., and Hovland, Carl I. (1958), "Assimilation and Contrast Effects of Anchoring Stimuli on Judgments," *Journal of Experimental Psychology*, 55, 150-155.

Shugan, Steven M. (1980), "The Cost of Thinking," *Journal of Consumer Research,* 7(2), 99-111.

Simon, Herbert A. (1957), *Models of Man*, New York: John Wiley and Sons.

Stevens, Stanley S. (1957), "On the Psychophysical Law," *The Psychological Review*, 64(3), 153-181.

Tetlock, Philip E. (1979), "Accountability: The Neglected Social Context of Judgment and Choice," *Research in Organizational Behavior*, 7, 297-332.

Tetlock, Philip E., Stitka, Linda, and Boettger, Richard (1989), "Social and Cognitive Strategies for Coping with Accountability: Conformity, Complexity, and Bolstering," *Journal of Personality and Social Psychology*, 57, 632-640.

_________ (1974), "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185, 1124-1131.

Viswanathan, Madhubalan (1997), "Individual Differences in Need for Precision," *Personality and Social Psychology Bulletin*, 23(7), 717-735.

Webster, Donna M., and Kruglanski, Arie W. (1994), "Individual Differences in Need for Cognitive Closure," *Journal of Personality and Social Psychology*, 67(6), 1049-1062.

Wood, Stacy L., and Lynch, John G. (2002), "Prior Knowledge and Complacency in New Product Learning," *Journal of Consumer Research*, 29 (Dec), 416-426.

Zajonc, Robert B. (1965), "Social facilitation," *Science*, 149, 149-169.

FIGURE 1

A STYLIZED DECISION MODEL DEPICTING ALL THE DECISION STAGES AND MECHANISMS PERTINENT TO THE EFFORT-ACCURACY UNDERADJUSTMENT BIAS

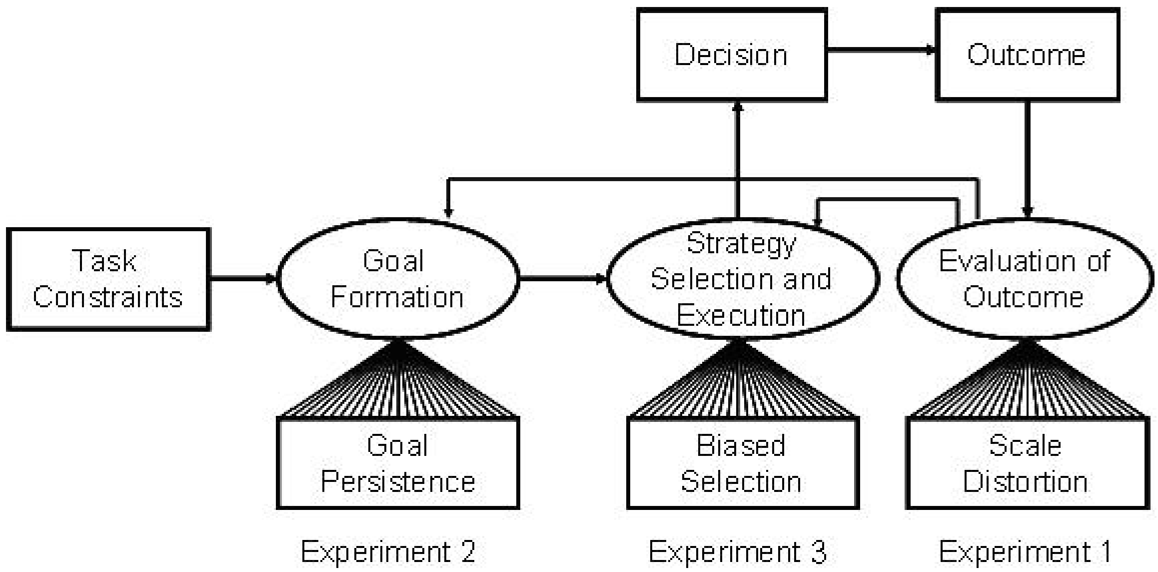


Figure 2

STUDY 1 STIMULUS

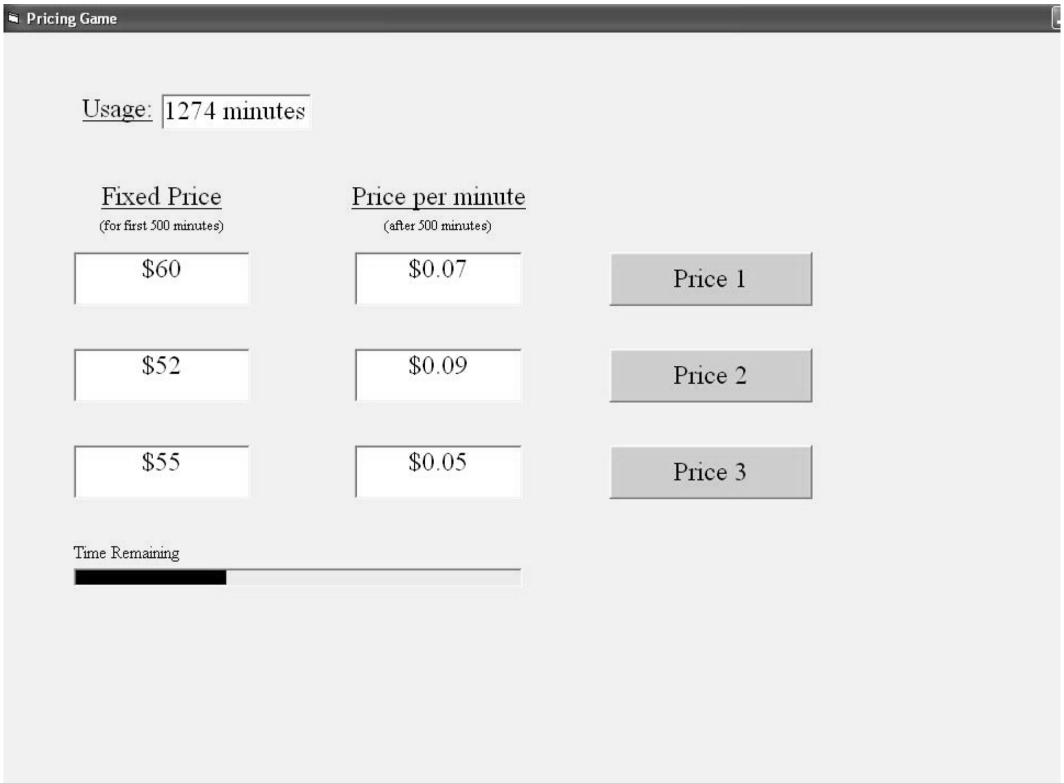
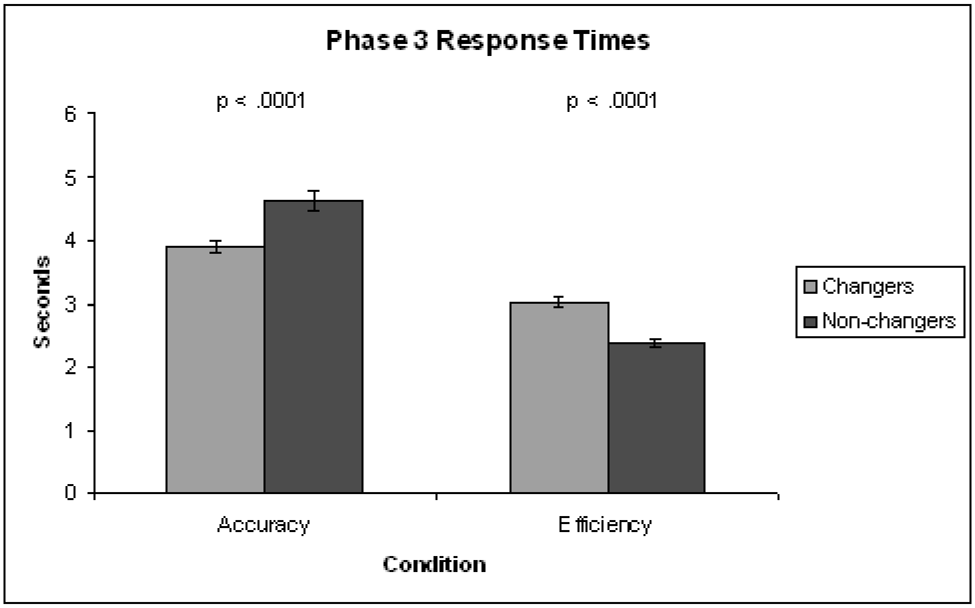


FIGURE 3

STUDY1: MEAN RESPONSE TIMES AND ACCURACY RATES BY CONDITION IN PHASE 3 (AFTER CHANGE)



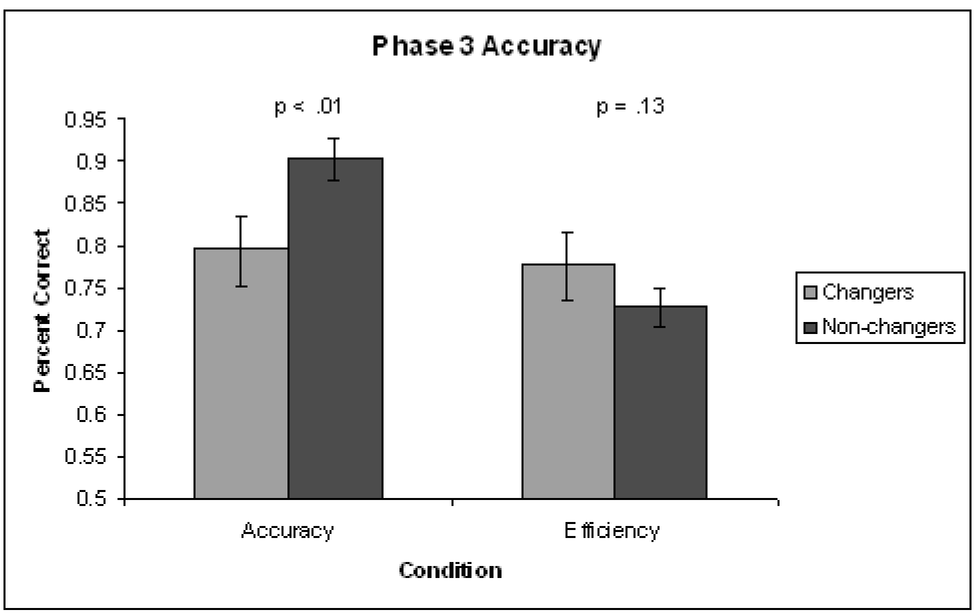
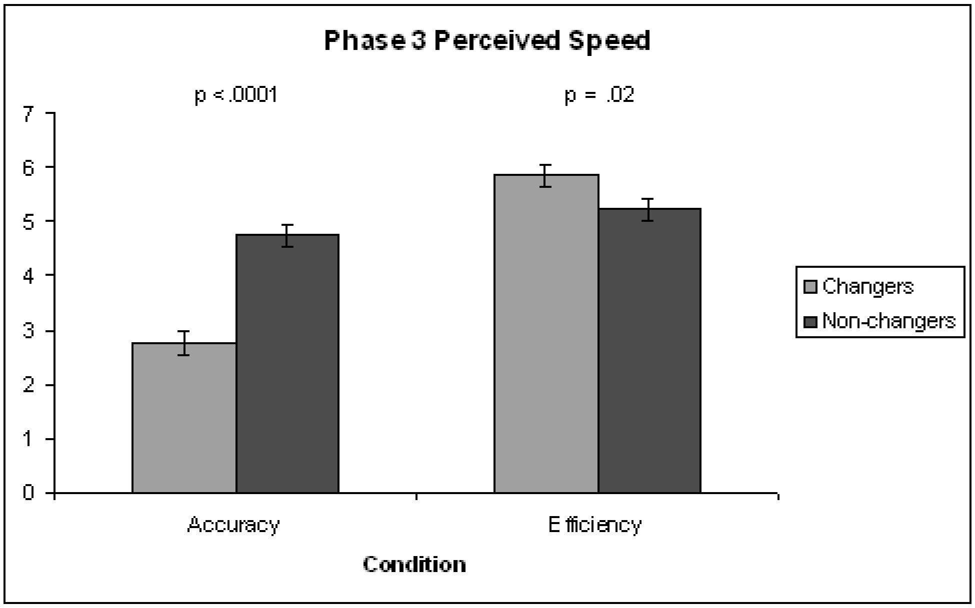


FIGURE 4

STUDY 1: MEAN PERCEIVED SPEED AND ACCURACY BY CONDITION AFTER PHASE 3



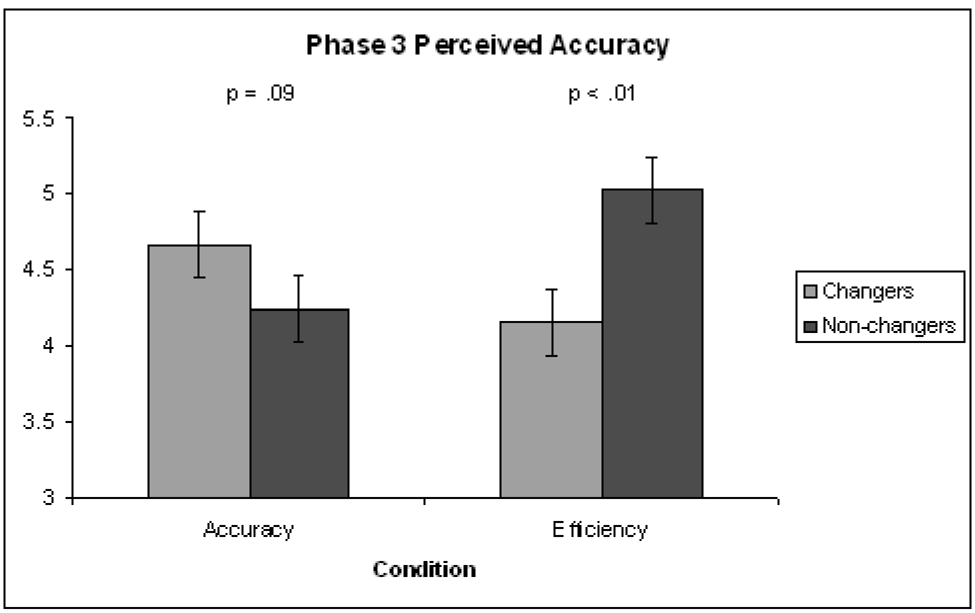
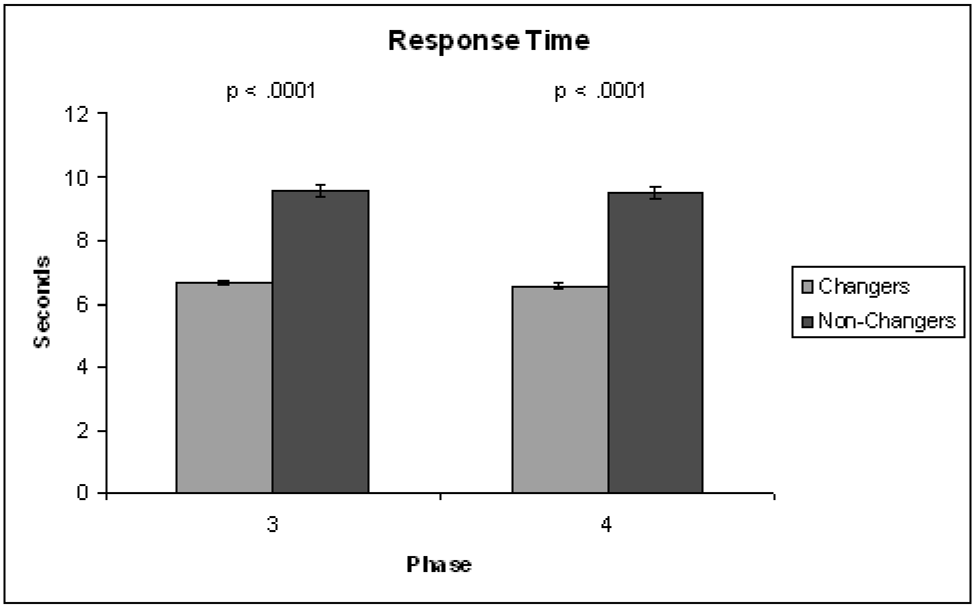


FIGURE 5

STUDY 2: MEAN RESPONSE TIMES AND ACCURACY RATES BY CONDITION IN PHASES 3 AND 4 (AFTER CHANGE)



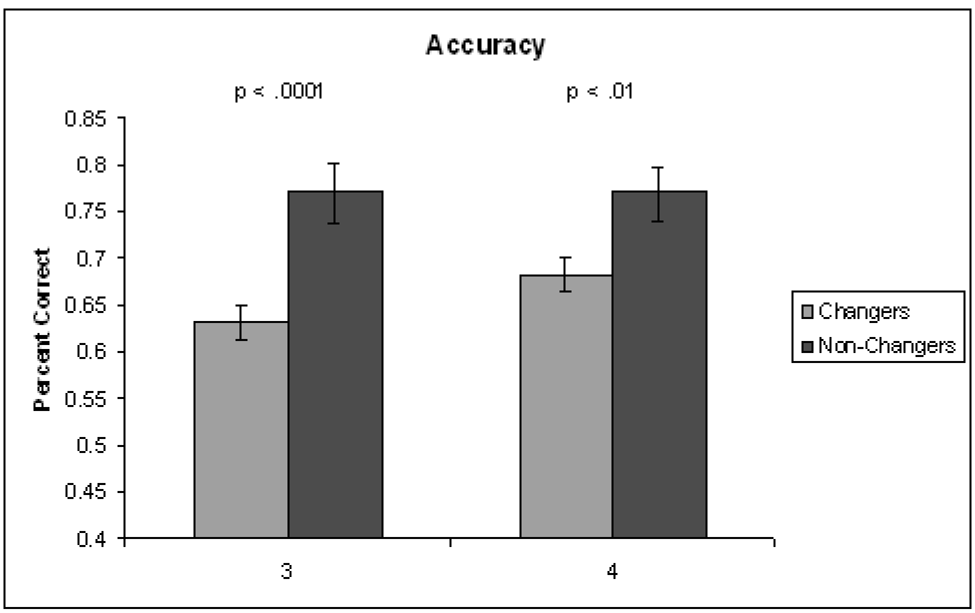


FIGURE 6

STUDY 3: MEAN RESPONSE TIMES AND ACCURACY RATES BY CONDITION IN PHASE 2 (AFTER CHANGE)