2014

Cognitive bias and learning from experience: Reflective processes for reducing bias

Dina Acklin
Louisiana State University and Agricultural and Mechanical College, dina.acklin@gmail.com

COGNITIVE BIAS AND LEARNING FROM EXPERIENCE: REFLECTIVE PROCESSES FOR REDUCING BIAS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

by
Dina M. Acklin
B.A., University of Pittsburgh, 2008
May 2015

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

While heuristic processing is often useful for quickly ascertaining information in everyday situations, it can lead to inaccuracies when task demands become complex and more systematic processing is required. These inaccuracies are often the result of confirmation bias, in which information that is consistent with our beliefs is noted at the expense of disconfirming evidence. The current decision making literature suggests that highlighting disconfirming evidence – termed negative feedback - might work to engage deliberate, systematic cognitive processes that lead to more accurate information acquisition. Using a probabilistic learning task where feedback is not consistently accurate (Matchmaker), the first experiment in this study will attempt to overcome confirmation bias by encouraging initial hypotheses to be considered from confirming and disconfirming vantages. It is proposed that errors resulting from bias will thus be made more salient and the testing of alternative solutions will be encouraged, resulting in greater accuracy. A second experiment will explore the cognitive processes involved in bias strengthening and determine if warnings of feedback error alter the way in which information is interpreted.

INTRODUCTION

As humans, we are tasked everyday with having to quickly and accurately sort through a vast amount of information. In many circumstances, this speedy and unsystematic type of heuristic processing works to our advantage and the conclusions we arrive at are sufficient to allow us to proceed on to the next task at hand. However, there are times when quick and intuitive judgments can lead to inaccurate conclusions due to a failure to systematically consider all available information and weigh alternatives. One example of this is referred to as confirmation bias, where learners only seek out information which is consistent with their initial conclusions (Silverman, 1992; Lilienfield, Ammirati, & Landfield, 2009; Nickerson, 1998). It is important to determine what cognitive processes allow such a bias to persist and to find strategies to enhance systematic consideration of information to reduce bias.

Previous research has shown that prompting learners to engage in systematic analysis of information can reduce reliance on heuristically based judgments. In one example, Natter and Berry (2005) have demonstrated how prompting can be used to improve the understanding of risk information for a fictitious medication. Participants were asked to imagine themselves as patients being prescribed a medication associated with mild side effects in some of the people to whom it is prescribed. Those in the control group passively read an informational pamphlet in which they were told of a 2% risk of side effects and were instructed to observe a bar graph depicting the 2% risk. The experimental group read the same pamphlet and information regarding the 2% risk of side effects and was additionally instructed to indicate the amount of risk by shading in the bar graph themselves. It has been found that using or producing graphs is an effective strategy for encouraging deliberate processing and correcting heuristically based assumptions due to an active engagement with the information to be learned (Stern, Aprea, &

Ebner, 2003; Cosmides & Tooby, 1996). A questionnaire presented after the pamphlet had been read asked participants to indicate the likelihood that they might experience the risk as well as the number of people in 1000 who would be expected to experience the risk. The results demonstrated that active engagement as a result of producing a graph led to more accurate risk assessments. A second experiment illustrated that engaging in reflective questioning about the presented information can have similar effects. By answering a reflective question regarding risk ("How many out of 100 people who take this medication will experience one or more side effects?") before proceeding to the questionnaire, those in the active group demonstrated a better understanding of risk than those who only read the pamphlet. These findings support the notion that questioning increases the interaction between readers and the information, leading to a deeper understanding of what is to be learned (Graesser, Baggett & Williams, 1996). This study shows that when the number of hypotheses to be tested is clear and small, prompting to be systematic can be quite effective.

However, relying on exhaustive systematic processing may not be optimal in all situations due to the focus and sustained mental effort required. For example, when in an emergency situation it may be necessary to act quickly without time to consider all of the alternatives (Klein, 1999). Further, in everyday, real world situations, feedback is often less than 100% accurate and there may be variability in results across situations. For example, doctors cannot presuppose that a medication that works for one patient will work equally well for another. It has also been shown that tasks involving heavy perceptual or intuitive information, such as wine tasting, are performed more poorly when participants attempt to systematically describe their knowledge (Melcher & Schooler, 1996). Therefore when encouraging deliberate, systematic evaluation of information it is also important to understand that this kind processing may not be optimal for a particular task (Gladwell, 2005; Klein, 1999).

Given the presence of bias in a multitude of every day contexts, from forming assessments of interpersonal behavior to interpreting news reports, it is important to understand practical ways in which unbiased processing can be encouraged. When attempting to engage in decision making *under the influence of an initial bias,* people are more likely to rely on their intuitive judgments at the expense of considering alternate possibilities (Kahneman, 2011). Reliance on intuitive judgments in such situations is likely to increase the bias by selective attention given to bias-consistent information. Current research indicates that information that contradicts initial hypotheses (negative feedback) during learning can result in more deliberative reasoning (Downer, Bhatt, & Montague, 2011). However, it remains unclear if such a strategy will remain effective in the face of an existing bias because deliberate processing might be subject to biased interpretation of feedback as well, particularly in cases were feedback is not consistently reliable.

This study involves a probabilistic learning task (in which feedback is not 100% accurate) and a situation where learners begin the task with a bias that is implanted in the laboratory. These conditions, while very challenging to learners, reflect common conditions in real world learning (for example, doctors prescribing medication or consumers purchasing a new gadget). The experimental manipulations focus on two issues that may alter the use of systematic processing in a probabilistic learning task: attention to negative outcomes and knowledge of error during feedback. Additionally, this paradigm permits an objective measure of bias strength initially and as it changes throughout the learning period to assess the effect of learning strategies on bias (Ledet, 2013).

The first experiment will require that participants consider their biased hypothesis from both a confirming and disconfirming vantage. It is predicted that errors resulting from bias will be more salient when learners are directed to make choices inconsistent with their bias, causing participants to seek out and test alternative solutions. The second experiment aims to gain an understanding of the underlying processes through which biases are reinforced or weakened by examining the effect warnings of feedback error might have on decision making.

Confirmation Bias: Confirming What May Not Be True

Confirmation bias arises when people continue to seek out evidence consistent with their beliefs even in light of information that speaks to the contrary (Silverman, 1992). In such situations, holding steadfastly to the first decision that comes easily to mind can prevent other more accurate decisions from being discovered. While our intuitive decision making processes are largely useful for a multitude of tasks, there are times when the belief in our first, automatic conclusions can be detrimental.

The most famous demonstration of the confirmation bias can be found in a task developed by Wason (1960). In his triplet task, participants were given a triplet of numbers (2 – 4 – 6, for example) and asked to test hypotheses about the rule until the believed they discovered the correct rule (which was any increasing sequence of numbers). After each triplet was generated, they were given feedback about whether or not it fit the experimenter's rule. Participants were more likely to generate a specific rule, such as 'increasing multiples of two' rather than a more general but more accurate rule, such as 'increasing numbers'. Even though the more general rules would generate more correct results, participants tended to stick with the first rule that allowed them to produce positive results. Thus, they rarely discovered the correct rule.

Biasing occurs when we are influenced to think or believe a particular way about problems and situations that prevents discovering a more accurate solution. Even when we are trained scientifically to seek out instances that disconfirm our hypotheses, in day to day situations we engage in confirmation bias, seeking out that which best conforms to our currently held beliefs (Downar et al., 2011). Because our mind first brings forth occurrences in which our solutions were effective and neglects those times when the same solutions have proven ineffective, we have a natural tendency to discount negative feedback (Gilbert, 1991). In neglecting to test other hypotheses, the participants in Wason's experiment failed to see that their first conclusion might not be the only outcome. Such a myopic focus on one outcome can have detrimental implications when expanded to complicated real world scenarios, such as media reporting and implementation of national policies.

Bodenhausen and Wyer (1985) highlighted the problems that can arise from biased reasoning in complex scenarios. Subjects in their study were asked to recommend disciplinary action for fictional employees based on workplace infractions and indicate the likelihood that the infraction would occur again in the future. These infractions were either consistent or inconsistent with American or Arab ethnic stereotypes, using names and infractions that were established in a pilot study as being stereotypical of each group. They found that when biased thinking occurred, subjects recommended more severe disciplinary actions for stereotypical offenses and saw them as more likely to recur. Further, when asked to recall the scenarios of the stereotypical cases, subjects recalled less information than for non-stereotypical scenarios. In a subsequent experiment, participants were given additional life circumstances information that might be used to illuminate a situational reason for the infraction and which could mitigate reliance on stereotypes. Information about life circumstances served to decrease severity of disciplinary action and chance of recurrence only when stereotypes were not activated. Providing the stereotypical name information lead to a discounting of life circumstances, which was inconsistent with or irrelevant to the activated stereotype, resulting in recall for only stereotype-consistent information. This study demonstrates that the activation of biases prevented the integration of other decision relevant information, causing subjects to make decisions based on partial information. Knowing this, it is important to consider what might be driving an innate reluctance to consider multiple hypotheses.

In later work, Bodenhausen (1988) provided evidence that this neglect of alternative hypotheses resulted from selective processing. When participants engaged in biased reasoning, they paid special attention to the information that conformed to their beliefs, recalling a significantly higher number of bias-consistent items than bias-inconsistent items. Acting as jurors, subjects were to read a case file that contained neutral, exonerating, and incriminating pieces of information for a defendant that was given either an ethically-nondescript or Hispanic name. The name of the defendant was presented either at the beginning or at the end of the case file. After reading the file they were asked to make guilt judgments and, after a delay period, recall all of the information they could remember from the case file. They were then presented with all of the evidence items and asked to assign a rating of the probative implications for each. Those who read the stereotypical Hispanic name prior to reading the case file were more likely to judge the defendant as guilty and more likely to rate all evidence items more negatively than those who read the defendant's name last. Once an expectation was in place, participants seemed determined to justify that expectation even if that meant discounting evidence that might disconfirm their initial reasoning.

It is also important to note that because heuristic processes are often sufficient for day-to-day decision making, individuals tend to feel overly confident about all of their everyday hypotheses regardless of actual accuracy (Gilovich, 1991). Due to the high confidence associated with quick decision making, systematic and effortful processing of information fails to occur (Thompson, Turner, & Penncook, 2011; Topolinski & Reber, 2010; Thompson & Morsanyi, 2012). Indeed, it has been shown that when an answer comes to mind with little strain, we actually experience pleasurable sensations that evoke a slight smile response (Topolinksi, Likowski, Weyers, & Strack, 2009).

If we are prone to automatically rely on our first assessments and we have little reason to doubt their veracity, how then might we be encouraged to engage in more deliberate and systematic processes to arrive at more accurate results?

Strategies for Overcoming Confirmation Bias

Consider the Opposite. Research has found that when participants are instructed to consider possible alternatives to their original hypothesis along with the confirming and disconfirming evidence for each case, flaws in initial, intuitive responses are revealed. This realization allows for an opportunity to discover the most accurate solution to a given problem. By encouraging participants to consider several hypotheses, they can be induced to think that their original solution is flawed and test alternative hypotheses to arrive at the appropriate conclusion (Arkes, Faust, Guilmette, & Hart, 1988; Kray & Galinsky; 2003).

In one experiment utilizing a covert method to induce considering an alternative hypothesis (Lord, Lepper, & Preston 1984), found evidence that this method helped participants ask less biased questions. Participants were instructed to gather information about a person in the next room and were given a list of questions to choose from in determining if the other person was an extrovert. The list included questions aimed at prompting about extroverted, introverted, and neutral behaviors and was the same for all groups. Participants in one condition were given a personality profile of an extrovert to aid in their selection, while those in a be-unbiased condition were given additional instructions to be as accurate and fair are possible is choosing questions to determine the person's character. Importantly, participants in a consider-the-opposite condition were given the personality profile of an introvert under the guise that the extroverted sheet had been misplaced. They were told that since introverts were the opposite of extroverts, the sheet should prove just as useful. In comparison to the other two conditions, participants instructed to consider-the-opposite were found to ask significantly fewer questions aimed at determining extroverted behaviors and significantly more questions aimed at determining introverted behaviors. This demonstrates that due to the covert suggestion of an alternate personality type, participants were more likely to diversify their question selection. The researchers were able to get participants to overcome biased hypothesis testing and successfully encourage the testing of alternative hypotheses.

While studies like the one described above have expounded the benefits of seeking alternative hypotheses (Koriat, Lichtenstein, & Fischoff, 1980; Mussweiler, Strack, & Pfeiffer, 2000; Hoch, 1985), still others have shown that confusion over relevant details (Cox & Popken, 2008) and an inability to spontaneously create alternative hypotheses (Sanna, Schwartz, & Stocker, 2002; Schwarz, et al., 1991) can in fact lead to a strengthening of original hypotheses. Findings like these have lead researchers to examine how confidence changes when participants are instructed to consider the opposite hypothesis, or take into account information that contradicts an initial belief. It has been demonstrated that that initial, heuristic judgments carry with them a 'feeling of rightness' (Thompson & Morsanyi, 2012). Their easy and effortless accessibility leads to a pleasurable feeling we come to associate with being correct. Therefore, when we cannot easily come up with an alternative to our initial hypotheses we become even more confident in its efficacy. When the consider-the-opposite strategy is effective, the acknowledgment of alternative solutions lessens our initial feelings of rightness and confidence, thereby encouraging us to test other solutions that may be more accurate (Koriat, Lichtenstein, & Fischoff, 1980; Arkes et al., 1988). When determining which circumstance can best enhance decision making, these studies illustrate that we must also examine the roll that confidence plays in these various scenarios.

Attention to Negative Feedback. To examine the ways in which people learn to make accurate decisions and what may lead to the formation of spurious beliefs, Downar et al. (2011) devised a probabilistic learning task through which they could study the performance of experienced physicians. In the task, physicians were asked to prescribe one of two fictional medications to simulated patients. Several relevant items of information were shown alongside each participant to aid in choosing the correct medicine. While only one of these pieces of information determined the best treatment, the physicians invariably created complex hypotheses in trying to arrive at the correct answer. In fact, nearly half did not prescribe the correct medications at much better than chance levels. Further analysis of the physicians' progress through the task indicated that provided feedback was not integrated into their initial hypotheses, leading to diagnoses based on spurious factors. In sum, after an initial rule was determined, the physicians failed to engage in systematic processing to test their hypotheses and instead relied on their first instinct as correct.

While it is true that overall the physicians performed at near chance levels, the researchers conducted a separate analysis of a high performing subset of the group who chose the optimal drug between 77% and 98% of the time. fMRI readings obtained during the task revealed that these high performers showed greater activation in the prefrontal cortex after experiencing a failed match whereas their lower performing peers showed this activation only after successful matches. This increased activation suggests something unique about this group in that, by being more attentive to failure, they were better able to integrate information in a way that led to more accurate decision making. Conversely, low-performers who focused only on information garnered from successes were more likely to fall victim to the confirmation bias and were significantly less accurate.

It is evident from the results of this experiment that the effects of confirmation bias and the formation of spurious beliefs can be reduced if learners are attentive to situations in which their original beliefs are shown to be inaccurate. It also appears as though, under most circumstances, paying special attention to this type of information is something we fail to do intuitively.

More recent research has explored the effects that highlighting negative feedback has on the decision making process. The Matchmaker task created by Ledet (2013) serves as a more accessible alternative to the medical decision task for a general population (not trained physicians). Importantly, the task was designed to produce a specific spurious hypothesis, thereby establishing experimental control over the bias created in the lab. In this paradigm, participants act as matchmakers for two simulated male bachelors with the goal of determining what compatibility factors best predict a match for each bachelor. After being introduced to the two bachelors with a photo and a brief description, participants are shown some sample profiles that either client has rated a 'good match' for themselves. In the biased conditions, participants were biased to believe that bachelors should be matched according to Entertainment Preference, with one bachelor preferring matches who liked sports and the other preferring matches who liked video games. Next participants are shown many new profiles and are tasked with matching the profile to the bachelor (see Figure 1).

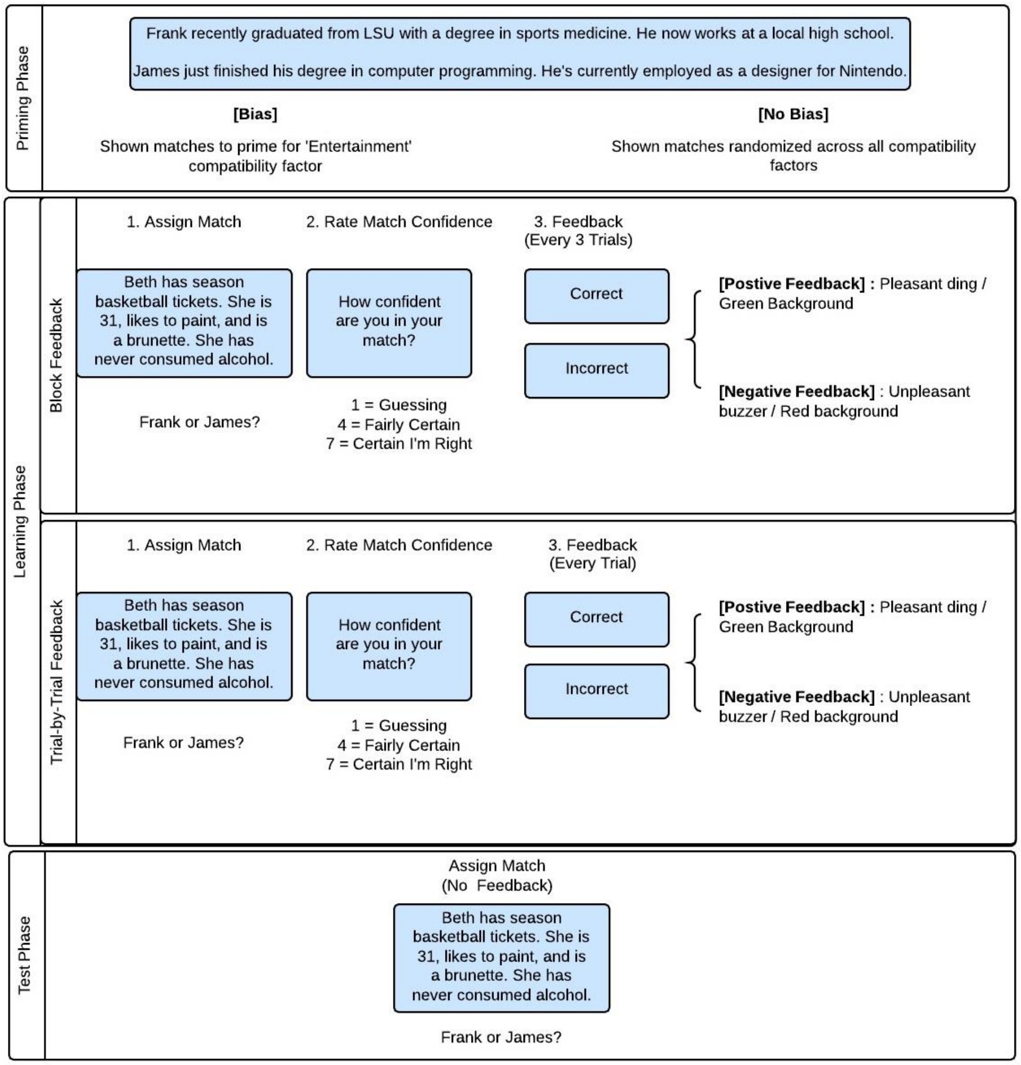


Figure 1. Schematic representation of Matchmaker task used by Ledet (2013). During the Priming Phase for Biased groups, Frank was introduced followed by four sample matches followed by James and four sample matches. Both block feedback and trial-by-trial feedback sections during the Learning Phase were comprised of 54 trials. The test phase consisted of 12 trials.

Unlike in the task carried out by Downer et al. (2011), Matchmaker primed participants with a specific bias (an Entertainment preference of sports for one bachelor and video games for the other). During a priming phase, participants were introduced to each bachelor by reading a paragraph that generally described the bachelor. Next, four sample matches were shown for each bachelor. The first sentence of each match described a trait consistent with the spurious, biased compatibility factor. This spurious factor had no bearing on whether or not the match would be a good one in future selections. Only the compatibility factor of Hair Color was critical to matching each client with a subsequent good match. During the priming phase, traits for the critical factor were always mentioned after the spurious factor. In keeping with the design of the Downar et al. (2011) probabilistic medical decision task, the critical factor resulted in subsequent good matches 75% of the time and a bad match 25% of the time in the subsequent learning phase. In the priming phase, the sample matches for a bachelor also corresponded to the matches 75% of the time (3 out of the 4 sample matches provided with respect to Hair Color).

This priming manipulation provides an added level of control, ensuring that all participants are biased toward testing the same initial hypothesis at the start of the learning phase. Additionally, establishing the same bias for all participants allows for the analysis of bias congruent trials (where both the correct biased and critical factors for one bachelor appear together) and bias incongruent trials (where the critical factor for one bachelor appears with the biased factor for the other). This allows for a better understanding of whether participants are relying on their bias to make decisions or if they are actively testing alternative hypotheses. When relying on bias, participants would be predicted to achieve high accuracy on bias congruent trials, as the primed and critical factors occur together. Correct rule learning (assigning matches based only on Hair Color) would appear most evident on bias incongruent trials. Because the biased and critical factors do not co-occur, reliance on bias will result in an incorrect response. However, if the participant learns that the critical factor most consistently produces correct answers, they should use this information to guide their match making. As such, accuracy on bias-incongruent trials should increase. Bias irrelevant trials were also presented one-third of the time (with Entertainment Preference listed as movies). With no bias-relevant information present, accuracy on bias irrelevant trials depends on learning how to correctly assign the critical factor of hair color.

Matchmaker also asked participants to provide a confidence rating after matching a profile with a bachelor. The inclusion of a confidence rating provides insight into the type of cognitive processes at work. When engaging in heuristic processing, answers come to mind with ease and fluency. This ease creates a ‘feeling of rightness’ that increases confidence in the accuracy of the immediately arrived at answer. Conversely, when engaging in systematic processing, the effort required to arrive at a solution reduces confidence in the answer’s correctness (Thompson, Turner, & Penncook, 2011; Topolinski & Reber, 2010; Thompson & Morsanyi, 2012). This over-confidence can be problematic in complex situations when finding the correct solution demands systematic thinking. For example, Zacharakis and Shepard (2001) conducted a study of venture capitalists in order to understand how confidence affected their ability to predict new venture success. Experienced venture capitalists were presented with cases containing either five or eight relevant items of information in order to make their prediction. Those who saw more information were found be more confident in their predictions for the venture, but also were found to be less accurate.

Using Matchmaker, Ledet manipulated the salience of either successful or unsuccessful trials by using red feedback screens and buzzer sounds to highlight *failures* in one condition and green feedback screens and dinging sounds to highlight *successes* in another. It was hypothesized that these external cues would highlight failure in such a way that would encourage systematic processing and increase accuracy, mimicking the successful strategy seemingly utilized by Downar et al.'s (2011) most successful subset of physicians. This manipulation succeeded in reducing participant confidence and increasing reaction time, factors consistent with systematic processing in previous research (Koriat, Lichtenstein, & Fischoff, 1980; Kahneman, 2011). However across all trial types, no significant differences in accuracy were found between the negative-feedback and positive-feedback groups. Post-hoc analyses examined the effect of feedback manipulation on each trial type. Learning would appear evident if responses to bias incongruent trials became more accurate over time, indicating that the influence of the biased compatibility factor has diminished and the critical factor alone was being used to guide decision making. While accuracy on bias incongruent trials did increase as the learning phase progressed, there were no differences between feedback groups and both achieved accuracy at well below chance, around 24%. Further, while both negative-feedback and positive feedback groups became less accurate on bias congruent trials toward the end of the learning phase, mean accuracy remained around 85%. This suggests that learning was the result of task exposure rather than feedback manipulation and that, while participants eventually recognized the biased factor to be suboptimal for making matches, they were unable to determine the actual critical factor. Ledet concluded that the ability to focus on failure is a primarily an internal trait and one that is particularly difficult to alter. He speculated that participants stuck with their initial biased reasoning because the emphasis on negative feedback reduced confidence in systematic decision making, rather than reducing confidence in heuristic decision making as intended.

Warnings of Feedback Error. In day-to-day situations, manipulations to avoid reliance on our preconceived notions are hardly plentiful and often nonexistent. More often than not, we are expected to determine for ourselves when the information we encounter warrants skepticism: the ubiquity of radio, television, Internet, and social media accounts of current events demand that we consider the reliability of a variety of sources on a daily basis. We know that in most situations a certain amount of error is present in the information we are receiving. How then do our initial biases affect how we interpret such knowledge?

To begin to address this question, we might first consider how our ability to detect error changes when we are explicitly made aware of its existence. Does knowledge of inconsistencies have any impact on our ability to detect the instances when this is occurring? A recent study by Biros, George, and Zmud (2013) aimed to determine what interventions might lead to better error detection. In an experimental task, participants enrolled in a standard human resources training course at a military base were required to make management decisions using data garnered from the base's human management information system. They were divided into three intervention groups and one control group. The first intervention consisted of traditional training to detect deceptive error within a widely used military human resource information system. The second consisted of a warning prior to beginning the task that an administrator had tampered with the system prior to being discharged and that this may affect their work on the exercises they were to accomplish. Although warned of the error explicitly, participants were not told the exact amount of error within the system, which was present in 14.5% of the data items they were to analyze. The third intervention received both the training and warning just prior to beginning the task. It was found that warning alone and warning combined with timely training were both significantly effective in increasing participants' ability to detect error within the system, thereby increasing accuracy. Further, participants were found to take more time on trials where error was successfully detected, indicating systematic and analytic processes at work.

While warnings about error lead to task performance with a higher rate of accuracy, the researchers did not consider how these results might change if participants were biased as to what features of the database to focus on or what elements of the task had been altered. For example, had the warning stated that tampering occurred because the administrator's score on the Enlisted Performance Report factor was not high enough to warrant an anticipated promotion, the participants might have assumed error was confined to this field in the database. Such an assumption could change task performance based on a bias to consider only one type of data item. To examine the effect that warnings might have on biased assumptions, researchers have taken advantage of the biases present within interpersonal relationships

A study carried out by Stiff, Kim, and Ramesh (1992) considered the effect that a warning of deception would have on an established truth bias. The truth bias is a presumption of truthfulness that can occur between two people in well-developed relationships. Repeated exposure to a partner's frequent truthful behavior leads to the assumption of truthfulness across situations. Much like the confirmation bias, the truth bias is used heuristically and creates a baseline of expectation. In their study, Stiff et al. recruited friendship dyads and split them into interviewers and interviewees. The interviewees were to watch two video clips eliciting either pleasant or unpleasant reactions and then told to either lie or be truthful about their reactions when asked by their interviewer friend. Interviewers were not shown the video clips but rather told that they were to question their friend about their reaction to what they were shown. Interviewers in a suspicion-arousing condition were also told that some interviewees are told to be truthful while others are told to respond deceptively during the interview. The researchers found that when suspicion-arousing information was provided, reliance on the truth bias was significantly reduced and those interviewers were less likely to judge their friend's behavior as truthful.

From Stiff et al. (1992), we begin to understand how the mere suggestion of deception can affect an otherwise well-established bias. Interrupting the use of biased, heuristic processes encouraged participants to become more cognitively involved in their decision making. Indeed, when Stiff and colleagues compared cognitive involvement across groups by coding the information-gathering questions posed in the interviews, they found that interviewers with strong truth biases asked significantly fewer cognitively demanding questions than those for whom the bias had been disrupted.

While arousing suspicions effectively reduced reliance on bias and led to more critical thinking, it is important to note that this did not make those participants any more accurate than those in the non-suspicious group. In other words, whether interviewers were more or less biased to believe that interviewees were being truthful had no influence on their ability to detect actual truthfulness. Similar results were also found by Toris and DePaulo (1985) under a comparable paradigm: interviewers warned that a paired interviewee may try to deceive them became more suspicious and less confident but were no more accurate in their assessments of truthfulness than controls.

These studies serve to show the strength of biases and the difficulties involved in overcoming them. Even though warnings effectively reduce reliance on biased reasoning, decrease confidence, and encourage critical thinking, biases persist to the extent that they prevent accurate decision making. Thus it becomes important to understand the cognitive processes underlying decision making under bias and how warnings of error affect those processes.

Bias and the Interpretation of Exceptions

Once it is understood how biases affect the interpretation and representation of information, strategies for increasing accuracy in spite of them can be developed. As previously mentioned, Bodenhausen (1988) found support for the selective processing hypothesis – that initial presentation of stereotype-activating information can lead to better recall of stereotype-consistent evidence items than for other item types. In a follow up to his initial experiment, Bodenhausen ensured that the stereotype itself was not simply being used as a retrieval cue. Participants were to act as jurors after being presented with a defendant who had either a stereotypical Hispanic or nondescript name. Exonerating, incriminating, and neutral items were then presented one at a time and given a probative rating (-5 = extremely unfavorable, +5 = extremely favorable). Rating evidence items as the information was first presented aimed to encourage equal depth of processing across item type, regardless of stereotypical consistency. The participants were then to make a guilt judgment and, after a short delay, given a free recall task in which they were to recall as much information about the case as possible. When asked to provide ratings of evidence on an item-by-item basis during encoding, participants recalled all evidence items at a better rate, regardless of if a stereotype was activated or not. When the stereotype was activated, both exonerating and incriminating items were recalled the same proportion of the time. These results illustrate that while biases serve to highlight only information that is consistent with a particular belief, recall of stereotype-inconsistent information can be fostered when all information is carefully considered during encoding.

However, more recently, Williams, Lombrozo, and Rehder (2013) have examined the effects of item-by-item processing when patterns are reliable or misleading (e.g. – patterns containing exceptions). The results of their study indicate a different set of processes and outcomes. In a task similar that of Ledet's (2013) Matchmaker, participants were shown 10 individuals along with some demographic and behavioral information and asked to determine if the person rarely or frequently donated to charity. In one condition all descriptor information perfectly predicated to donation likelihood. In another condition, some items were unreliable with exceptions occurring twice in two of the descriptors. For example, while young age and extroverted behavioral descriptors were typically associated with rare donators, two times it occurred than an older individual or an introvert would also be seen to donate rarely. Participants were either told to explain an individual's behavior for each trial or that they would be asked to explain the behavior later. After answering for donation frequency, they were given the correct answer and briefly shown the individual again before continuing on to the next trial.

Those in the misleading conditions produced more errors than when the pattern was reliable, and among explainers, those who were asked to provide trial-by-trial explanations performed significantly worse than those who did not, particularly when items contained exceptions. Rather than forming a nuanced representation of each individual, as would be predicted from the results in the Bodenhausen (1988) study, participants focused only on those features that adhered to largely supported patterns and ignored idiosyncratic information that failed to fit. These findings would suggest that effortful attempts to explain feedback tap into our human desire to form generalities and seek patterns. While this behavior is normally adaptive, acting in such a way in complex scenarios can have adverse effects. These findings suggest that, under certain conditions, deliberate processing could increase bias.

Focus of the Present Study

Taken together, the aforementioned studies appear to provide a conflicting picture of decision making under bias. The current study aims to provide further insight into effective strategies for overcoming bias while examining factors that work to strengthen or weaken reliance on bias in decision making.

The first experiment requires participants to deliberately make bad matches, intending to stimulate opposite hypothesis testing (e.g. Entertainment Preference does not matter). Lord et al. (1984) demonstrated that the considering of alternative hypotheses can be successfully promoted through non-explicit suggestion, but did not examine how this strategy could be used to improve accuracy performance. It is supposed that the physicians in the Downar et al (2011) study who were able to improve accuracy through attention to negative feedback were relying on some internal cue or trait to direct their attention. Conversely, students in the Ledet (2013) study were presented with an external cue (a buzzer sound and red screen) as a means to highlight negative feedback, which did not activate the cognitive processes necessary for systematic thinking that would result in high accuracy. In directing participants to establish and answer a disconfirming question internally, flaws in initial hypotheses are expected to become more salient. By disrupting a biased approach to the task, participants are predicted to test alternatives and become more accurate.

The Matchmaker program will be used to create an initial, consistent experimental bias across all participants. The learning phase will be restructured such that one group will be told to make bad matches half the time, allowing these participants to engage in deliberate disconfirmation by considering the initial task question ("what makes a good match?") from an opposite vantage ("what makes a bad match?). By getting participants to pose the disconfirming question themselves, it is possible that more systematic thought processes will be engaged.

A second experiment will further explore the results of Bodenhausen (1988) and Williams et al. (2013) by examining the effects of biased reasoning when participants are warned of feedback error. The Matchmaker task is probabilistic, designed such that there is a 25% chance of receiving incorrect feedback when a correct match is made. In this experiment, some participants will be biased, warned of feedback error, and required to make trial-by-trial assessments of feedback accuracy during the learning phase. Providing a warning should encourage more systematic processing of information and increase overall accuracy (Biros, George, & Zmut, 2013; Stiff, Kim, & Ramesh, 1992). By requiring that participants make ratings of feedback veracity on a trial-by-trial basis, participants should be encouraged to think critically about why the feedback for each trial is true or false. Should high accuracy occur across all trial types, this would indicate that Bodenhausen's (1988) strategy providing item-by-item Likert ratings is an effective way to distribute attention across all information items. However, if bias-congruent trials are answered more accurately than bias-incongruent or bias-irrelevant trials, this would indicate that, similar to the findings of Williams et al. (2013), item-by-item processing leads to increased reliance on biased reasoning when patterns are not always consistent.

EXPERIMENT 1

While direct instruction to consider opposite hypotheses has been shown effective in getting participants to engage in more systematic processing (Lord, Lepper, & Preston, 1984; Koriat, Lichtenstein, & Fischoff, 1980), such instruction has been noted as being cognitively demanding and not practical in real-world scenarios where biases exist (Cox & Popken, 2008; Sanna, Schwartz, & Stocker, 2002). Previous attempts by Ledet (2013) illuminated the persistence of the confirmation bias under various conditions designed to overcome it, whether by manipulating saliency of failures or increasing perceived task difficulty.

The proposed experiment will restructure the Matchmaker task (Ledet, 2013) by encouraging participants to actively generate hypotheses counter to their initial bias. In the new version of the task participants will be encouraged to view their initial (biased) hypothesis from confirming and disconfirming vantages. Errors resulting from a biased approach are predicted to be more salient. This should lead to exploration of alternative hypotheses and increased task accuracy.

While Ledet's version of the Matchmaker task consisted of a priming phase followed by two learning blocks and a test phase, the first block of the learning phase (feedback provided after three matches) will be removed (see Figure 1). Ledet included the initial learning block because he was not sure that the bias would continue with more precise trial-by-trial feedback. However, his study demonstrated that experimentally-induced bias persists throughout the second block with trial-by-trial feedback.

To measure the strength of the initially induced bias, a baseline phase will be added prior to the learning phase in which participants will make matches and indicate confidence in their match but receive no feedback. This will provide a measure of bias strength at the outset of the task and as well as a measure of initial confidence to compare to confidence during the learning and test phases. It is predicted that bias will be strongly present during the baseline phase, as indicated by inaccuracy on bias incongruent trials along with high match confidence.

During the learning phase, participants will assign matches and indicate their confidence in their performance. A biased task approach would be evident in consistent high accuracy and confidence on bias congruent trials and inaccuracy on bias incongruent trials. A systematic approach should reveal increasing accuracy on bias incongruent trials with lower confidence across all trial types and more careful consideration of feedback as indicated by longer reaction times.

Participants

Participants were 181 college students currently enrolled in psychology courses at Louisiana State University participating for course credit. Participants were randomly assigned to one of two groups: a task consistent (n = 90) or task switch (n = 91) group.

Materials and Procedure

The Matchmaker task was presented via EPrime software on standard PCs. Keyboards were used to provide responses. The task was programmed to provide correct feedback 75% of the time and incorrect feedback 25% of the time for all groups. Ratings of confidence and response times were collected in addition to match assignments.

Participants were randomly assigned to two groups, a task-consistent group and a task-switch group. Both groups proceeded first through a priming phase, in which the experimental bias was established (See Figure 2). Specifically, participants were biased to believe that one bachelor (Frank) prefers only matches who list sports as their Entertainment Preference and the other bachelor (James) prefers only matches who list video games as their Entertainment

Preference. After each bachelor was introduced, four sample matches were presented. For Frank, the first sentence for each of his four sample matches emphasized an interest in sports. The first sentence for each of James' sample matches emphasized an interest in video games (see Figure 3).

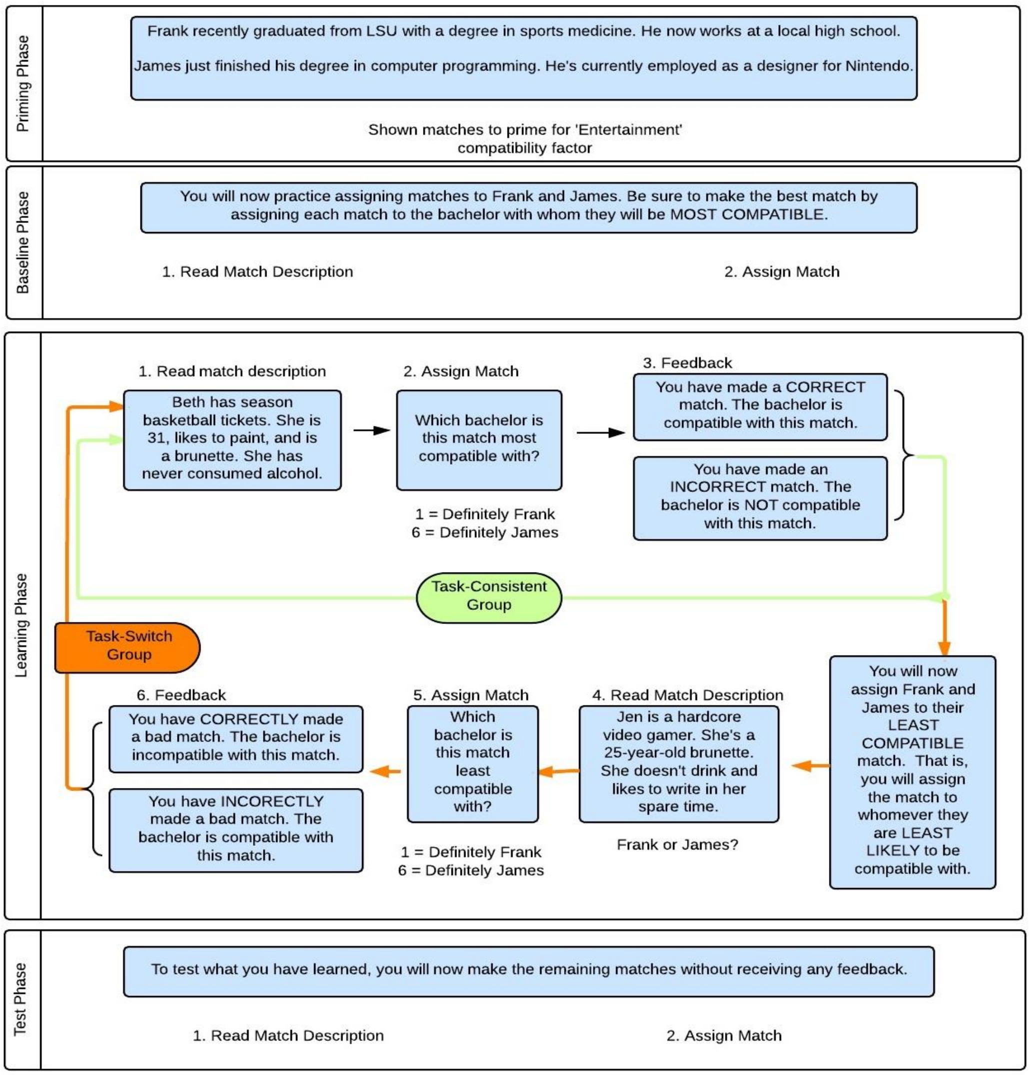


Figure 2. Schematic representation of Matchmaker task for Experiment 1.

Frank recently graduated from LSU with a degree in sports medicine. He now works at a local high school.

-

Mary is an avid fan of LSU football. She's a 26 year old with long blond hair. She writes poetry in her spare time and likes going to bars.

Susan is an LSU cheerleader. She is 21 years old and an amateur singer. She has short brown hair. She does not like to drink.

Karen has season basketball tickets. Karen is 31 and likes to paint. She is a brunette. She has never consumed alcohol.

Lisa is the third generation of her family to play softball at LSU and a proud supporter of the Tigers. She has long brown hair and is 28 years old. She enjoys sketching when she isn't bar hopping.

James just finished his degree in computer programming. He's currently employed as a designer for Nintendo.

-

Jenny considers herself a hardcore video gamer. She's a 25 year old brunette. She doesn't drink but loves singing karaoke.

Heather is an avid player of first person shooter video games. She is a 27 year old blond and likes to write in her spare time. She enjoys drinking with friends.

Beth loves playing games on her xbox. She has long blond hair and enjoys classical music. She does not like to drink and is 23 years old.

Nancy really enjoys playing World of Warcraft and other online games. She has short blond hair and is 31 years old. She enjoys painting in her spare time and is a wine enthusiast.

Figure 3. Bachelor descriptions and sample matches for bias priming.

After priming, participants were to fill out a knowledge questionnaire. This questionnaire listed each compatibility factor (Entertainment Preference, Age, Drinking Habits, Hair Color, and Artistic Hobby) and asked participants to rate how important a particular factor was to each bachelor on a 1 – 5 scale. Additionally, participants were to indicate which specific factor the bachelor likes best (e.g. – Hair Color could be ranked at an importance level of 5, with brown hair as most preferable).

A baseline phase followed, during which participants attempted to make good matches for the bachelors over 30 trials with no feedback. Matches were assigned using a 1 to 6 Likert scale ranging from "Definitely Frank" to "Definitely James" for both groups. During analysis, responses were be collapsed across "Frank" and "James" to provide an accuracy measure. The entire scale was used to assess confidence, with high confidence indicated by use of "definitely" assignments (e.g., "Definitely Frank"), moderate confidence indicated by use of "probably" assignments, (e.g., "Probably Frank"), and low confidence indicated by use of "unsure" (e.g., "Unsure Frank").

Equal exposure was given to bias congruent, bias incongruent, and bias irrelevant trials. Bias congruent trials are those in which the biased compatibility factor (Entertainment) is paired with the correct critical factor (Hair Color). Bias incongruent trials are those in which the biased factor is paired with the incorrect critical factor. Bias irrelevant trials are those in which the bias cannot be utilized (i.e., Entertainment preference is listed as Movies).

The groups then proceeded to the learning phase which consists of 60 trials, again with equal exposure to all trial types. During the learning phase matches were made with those in the task-consistent group making only good matches and those in the task-switch group alternating between attempting to make good and bad matches. All participants were given feedback on a trial-by-trial basis.

Those in the task-consistent group were told only to pair each match with the most compatible bachelor. Feedback was shown stating either "You have made a CORRECT match. The bachelor is compatible with this match" or "You have made an INCORRECT match. The bachelor is NOT compatible with this match".

Those in the task switch-group were to alternate between making "good" and "bad" matches every other trial. Before each good match trial, participants were told "You will now assign Frank and James to their MOST COMPATIBLE match. That is, you will assign the match to whomever they are MOST LIKELY to be compatible with". Before each bad match trial, participants will be told "You will now assign Frank and James to their LEAST COMPATIBLE match. That is, you will assign the match to whomever they are LEAST LIKELY to be compatible with". On good match trials, feedback was the same as for the task consistent group. For "bad match" trials, feedback stated "You have CORRECTLY made a bad match. The bachelor is incompatible with this match" or "You have INCORECTLY made a bad match. The bachelor is compatible with this match".

Following the learning phase, participants then completed a test phase with instructions to make the most compatible match for the bachelors on 30 trials, with no feedback given. Lastly, participants were to complete a written system knowledge test, indicating the probability for each of the bachelors in liking a match based on a particular compatibility factor. Additionally, participants were able to make open-ended comments indicating their interpretation of a best match for each bachelor

Results

All results are reported as significant at $p < .05$ unless otherwise indicated. For all instances when Mauchly's test indicated that sphericity had been violated, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Accuracy Performance. In order examine how bias and warning affected performance throughout the task, separate 2 (group: task-consistent, task-switch) x 3 (phase: baseline, learning, test) mixed-model ANOVAs were conducted on the accuracy of bias incongruent, bias congruent, and bias irrelevant trials.

First, a 2 x 3 mixed-model ANOVA examined performance on bias incongruent trials throughout the task phase for task consistent and task switch groups (see Figure 4).

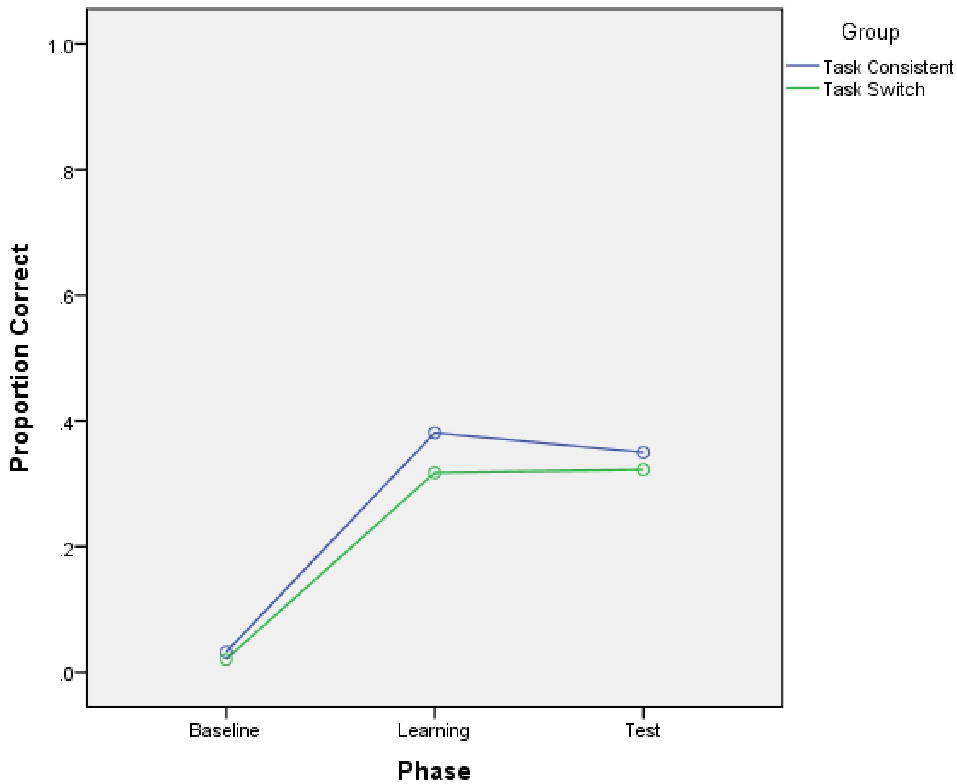


Figure 4. Accuracy performance on bias incongruent trials for Experiment 1.

There was a significant main effect of phase in incongruent trial accuracy, $F(1.56, 277.68) = 228.96$, $MSE = .03$ $\eta_p^2 = .56$. Pairwise comparisons revealed that accuracy was significantly lower at baseline ($M = .03$, $SE = .01$) than at learning ($M = .35$, $SE = .02$) or test ($M = .34$, $SE = .34$, $SE = .01$), but that learning and test accuracy did not differ from each other. There was no main effect of group, $F(1, 179) = 2.55$, and no significant interaction between phase and group, $F(1.55, 277.68) = 1.22$.

In order to determine if making bad matches had any effect on accuracy for bias incongruent trial types, a dependent samples t-test was conducted on the accuracy of good matches and bad matches for learning trials in the task-switch group. Participants in this group were found to be more accurate when making bad matches on bias incongruent trials (M = .40, SE = .02) than when making good matches (M = .25, SE = .02), $t(90) = -5.91$, d = 1.25.

A second 2 x 3 mixed-model ANOVA compared accuracy on bias congruent trials at baseline, learning, and test for task-consistent and task-switch groups (see Figure 5).

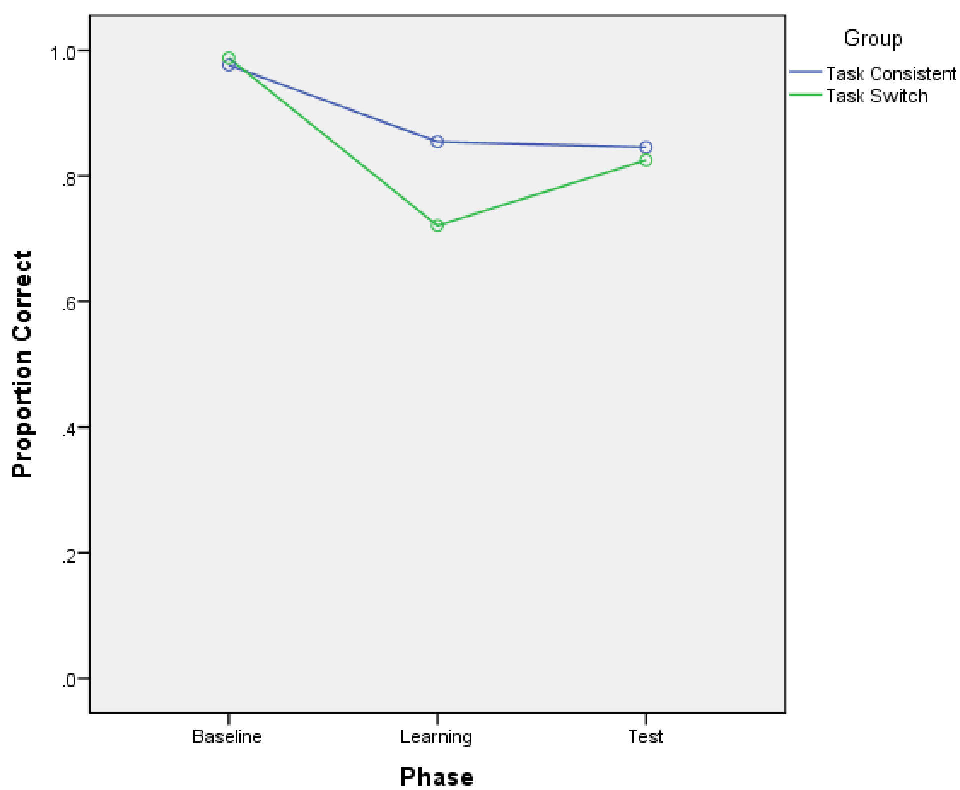


Figure 5. Accuracy performance on bias congruent trials for Experiment 1.

A main effect of phase was observed for accuracy on bias congruent trials, $F(1.82, 325.05) = 140.61$, MSE = .02, $\eta_p^2 = .44$. Pairwise comparisons revealed that all phases significantly differed from one another, with accuracy performance highest at baseline (M = .98, *SE* = .00), dropping at learning (*M* = .79, *SE* = .01), and rising again at test, (*M* = .84, *SE* = .01). A main effect of group was observed, $F(1, 179) = 13.2$, $MSE = .01$, $\eta_p^2 = .07$, with pairwise comparisons demonstrating that that the overall performance of the task-consistent group was significantly better than the task-switch group on bias congruent trials throughout the task (*M* = .89, *SE* = .01 and *M* = .86, *SE* = .01 respectively). A significant interaction was also observed between phase and group, $F(1.82, 325.04) = 19.82$, $MSE = .29$, $\eta_p^2 = .10$, indicating that accuracy on bias congruent trials at each phase of the task differed between groups. Independent samples t-tests compared accuracy on bias congruent trials and revealed that the task-consistent group was significantly more accurate on these trials at leaning (*M* = .85, *SE* = .13) than the task-switch group (*M* = .72, *SE* = .13), $t(179) = 6.91$, $d = 1.03$. No differences were found between the groups at baseline, $t(179) = .15$, or test, $t(179) = .79$.

In order to determine if making bad matches had any effect on accuracy for bias congruent trial types, a dependent samples t-test was conducted on the accuracy of good matches and bad matches for learning trials in the task-switch group. Participants in this group were found to be more accurate when making good matches on congruent trials (*M* = .83, *SE* = .02) than when making bad matches (*M* = .61, *SE* = .02), $t(90) = 8.60$, $d = 1.81$.

A 2 x 3 mixed-model ANOVA was conducted to analyze performance on bias irrelevant trials at baseline, learning, and test between the two groups (see Figure 6).

A main effect of phase was observed for accuracy on bias irrelevant trials, $F(1.83, 328) = 12.77$, $MSE = .03$, $\eta_p^2 = .07$. Pairwise comparisons revealed that accuracy on these trials was significantly worse at baseline (*M* = .53, *SE* = .02) than at learning (*M* = .60, *SE* = .01) or test (*M* = .62, *SE* = .02), but that no significant differences existed between learning and test.

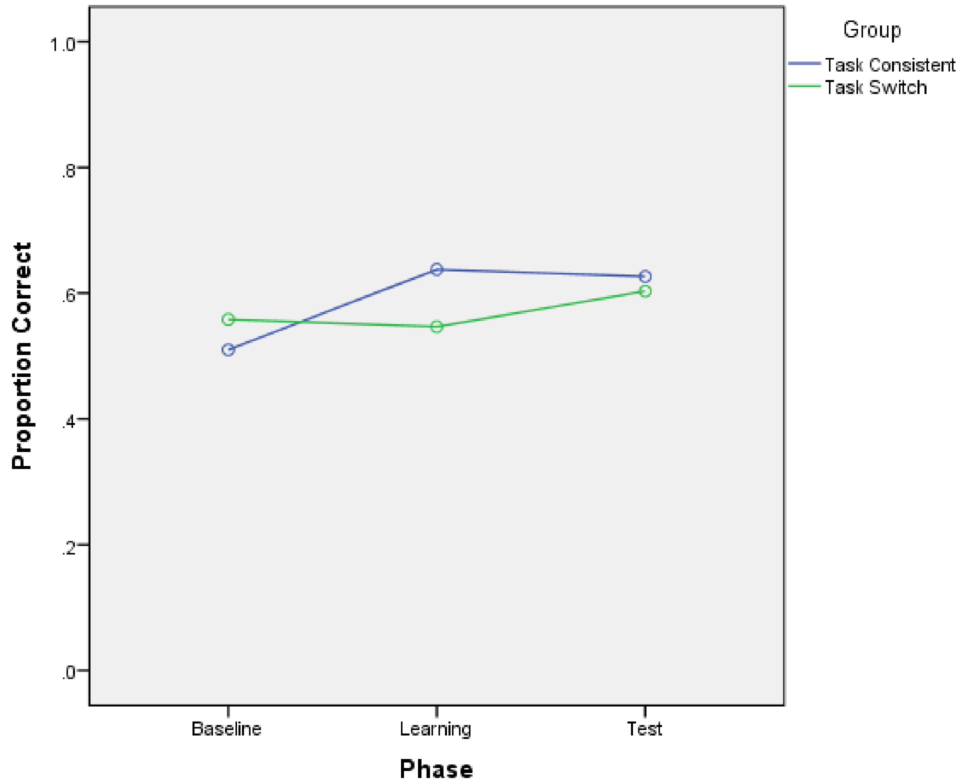


Figure 6. Accuracy performance on bias irrelevant trials for Experiment 1.

While there was no main effect of group, $F(1, 179) = 1.09$, a significant interaction was observed between phase and group, $F(1.83, 328) = 8.91$, $MSE = .03$, $\eta_p^2 = .05$, suggesting that accuracy performance on bias irrelevant trials at each phase of the task differed between groups. Independent samples t-tests revealed that the groups significantly differed in accuracy on bias irrelevant trials at the learning phase, $t(179) = 3.89$, $d = .58$, with the task-consistent group achieving higher accuracy ($M = .64$, $SE = .17$) than the task-switch group ($M = .55$, $SE = .15$). No significant differences in accuracy were found between the two groups at baseline, $t(179) = -1.63$, or at test, $t(179) = .74$.

In order to determine if making bad matches had an effect on accuracy for bias irrelevant trial types, a dependent samples t-test was conducted on the accuracy of good matches and bad matches for learning trials in the task-switch group. Participants in this group were found to be more accurate when making good matches on irrelevant trials (M = .58, SD = .02) than when making bad matches (M = .51, SE = .02), $t(90) = 2.76$, d = .58.

Confidence. In order to determine how confidence changed throughout the task, separate 2 (group: task-consistent, task-switch) x 3 (phase: baseline, learning, test) mixed-model ANOVAs were conducted on confidence ratings for bias incongruent, bias congruent, and bias irrelevant trials.

A 2 x 3 mixed-model ANOVA examined confidence on bias incongruent trials for both groups at each phase of the task (see Figure 7). A significant main effect of phase was found, $F(1.73, 309.18) = 73.91$, $MSE = .12$, $\eta_p^2 = .30$. Pairwise comparisons revealed significant decreases in confidence on bias incongruent trials as participants progressed from baseline (M = 2.42, SE = .03) to learning (M = 2.09, SE = .03), and from learning to test (M = 2.03, SE = .03). A main effect of group was found, $F(1, 179) = 4.85$, $MSE = .10$, $\eta_p^2 = .03$, and pairwise comparisons revealed that the task-switch group was significantly more confident overall (M = 2.23, SE = .03) than the task consistent group (M = 2.13, SE = .03). The interaction between phase and group was also found to be significant, $F(1.27, 309.18) = 3.24$, $MSE = .12$, $\eta_p^2 = .02$, suggesting that the task-consistent and task-switch groups were differing their confidence on bias incongruent trials during the phases of the task.

A follow-up independent samples t-test revealed that task-consistent groups were significantly less confident (M = 2.00, SD = .42) than task-switch groups (M = 2.19, SD = .41) on bias incongruent trials during the learning phase, $t(179) = -2.99$, d = -.44, but that the two groups did not differ in confidence at baseline, $t(179) = .18$, or test, $t(179) = -1.71$.

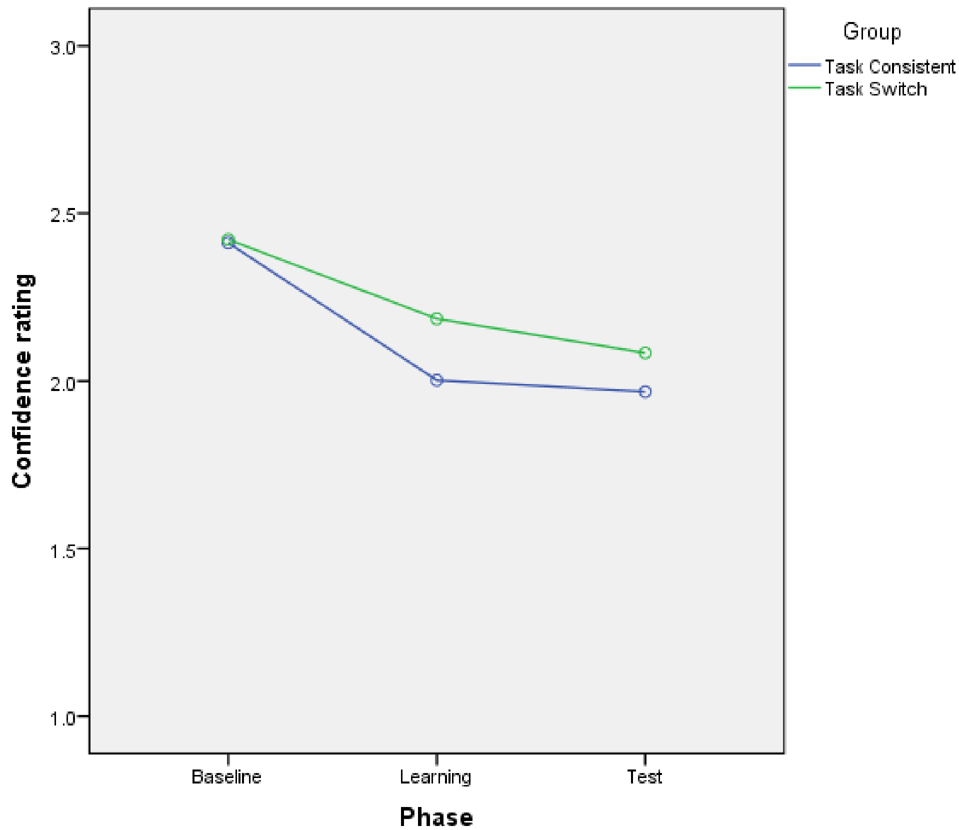


Figure 7. Confidence on bias incongruent trials for Experiment 1.

An additional dependent samples t-test was conducted in order to determine if confidence differed for participants in the task-switch group when making good or bad matches on bias incongruent learning trials. Confidence was significantly higher when making good matches (*M* = 2.26, *SD* = .42) than when making bad matches (*M* = 2.11, *SD* = .45) for these trial types, *t*(90) = 4.44, *d* = .93.

A second ANOVA examined confidence on bias congruent trials for both groups at each phase of the task (see Figure 8). The main effect of phase was found to be significant, *F*(1.81, 324.41) = 67.91, *MSE* = .10, η_p^2 = .28. Pairwise comparisons revealed that confidence on bias congruent trials decreased significantly both from baseline (*M* = 2.45, *SE* = .03) to learning (*M* = 2.22, *SE* = .03) and from learning to test (*M* = 2.08, *SE* = .03).

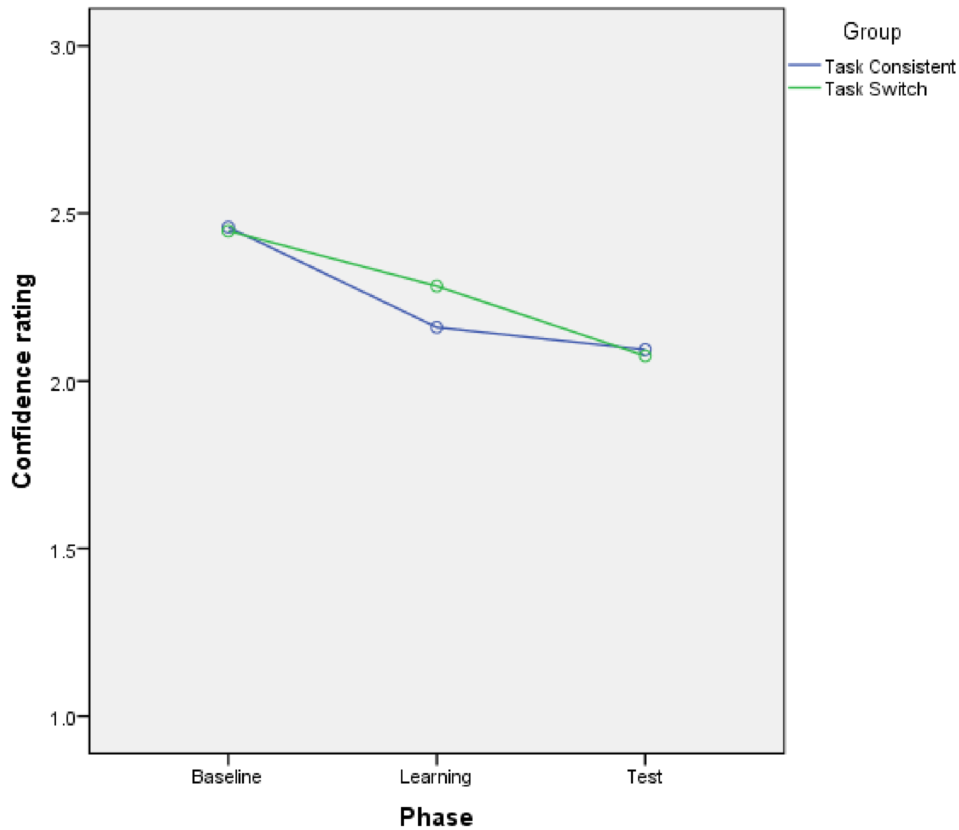


Figure 8. Confidence on bias congruent trials for Experiment 1.

While no main effect was found for group, $F(1, 179) = .50$, the interaction between phase and group was found to be significant, $F(1.82, 324.41) = 3.11$, $MSE = .10$, $\eta_p^2 = .02$, suggesting that the task-consistent and task-switch groups were differing in their confidence on bias congruent trials during the phases of the task. Follow-up independent samples t-tests revealed that task-consistent groups were significantly less confident ($M = 2.16$, $SD = .40$) than task-switch groups ($M = 2.28$, $SD = .36$) on bias congruent trials during the learning phase, $t(179) = -2.16$, $d = -.32$, but that the two groups did not differ in confidence at baseline, $t(179) = .23$, or test, $t(179) = .30$.

An additional dependent samples t-test was conducted in order to determine if confidence differed for participants in the task-switch group when making good or bad matches on bias congruent trials. Confidence was significantly higher when making good matches (*M* = 2.36, *SD* = .35) than when making bad matches (*M* = 2.20, *SD* = .45) for these trial types, *t*(90) = 4.59, *d* = .98.

A last ANOVA examined confidence on bias irrelevant trials for both groups at each phase of the task (see Figure 9). A main effect of phase was found to be significant for confidence on bias irrelevant trials, *F*(1.80, 321.54) = 45.23, *MSE* = .12, η_p^2 = .20. Pairwise comparisons revealed that confidence increased significantly from baseline (*M* = 1.70, *SE* = .02) to learning (*M* = 2.0, *SE* = .03), but not from learning to test (*M* = 2.0, *SE* = .03). The main effect of group was not significant, *F*(1, 179) = 1.66, however, a significant interaction was found between phase and group, *F*(1.80, 321.54) = 5.40, *MSE* = .12, η_p^2 = .03, suggesting that the two groups were assigning different levels of confidence to bias irrelevant trials at different phases of the task. Follow-up independent samples t-tests revealed the confidence at test was trending toward being significantly higher at test for the task-switch group (*M* = 2.05, *SD* = .42) than for task-consistent group (*M* = 1.91, *SD* = .50), *t*(179) = -1.94, *d* = -.29, but significant differences did not exist at baseline, *t*(179) = 1.40, or learning, *t*(179) = -1.80.

An additional dependent samples t-test was conducted in order to determine if confidence differed for participants in the task-switch group when making good or bad matches on bias irrelevant learning trials. Confidence was significantly higher when making good matches (*M* = 2.08, *SD* = .40) than when making bad matches (*M* = 2.00, *SD* = .46) for these trial types, *t*(90) = 3.21, *d* = .68.

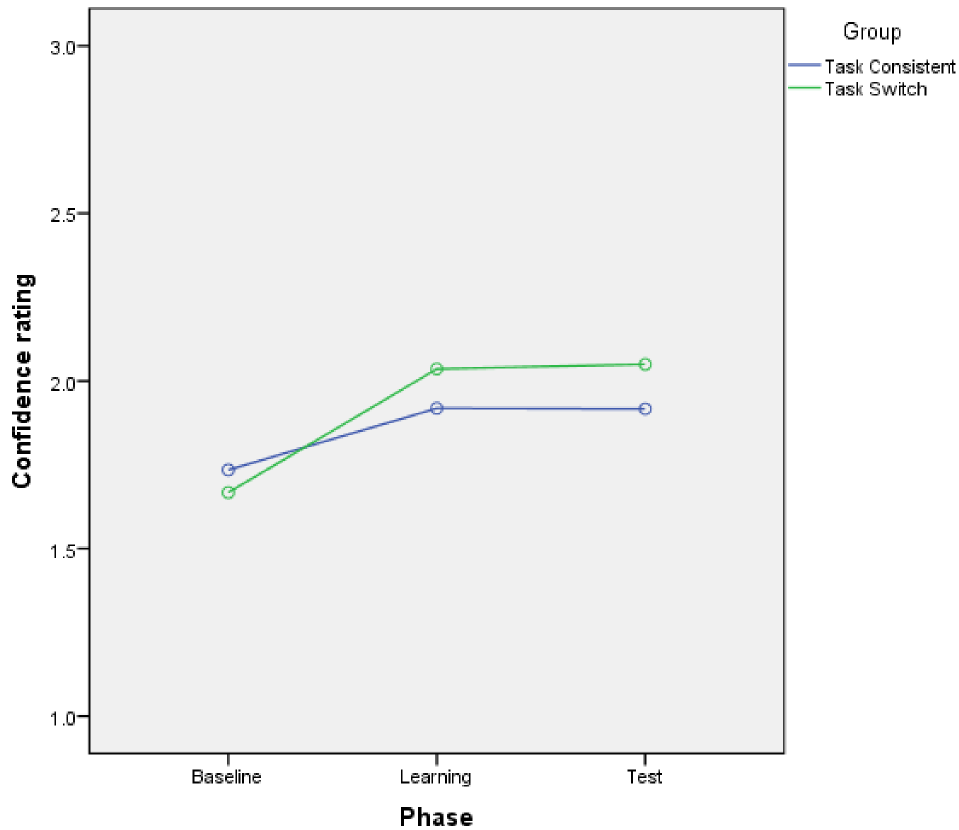


Figure 9. Confidence on bias irrelevant trials for Experiment 1.

Response Time. In order to determine how response time changed throughout the task, separate 2 (group: task-consistent, task-switch) x 3 (phase: baseline, learning, test) mixed-model ANOVAs were conducted on response times for bias incongruent, bias congruent, and bias irrelevant trials.

A first ANOVA examined response time on bias incongruent trials for both groups at each phase of the task (see Figure 10). The main effect of phase was found to be significant for bias incongruent trials, $F(1.82, 325.80) = 2.94$, $MSE = 4233.66$, $\eta_p^2 = .02$, with pairwise comparisons indicating that participants were significantly faster at test ($M = 4.08$, $SE = .14$) than at learning ($M = 4.36$, $SE = .13$).

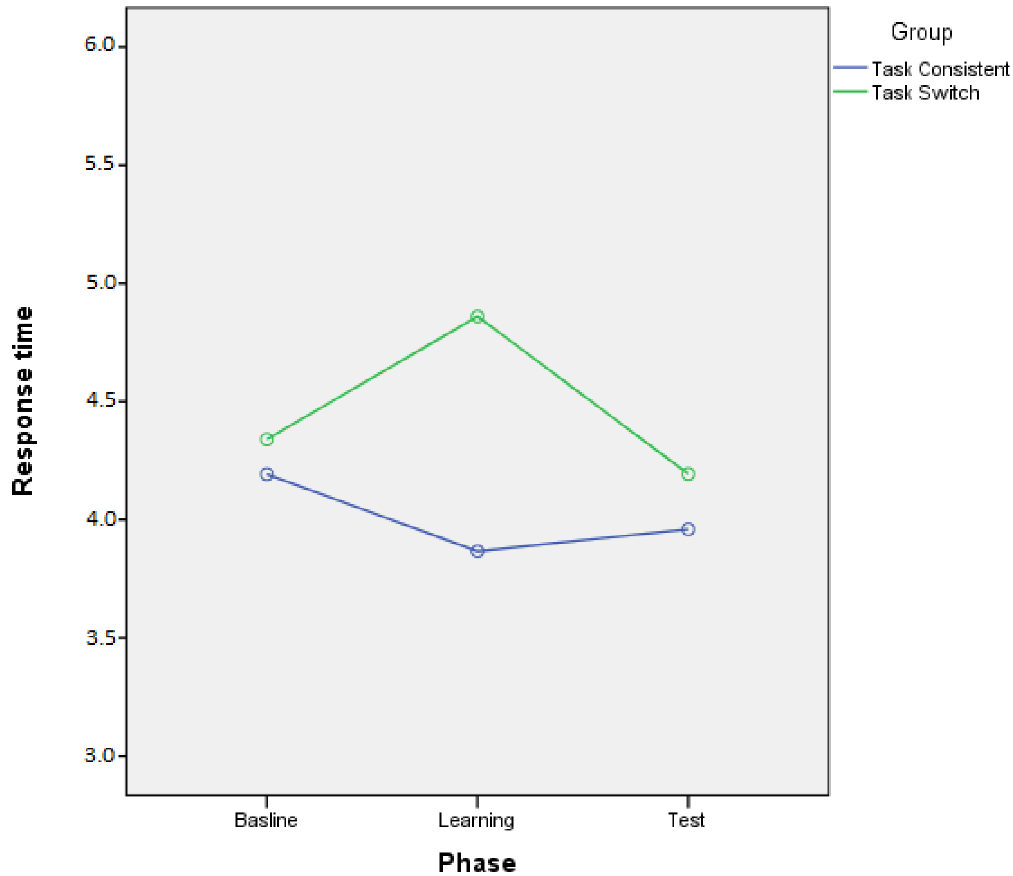


Figure 10. Response time in seconds on bias incongruent trials for Experiment 1.

A main effect was also found for group, $F(1, 179) = 3.91$, $MSE = 9516.17$, $\eta_p^2 = .02$, with task-consistent participants exhibiting faster response times ($M = 4.01$, $SE = .16$) than task-switch participants ($M = 4.46$, $SE = .16$). The interaction between phase and group was also found to be significant, $F(1.82, 325.80) = 7.48$, $MSE = 10785.92$, $\eta_p^2 = .04$, suggesting that the two groups were exhibiting different response times on bias incongruent trials at different phases of the task. Follow-up independent samples t-tests revealed that, during the learning phase, the task-consistent group was significantly faster ($M = 3.87$, $SD = 1.32$) than the task-switch group ($M = 4.86$, $SD = 2.05$), $t(179) = -3.87$, $d = -.58$, but that the groups did not differ in response times at baseline, $t(179) = -.55$, or test, $t(179) = -.82$.

An additional dependent samples t-test was performed to determine if response time differed for participants in the task-switch group when making good or bad matches on bias incongruent trials. Response times differed significantly between good and bad matches, $t(90) = -6.04$, $d = -1.27$, with good matches being made faster ($M = 4.51$, $SD = 2.05$) than bad matches ($M = 5.48$, $SD = 2.43$).

A second ANOVA examined response time on bias congruent trials for both groups at each phase of the task (see Figure 11). A main effect of phase was found for response times on bias congruent trials, $F(1.72, 307.36) = 21.36$, $MSE = 27569.41$, $\eta_p^2 = .12$, with pairwise comparisons indicating that response times on these trials decreased significantly from baseline ($M = 4.60$, $SE = .14$) to learning, ($M = 4.07$, $SE = .12$), and from learning to test ($M = 3.83$, $SE = .12$). A main effect of group was also found, $F(1, 179) = 6.13$, $MSE = 12165.50$, $\eta_p^2 = .03$. Pairwise comparisons revealed that the task-consistent group exhibited significantly faster reaction times overall on bias congruent trials ($M = 3.91$, $SE = .21$) than task-switch groups ($M = 4.43$, $SE = .15$).

The interaction between phase and group was also found to be significant, $F(1.72, 307.36) = 4.00$, $MSE = 5169.06$, $\eta_p^2 = .02$, suggesting that the two groups were exhibiting different response times on bias congruent trials at different phases of the task. Follow-up independent samples t-tests revealed that response times differed significantly between the groups during the learning phase, $t(179) = -3.90$, $d = -.58$, with task-consistent participants exhibiting faster times ($M = 3.63$, $SD = 1.72$) than task-switch participants ($M = 4.51$, $SD = 1.80$). There were no differences in response time between the groups at baseline, $t(179) = -1.65$, or test, $t(179) = -.88$.

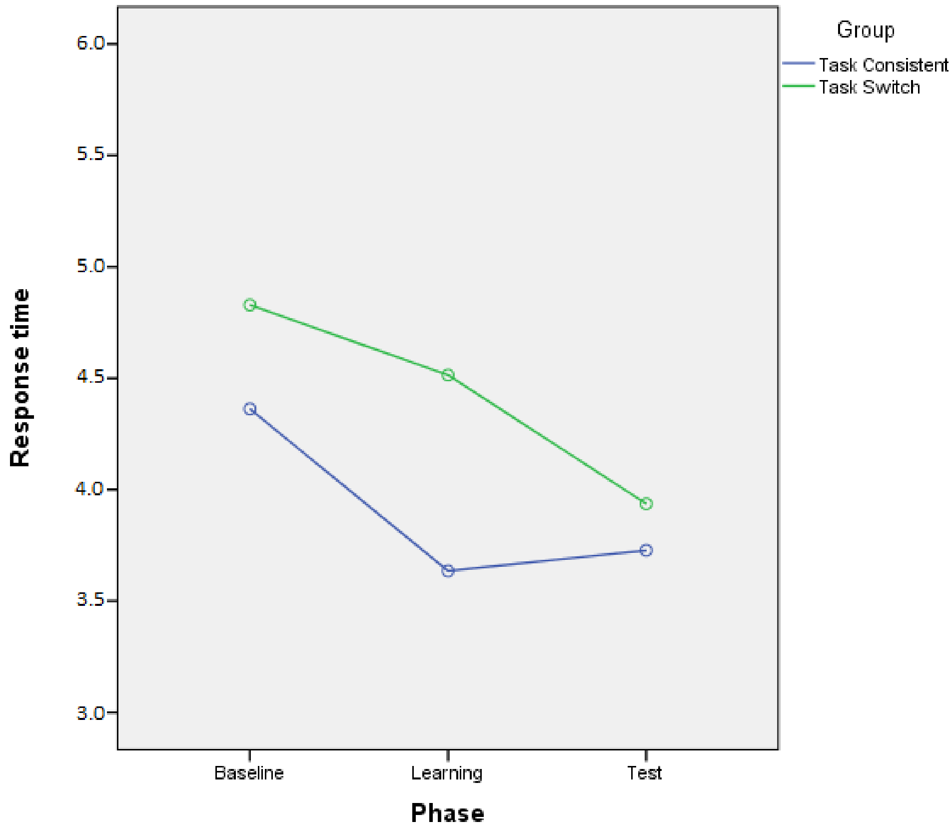


Figure 11. Response time in seconds on bias congruent trials for Experiment 1.

An additional dependent samples t-test was performed to determine if response time differed for participants in the task-switch group when making good or bad matches on bias congruent learning trials. Participants did not differ in their response times when making good or bad matches, $t(90) = -1.36$, $d = -.29$.

A third ANOVA examined response time on bias irrelevant trials for both groups at each phase of the task (see Figure 12).

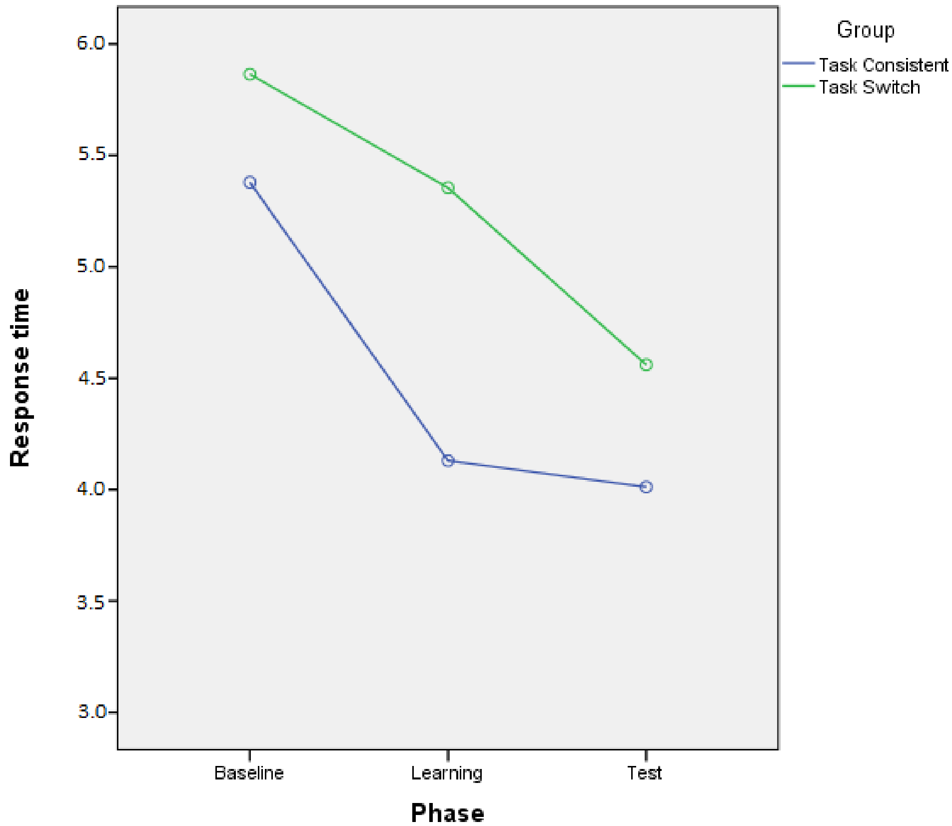


Figure 12. Response time in seconds on bias irrelevant trials for Experiment 1.

A main effect of phase was found, *F*(2.358) = 48.75, *MSE* = 83483.10, η_p^2 = .21, with pairwise comparisons revealing that response time significantly decreased from baseline (*M* = 5.62, *SE* = .17) to learning (*M* = 4.74, *SE* = .14), and from learning to test (*M* = 4.29, *SE* = .15). A main effect of group was also found, *F*(1, 179) = 8.41, *MSE* = 25649.63, η_p^2 = .05, with pairwise comparisons revealing that the task-consistent group was faster overall (*M* = 4.51, *SE* = .18) as compared to the task-switch group (*M* = 5.26, *SE* = .18). The interaction between phase and group was also found to be significant, *F*(2, 358) = 4.43, *MSE* = 7593.28, η_p^2 = .02, suggesting that the two groups were exhibiting different response times on bias irrelevant trials at different phases of the task. Follow-up independent samples t-tests revealed that the groups exhibited significant differences in response time on bias irrelevant trials during the learning phase, $t(179) = -4.42$, $d = -.66$, with the task-consistent group responding faster ($M = 4.13$, $SD = 1.43$) than the task-switch group ($M = 5.35$, $SD = 2.21$).

An additional dependent samples t-test was performed to determine if response time differed for participants in the task-switch group when making good or bad matches on bias irrelevant learning trials. Participants did not differ in their response times when making good or bad matches, $t(90) = -1.40$, $d = -.30$.

Bias Strength. In order to determine how bias strength varied throughout the task, a 2 (group: task-consistent, task-switch) x 3 (phase: baseline, learning, test) mixed-model ANOVA was conducted on the computed difference score (accuracy on bias congruent trials minus accuracy on bias incongruent trials) at baseline, learning, and test for both groups (see Figure 13).

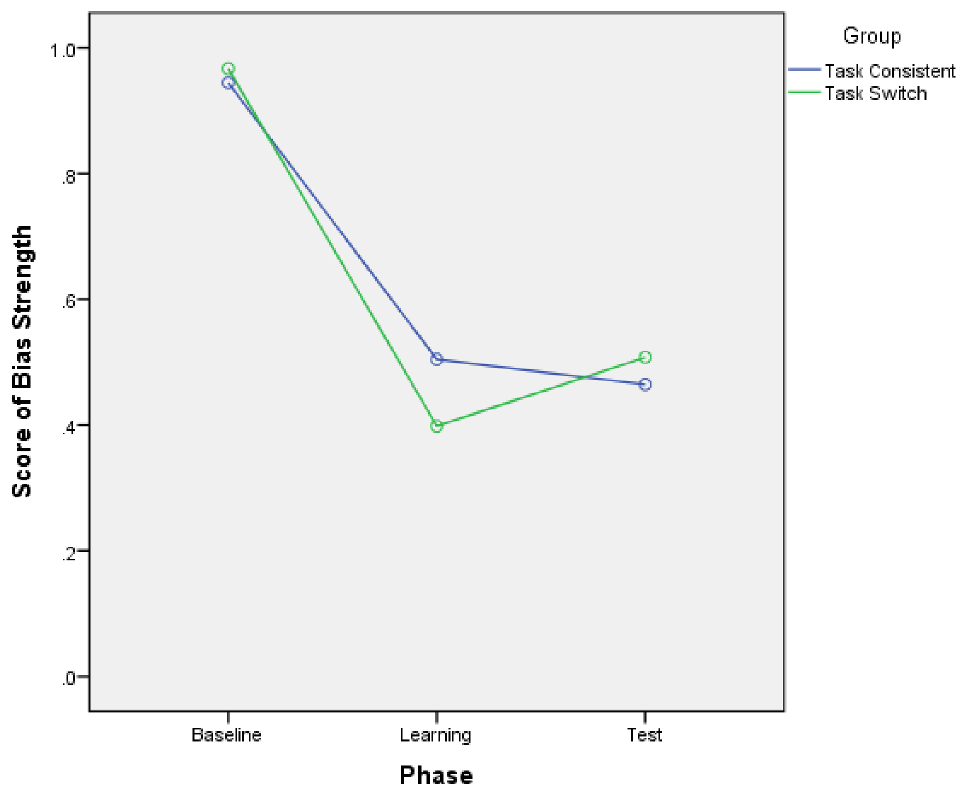


Figure 13. Bias strength score throughout task for Experiment 1.

A significant main effect of phase was found, *F*(1.76, 314.32) = 328.54, *MSE* = .05, η_p^2 =.65. Pairwise comparisons revealed that bias strength was significantly higher at baseline (*M* = .96, *SE* = .01) than at learning (*M* = .45, *SE* = .02) or test (*M* = .49, *SE* = .03), but that strength was significantly different between learning and test. While no main effect of group was found, *F*(1, 179) = .285, a significant interaction between phase and group was found, *F*(2, 358) = 6.78, *MSE* = .05, η_p^2 = .04, indicating that bias strength differed between groups throughout the phases of the task. Follow-up independent samples t-tests revealed that bias strength differed between the groups during the learning phase, *t*(179) = 3.11, *d* = .46, but not at baseline, *t*(179) = -1.43, or test, *t*(179) = -.88.

Explicit Measure of Bias. In order to compare participants' choice bias as indicated by the above performance measures on the computerized matching task to their explicit indication of bias, analyses were performed on ratings of compatibility factor importance provided by participants in writing on hypothesis tests at baseline and test phases. Composite scores of compatibility factor importance were calculated for each compatibility factor by averaging the importance ratings for both bachelors. Of particular interest are the compatibility factors of Entertainment Preference (biased factor) and Hair Color (critical factor).

A 2 (group: task-consistent, task-switch) x 2 (phase: baseline, test) mixed-model ANOVA was conducted for ratings of Entertainment Preference at baseline and test for both groups (see Figure 14). All results are reported as significant at $p < .05$ unless otherwise indicated.

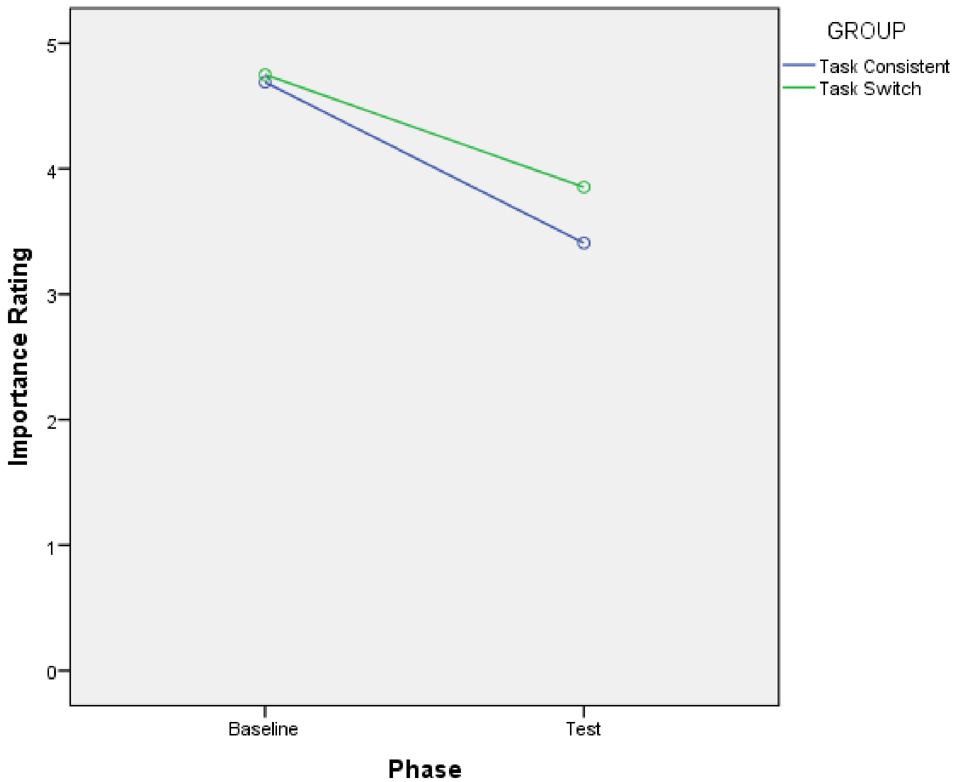


Figure 14. Importance rating of entertainment preference for Experiment 1.

A main effect of phase was found on ratings of Entertainment Preference importance, $F(1, 167) = 147.82$, $MSE = .68$, $\eta_p^2 = .47$. Pairwise comparisons indicated that Entertainment Preference was rated as significantly more important at baseline ($M = 4.72$, $SE = .05$) than at test ($M = 3.63$, $SE = .08$). A significant main effect was found for group, $F(1, 167) = 7.04$, $MSE = .38$, $\eta_p^2 = .04$, indicating that task consistent and task switch groups rated Entertainment Preference differently overall from baseline to test. Pairwise comparisons revealed that the task-switch groups rated this factor as significantly more important ($M = 4.30$, $SE = .07$) than the task-consistent group ($M = 4.05$, $SE = .07$). The interaction between phase and group was also found to be significant, $F(1, 167) = 4.63$, $MSE = .68$, $\eta_p^2 = .03$, suggesting that groups were rating Entertainment Preference differently at baseline than at test. Post-hoc independent samples t-tests revealed that task-consistent and task-switch groups did not rate Entertainment Preference differently at baseline (*M* = 4.68, *SD* = .72 and *M* = 4.75, *SE* = .46 respectively), *t*(170) = -.8, *d* = -.12. However, the task consistent group rated this factor lower at test (*M* = 3.41, *SD* = 1.10) than did the task switch group (*M* = 3.88, *SD* = .99). This difference was found to be statistically significant, *t*(169) = -2.97, *d* = -.46.

A 2 (group: task-consistent, task-switch) x 2 (phase: baseline, test) mixed-model ANOVA was also conducted for rating of Hair Color importance for the groups at baseline and test (see Figure 15).

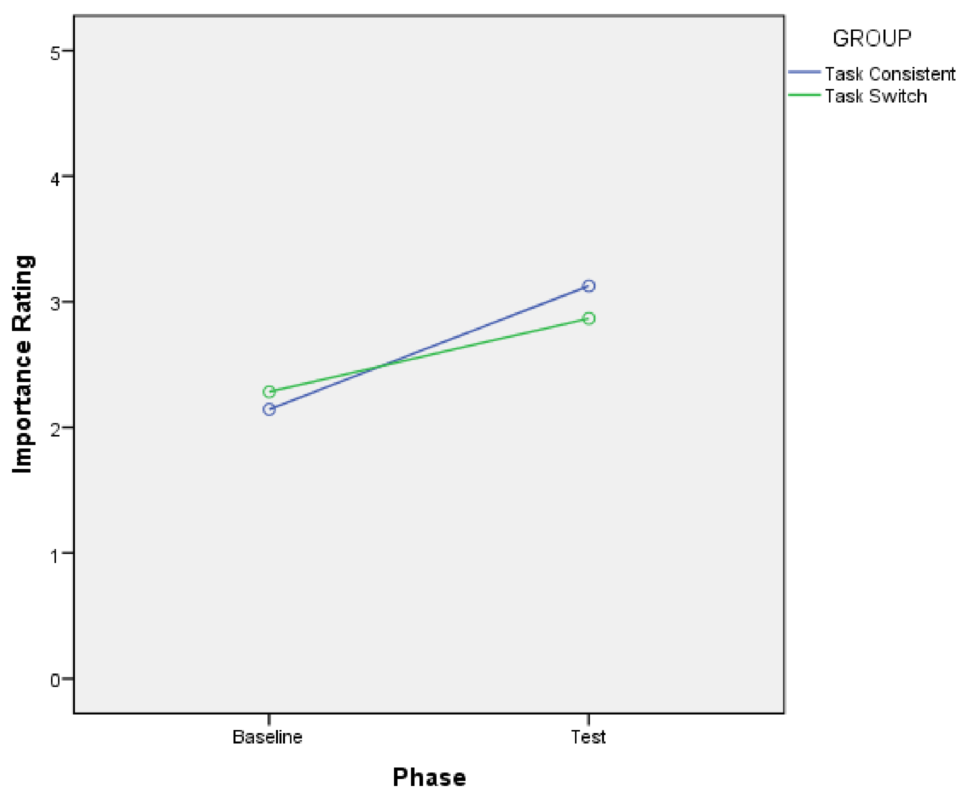


Figure 15. Importance rating of hair color for Experiment 1.

This analysis revealed a main effect of phase, $F(1, 168) = 55.97$, $MSE = .93$, $\eta_p^2 = .25$, with pairwise comparisons demonstrating that both groups rated this compatibility factor as lower at baseline (*M* = 2.21, *SE* = .10) than at test (*M* = 3.0, *SE* = .11). The main effect of group was insignificant, $F(1, 168) = .12$, as was the interaction between phase and group, $F(1, 168) = 3.37$.

Although participants rated the biased compatibility factor lower at test than at baseline and believed that the critical factor was more important at test than at baseline, descriptive measures show that the biased factor was still seen as the most important factor in making a good match at test (see Figure 16). The number of participants that endorsed the biased factor important (rating it 4 or 5 on written tests after task completion) was greater for Entertainment Preference than any other compatibility factor.

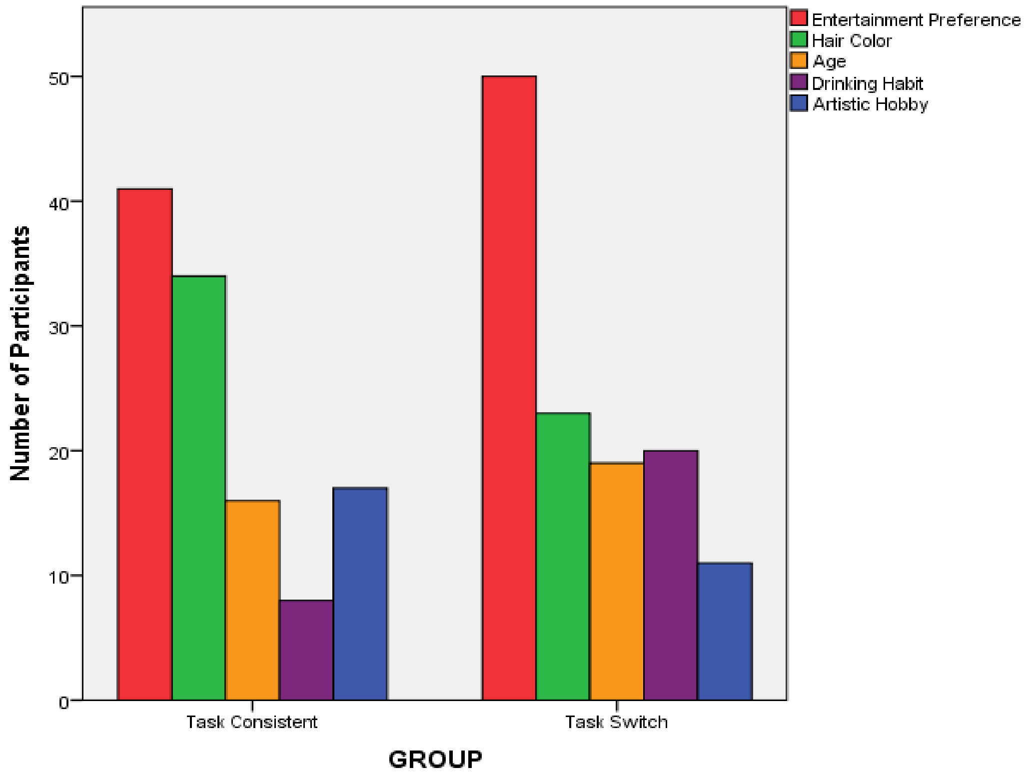


Figure 16. Compatibility factors endorsed as important at test.

DISCUSSION

Encouraging participants to consider opposite hypotheses by making both good and bad matches did not appear to be a particularly potent strategy for reducing reliance on bias in decision making. However, both the task consistent and the task switch groups demonstrated a reduction in bias strength as the task progressed, suggesting that simply proceeding through the task and receiving feedback was enough to lower reliance on bias by nearly half. Although bias was still being utilized, it appears as though some learning was occurring for both groups, as demonstrated by increased accuracy on bias incongruent trials over the course of the task and lower ratings of importance for the biased factor of Entertainment Preference at test. Participants in the task switch group also demonstrated lower accuracy on bias congruent trials during the learning phase, suggesting that alternative hypothesis testing may have been occurring at this time. If such testing was in fact taking place, the gains in doing so were small. Participants in the task switch group showed a return to reliance on bias at test, evidenced by increased accuracy performance on bias congruent trials at this phase, as well as high and frequent ratings of the importance of the biased factor of Entertainment Preference at this time.

In general, confidence for both groups decreased as the task progressed. However the task switch group was shown to display higher ratings of confidence on both bias congruent and bias incongruent trial types during the learning phase than the task consistent group. Because lower confidence is associated with more systematic thinking, this finding provides further evidence that the task switch group was not utilizing systematic reasoning processes throughout the task. Similarly, slower response times in the task switch group are likely due to the effort required to keep the proper goal in mind for each trial (e.g. making a good or bad match) rather than engagement in more cognitively difficult information processing.

EXPERIMENT 2

While some research has shown that providing warning of error can lead to increased accuracy in decision making (Biros, George, & Zmut, 2013), other studies have shown that warnings may not lead to the same results during decision making under bias (Stiff, Kim, and Ramesh, 1992; Toris & DePaulo, 1984). Although presenting warnings under bias can lower confidence and lead to deeper cognitive processing, it does not seem to lead to increased accuracy. However, while Stiff et al (1992) found that warning biased participants of error in feedback did not lead to increases in accuracy, they did not examine if elaborative encoding of instances containing perceived error would improve performance.

Bodenhausen (1988) found that trial-by-trial ratings of information items can allow more attention to be devoted to all items, not just those consistent with a bias. This should lead to an integrated and precise understanding of information that serves to overcome the heuristic processing that is a hallmark of biased reasoning. More recent work by Williams, Lombrozo, & Rehder (2013) has found that focusing attention by generating explanations of feedback on a trial-by-trial basis has the opposite effect: bias inconsistent items were less likely to be recalled and a generalized impression was formed that focused on only those items that justified bias. The proposed design seeks to explore these findings in a complex decision making task.

Experiment 2 used the Matchmaker task to examine the effect of warnings on biased decision making. In addition, this experiment aimed to determine how focusing attention on the information gleaned from each learning instance affects the interpretation of feedback veracity.

Under the Williams et al. (2013) theory of biased processing, we would predict that providing an assessment for each instance of feedback will lead to a discounting of negative feedback shown by low accuracy on bias incongruent trials, short reaction times, high match confidence, and low confidence of feedback when told matches are incorrect. However, if making assessments for each instance of feedback encourages more systematic deliberation of information, as indicated by Bodenhausen (1988), accuracy on bias incongruent trials should increase as the task progresses and a holistic understanding of the task occurs. This systematic processing would be evident as well in longer reaction times and lower match confidence.

Participants

Participants were 203 college students currently enrolled in psychology courses at Louisiana State University participating for course credit. Participants were randomly assigned to one of four groups: No Warning/Bias (*n* = 52), No Warning/No Bias (*n* = 51), Warning/Bias (*n* = 52), and Warning/No Bias (*n* = 48).

Materials and Procedure

The same materials used in Experiment 1 were used in Experiment 2.

Participants assigned to the Biased conditions proceeded through the priming phase as described in Experiment 1 (see Figure 17). They were biased to believe that Frank prefers only matches who list sports as an Entertainment Preference and that James prefers only matches who list video games as an Entertainment Preference. Participants in the Unbiased conditions were not biased to believe the bachelors preferred any one compatibility factor. Additionally, Unbiased participants viewed descriptions for sample matches that were randomized across compatibility factors and told that these are simply random example matches (see Figure 18). After the priming phase, a knowledge test was conducted as in Experiment 1.

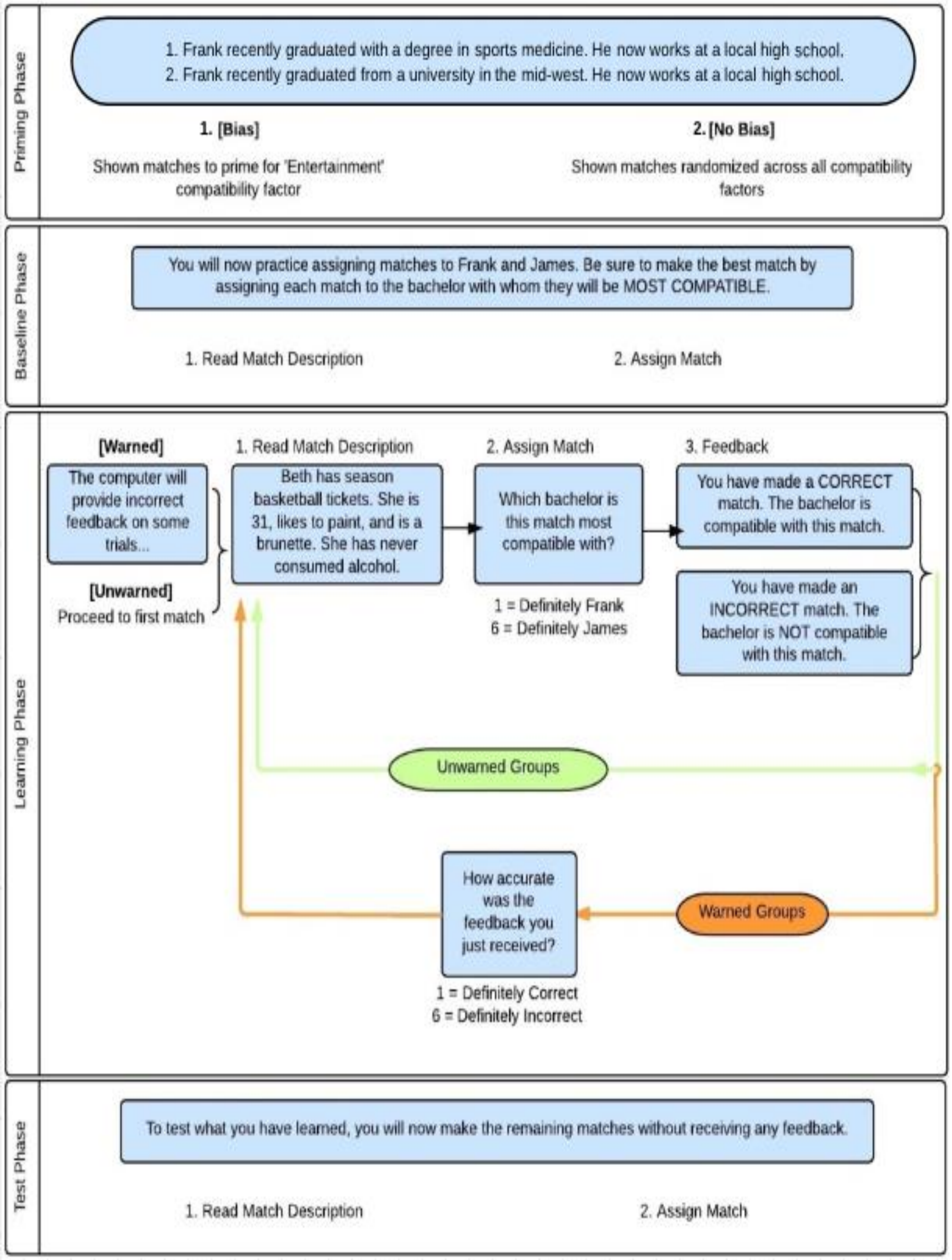
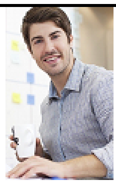


Figure 17. Schematic representation of Matchmaker for Experiment 2.

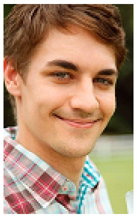


Frank recently graduated from a university in the mid-west. He now works at a local high school.
-
Mary is an avid fan of LSU football. She's a 26 year old with long blond hair. She writes poetry in her spare time and likes going to bars.

Karen is a brunette who likes to paint. She is 31 years old and has never consumed alcohol. She enjoys film noir movies.

Lisa is a 28 year old who enjoys bar hopping. She enjoys sketching in her spare time. She has blond hair and enjoys watching Netflix.

Susan is 21 years old and an amateur singer. She has short brown hair. She loves playing games on her xbox. She does not like to drink.



James just finished his degree at a college in his hometown. He's currently employed at a nearby community college.
-
Jenny considers herself a hardcore video gamer. She's a 25 year old brunette. She doesn't drink but loves singing karaoke.

Heather is a 27 year old blond and likes to write in her spare time. She has season basketball tickets. She enjoys drinking with her friends.

Beth is a brunette who enjoys classical music. She does not like to drink and she is 23 years old. She loves watching scary movies.

Nancy is 31 years old and a wine enthusiast. She has short blond hair and has an extensive DVD collection. She enjoys painting in her spare time.

Figure 18. Bachelor descriptions and sample matches for unbiased groups.

In the subsequent baseline phase, all participants were presented with 30 trials on which they received no feedback, with equal exposure to all trial types. Again, ratings were given on a Likert scale from 1 ("Definitely Frank") to 6 ("Definitely James").

All participants next proceeded to the learning phase and were told that now they will receive feedback after making each match in order to gain a better understanding of their performance. Those in Warning conditions were also made aware that the computer will provide incorrect feedback on some trials to simulate the experience of making the best matches when client preferences are not always consistent. They were told that the incorrect feedback will occur throughout the task, but overall feedback will be consistent enough that they should have no trouble in determining the real qualities a client wants in a match.

During the learning phase participants made a match and received feedback after every trial. For participants in the Warning conditions, after each match they were required to respond to the question "Was the feedback accurate on this trial?" A Likert scale was used to assess feedback accuracy from 1 ("Definitely Incorrect") to 6 ("Definitely Correct"). Those in the No Warning conditions did not make this judgment. All groups completed a total of 60 learning trials and were exposed to equal numbers of trial types.

At the completion of the learning phase, participants moved on to the test phase. The test phase consisted of 30 trials without feedback. Finally, a system knowledge test asked participants to indicate the probability that each bachelor would prefer a match for each given compatibility factor in order to determine changes in bias at the conclusion of the task.

Results

All results are reported as significant at $p < .05$ unless otherwise indicated. For all instances when Mauchly's test indicated that sphericity had been violated, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Manipulation Check. Independent sample t-tests were conducted for accuracy on bias congruent and bias incongruent trials for biased and unbiased groups at baseline in order to test the effectiveness of the biasing and non-biasing manipulations. Accuracy scores were calculated as the proportion of correct responses out of ten trials for each trial type.

Both Biased/No Warning and Biased/Warning groups preformed significantly different on bias congruent and bias incongruent trials, $t(51) = 39.50$, $d = 11.06$, $p < .001$ and $t(51) = 31.53$, $d = 8.83$, $p < .001$ respectively. As expected from the biasing manipulation, performance was near ceiling for both groups on bias congruent trials ($M = .98$, $SD = .07$ for Biased/No Warning and $M = .96$, $SD = .11$ for Biased/Warning). Accuracy on bias incongruent trials was markedly lower for both groups ($M = .05$, $SD = .11$ for Biased/No Warning and $M = .05$, $SD = .13$ for Biased/Warning), indicating that the biasing manipulation was successful.

Unexpectedly, participants in Unbiased/No Warning and Unbiased/Warning groups also demonstrated significant differences in accuracy on bias congruent and bias incongruent trials at baseline, $t(50) = 6.07$, $d = 1.71$, $p < .001$ and $t(47) = 4.66$, $d = 1.36$, $p < .001$ respectively. Means for each group indicated that participants were more accurate on bias congruent trials ($M = .69$, $SD = .24$ for Unbiased/No Warning and $M = .67$, $SD = .28$ for Unbiased/Warning) than on bias incongruent trials ($M = .35$, $SD = .25$ for Unbiased/No Warning and $M = .32$, $SD = .28$ for Unbiased/Warning).

To determine if bias strength was similar between biased and unbiased groups at baseline, a difference score was created by subtracting baseline accuracy on incongruent trials from accuracy on bias congruent trials. A difference score of 1 would indicate that matches are being made perfectly in line with the manipulated bias, a score of 0 would indicate no effect of bias, and a score of -1 would indicate that matches are being made perfectly against the bias. Correcting for variance, this test revealed that Biased groups were significantly more biased ($M = .92$, $SD = .19$) than Unbiased groups ($M = .34$, $SD = .45$), $t(129.5) = 11.69$, $d = 2.05$, $p < .001$.

Thus, although some bias appeared to be present in the Unbiased group, the strength of that bias is significantly less than the bias that exists in the Biased group.

Accuracy Performance. In order examine how bias and warning affected performance throughout the task, separate 3 (phase: baseline, learning, test) x 4 (group: Biased/No warning, Unbiased/No Warning, Biased/Warning, Unbiased/Warning) mixed-model ANOVAs were conducted on the accuracy of bias incongruent, bias congruent, and bias irrelevant

First, a 3 x 4 mixed-model ANOVA examined performance on bias incongruent trials throughout the task phase for all groups (see Figure 19). There was a significant main effect of phase on incongruent trial accuracy $F(1.81, 359.42) = 187.6$, $MSE = .05$, $\eta_p^2 = .49$. Bonferroni corrected pairwise comparisons that accuracy was significantly lower at baseline ($M = .19$, $SE = .01$) than at learning ($M = .38$, $SE = .01$) or test ($M = .60$, $SE = .01$), and that learning accuracy was significantly less than test accuracy.

There was also a main effect of group on accuracy performance for bias incongruent trials, $F(3, 199) = 28.52$, $MSE = .02$, $\eta_p^2 = .3$. Bonferroni corrected pairwise comparisons indicated that there were no significant differences in accuracy performance on bias incongruent trial between Biased/No Warning ($M = .33$, $SE = .02$) and Biased/Warning ($M = .29$, $SE = .02$) groups, nor did differences in accuracy exist between Unbiased/No Warning ($M = .48$, $SE = .02$) and Unbiased/Warning ($M = .46$, $SE = .02$) groups. Both Biased and Unbiased groups were significantly different from each other, however.

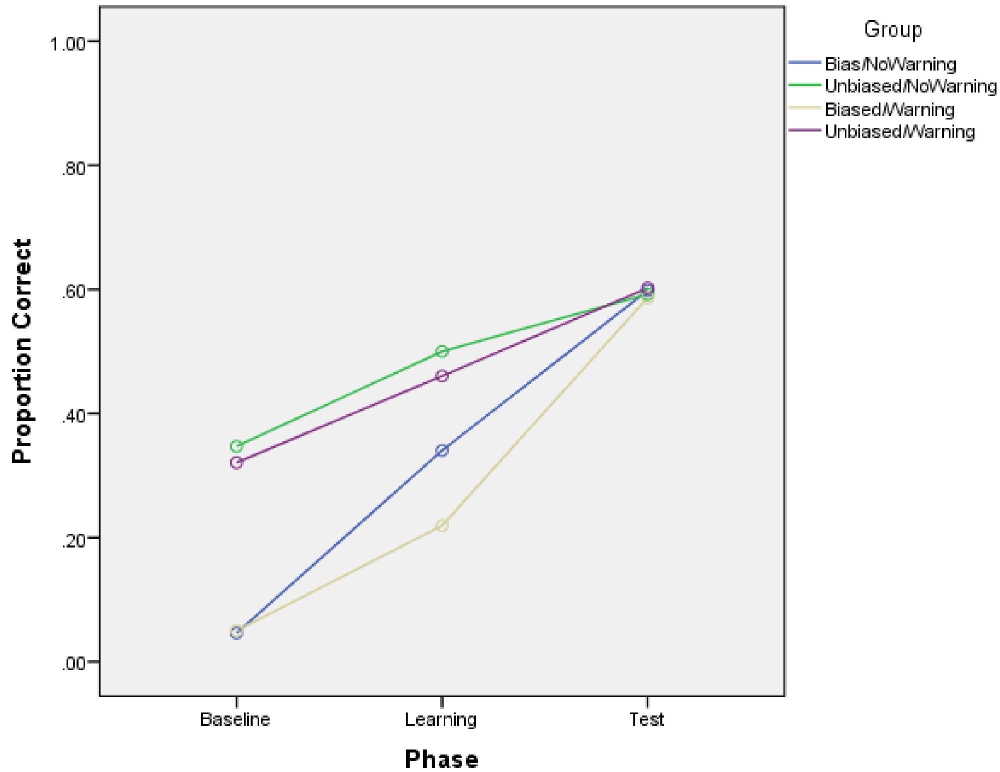


Figure 19. Accuracy performance on incongruent trials for Experiment 2.

A significant interaction was also found to exist between Phase and Group, $F(5.42, 359.42) = 9.03$, $MSE = .05$, $\eta_p^2 = .12$, indicating accuracy on bias incongruent trials at each phase of the task differed between groups. Independent samples t-tests were conducted collapsing across warning condition for biased and unbiased groups to examine how these groups differed on incongruent trial type accuracy throughout the task. These tests revealed that, at baseline, accuracy on these trial types was significantly worse for Biased groups ($M = .05$, $SD = .12$) than Unbiased groups ($M = .33$, $SD = .26$), $t(135.66) = -9.98$, $d = -1.71$. Levene's test indicated unequal variances ($F = 85.14$, $p < .001$), so degrees of freedom were adjusted from 201 to 135.66. Biased groups continued to perform worse during the learning phase ($M = .28$, $SD = .02$) than Unbiased groups ($M = .48$, $SD = .18$), $t(196.96) = -7.11$, $d = 1.0$. Levene's test again indicated unequal variances ($F = 6.93$, $p < .05$), and degrees of freedom was adjusted from 201 to 196.96. However, Biased and Unbiased groups were showing equal accuracy performance at test, $t(201) = -.14$, $p > .05$, ($M = .59$, $SD = .24$ and $M = .60$, $SD = .24$ respectively).

A second 3 x 4 mixed-model ANOVA compared accuracy on bias congruent trials at baseline, learning, and test for each of the four groups (see Figure 20).

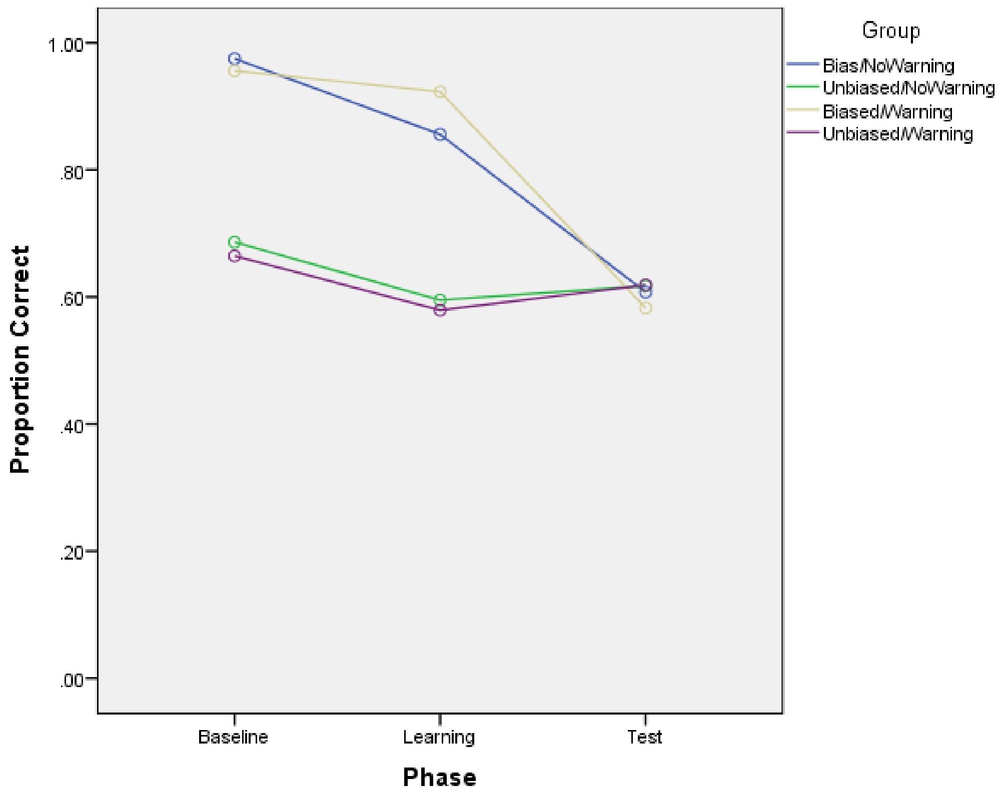


Figure 20. Accuracy performance on bias congruent trials for Experiment 2.

There was a significant main effect of phase on congruent trial accuracy $F(1.71, 340.64) = 69.13$, $MSE = .04$, $\eta_p^2 = .26$. Bonferroni corrected pairwise comparisons indicated that baseline accuracy ($M = .82$, $SE = .01$) was significantly better than learning accuracy ($M = .74$, $SE = .01$) and test accuracy ($M = .61$, $SE = .02$), and that learning accuracy was significantly better than test accuracy.

A main effect of group on accuracy performance for bias congruent trials was found to exist, $F(3, 199) = 37.63$, $MSE = .02$, $\eta_p^2 = .36$. Bonferroni corrected pairwise comparisons indicated that there were no significant differences in accuracy performance on bias incongruent trial between Biased/No Warning ($M = .81$, $SE = .02$) and Biased/Warning ($M = .82$, $SE = .02$) groups, nor did differences in accuracy exist between Unbiased/No Warning ($M =.63$, $SE = .02$) and Unbiased/Warning ($M = .62$, $SE = .02$) groups. Biased and Unbiased groups were found to be significantly different from each other, however.

A significant interaction was also found to exist between Phase and Group, $F(5.14, 340.64) = 17.54$, $MSE = .04$, $\eta_p^2 = .21$, indicating accuracy on bias congruent trials at each phase of the task differed between groups. Independent samples t-tests were conducted collapsing across warning condition for biased and unbiased groups to examine how these groups differed on congruent trial type accuracy throughout the task. These tests revealed that the groups differed significantly at baseline, with Biased groups ($M = .97$, $SD = .09$) outperforming Unbiased groups ($M = .68$, $SD = .26$), $t(122.09) = 10.54$, $d = 1.91$. Levene's test indicated unequal variances ($F = 95.63$, $p < .01$) and degrees of freedom were adjusted from 201 to 122.09. Biased groups were more accurate on bias congruent trials during learning as well ($M = .89$, $SD = .12$), outperforming the Unbiased groups ($M = .59$, $SD = .18$), $t(185.59) = 13.52$, $d = 1.98$. Levene's test again indicated unequal variances ($F = 9.17$, $p < .05$) and so degrees of freedom were adjusted from 201 to 185.59. Biased and Unbiased groups were comparable in accuracy by the time of test. Biased and Unbiased groups were comparable in accuracy on bias congruent trials by the time of test, $t(201) = -.70$, $p > .05$, ($M = .60$, $SD = .24$ and $M = .62$, $SD = .23$ respectively).

A 3x4 mixed-model ANOVA comparing accuracy on bias irrelevant trials at baseline, learning, and test for each of the four groups was also conducted (see Figure 21).

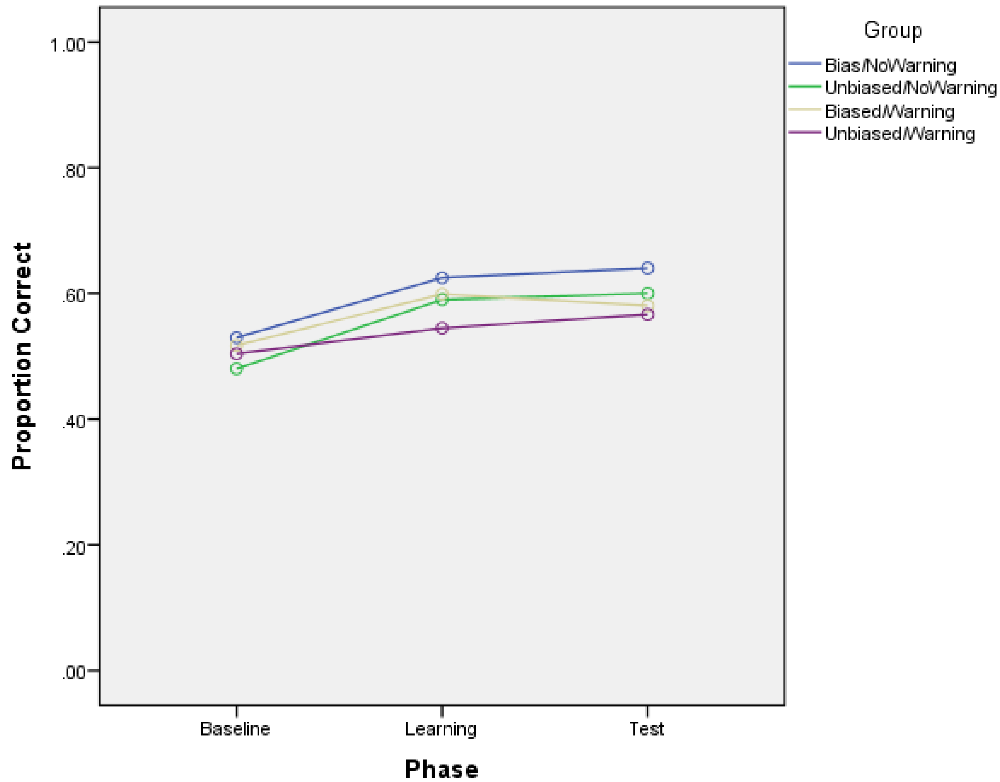


Figure 21. Accuracy performance on bias irrelevant trials for Experiment 2.

A main effect of phase was found for accuracy on bias irrelevant trials, $F(1.82, 362.74) = 13.64$, $MSE = .04$, $\eta_p^2 = .06$. Bonferroni corrected pairwise comparisons indicated that baseline accuracy ($M = .51$, $SE = .01$) was significantly lower than learning ($M = .59$, $SE = .01$) or test ($M = .60$, $SE = .02$) accuracy, but that no significant differences were found between learning and test. There was no main effect of group, $F(1, 199) = 2.21$, and no significant interaction between phase and group, $F(6, 398) = .48$.

Confidence. In order to determine how confidence changed throughout the task, individual 3 (phase: baseline, learning, test) x 4 (group: Biased/No Warning, Unbiased/No

Warning, Biased/Warning, Unbiased/Warning) mixed-model ANOVAs were carried out on confidence ratings for each of the three trial types.

A first 3 x 4 mixed-model ANOVA was conducted to examine confidence on bias incongruent trials for all groups throughout each phase of the task (see Figure 22). A main effect of phase was found indicating that confidence differed for all groups across each phase of the task, $F(1.93, 383.04) = 68.67$, $MSE = .12$, $\eta_p^2 = .26$. Pairwise comparisons with Bonferroni correction revealed that across all groups, ratings of confidence were higher at baseline (M = 2.01, SE = .03) than at learning (M = 1.73, SE = 1.73) or test (M = 1.78, SE = .02). Confidence did not change on bias incongruent trials from learning to test.

The main effect of group was also found to be significant, $F(1, 199) = 7.72$, $MSE = .25$, $\eta_p^2 = .29$. Bonferroni corrected pairwise comparisons revealed that the Biased/Warning group provided significantly different confidence ratings (M = 2.00, SE = .04) than both the Unbiased/Warning group (M = 1.78, SE = .04) and the Unbiased/No Warning group (M = 1.81, .04). There was a significant interaction between phase and group, $F(5.77, 383.76) = 3.86$, $MSE = .12$, $\eta_p^2 = .06$, suggesting that the groups expressed different levels of confidence on bias incongruent items at the different phases of the task. Independent repeated measures ANOVAs were conducted for examine this interaction for each group.

The main effect of phase for the Biased/No Warning group was significant, $F(2, 102) = 27.37$, $MSE = .10$, $\eta_p^2 = .35$, with Bonferroni corrected pairwise comparisons indicating that confidence on bias incongruent trials was greater at baseline (M = 2.16, SE = .06) than at learning (M = 1.73, SE = .06) or test (M = 1.80, SE = .04), but that learning and test were not significantly different from one another.

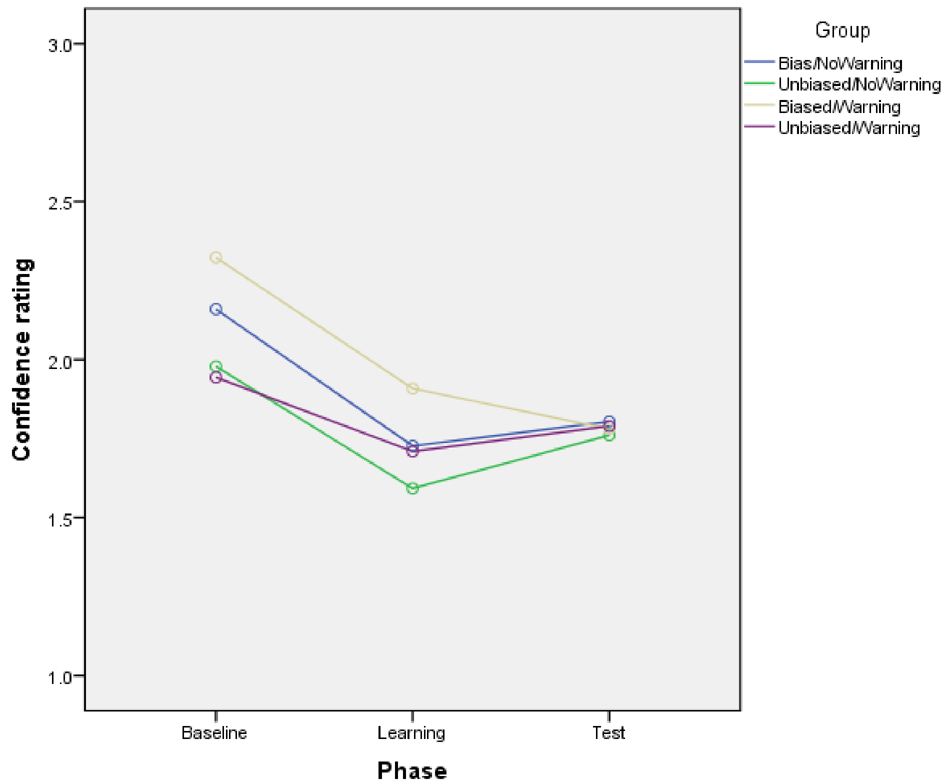


Figure 22. Confidence on bias incongruent trials for Experiment 2.

The main effect of phase was also significant for the Unbiased/No Warning group, $F(2, 100) = 18.10$, $MSE = .11$, $\eta_p^2 = .27$, with Bonferroni corrected pairwise comparisons indicating that baseline confidence ($M = 1.98$, $SE = .05$) was significantly higher than confidence at learning ($M = 1.60$, $SE = .06$) and learning confidence was significantly lower than confidence at test ($M = 1.77$, $SE = .05$). Confidence did not significantly differ between baseline and test for this group. The main effect of phase was significant for the Biased/Warning group, $F(2, 102) = 27.17$, $MSE = .15$, $\eta_p^2 = .35$, with Bonferonni corrected pairwise comparisons indicating that baseline confidence ($M = 2.32$, $SE = .06$) was significantly higher than learning ($M = 1.91$, $SE = .07$) and test confidence ($M = 1.78$, $SE = .05$), but that learning and test did not significantly differ from one another. The main effect of phase was significant for the Unbiased/Warning group as well, $F(2, 94) = 6.39$, $MSE = .11$, $\eta_p^2 = .12$, with Bonferroni corrected pairwise comparisons indicating that baseline confidence (*M* = 1.94, *SE* = .05) was significantly higher than learning confidence (*M* = 1.71, *SE* = .05), but that no significant difference existed in confidence between learning and test (*M* = 1.79, *SE* = .04), or baseline and test.

A second 3 x 4 mixed-model ANOVA examined confidence on bias congruent trials for all groups at each phase of the task (see Figure 23). A main effect of phase was found, $F(1.90, 377.09) = 52.29$. $MSE = .11$, $\eta_p^2 = .21$. Pairwise comparisons with Bonferroni correction revealed that baseline confidence on congruent trials (*M* = 2.10, *SE* = .03) was significantly different than learning (*M* = 1.88, SE = .03) and test (*M* = 1.78. *SE* = .02), and that confidence at learning and test were different from one another. The main effect of group was found to be significant, $F(1, 199) = 20.08$. $MSE = .06$, $\eta_p^2 = .23$. Bonferroni corrected pairwise comparisons revealed that the Biased/No Warning (*M* = 2.00, *SE* = .03) and Biased/Warning (*M* = 2.03, *SE* = .03) groups did not differ from each other, nor did the Unbiased/No Warning (*M* = 1.77, *SE*= .03) or Unbiased/Warned (*M* = 1.82, *SE* = .04) groups differ from each other. The Unbiased/No Warning group was significantly less confident than both biased groups and the Biased/Warning group was significantly more confident than both of the Unbiased groups.

The interaction between phase and group was also found to be significant $F(5.69, 377.09) = 8.19$, $MSE = .11$, $\eta_p^2 = .11$. Independent samples t-tests were conducted, collapsing across biasing condition. These tests revealed that confidence was significantly different at baseline, $t(201) = 7.45$, $d = 1.05$, with Biased groups displaying more confidence than Unbiased groups (*M* = 2.29, *SD* = .40 and *M* = 1.91, *SD* = .33 respectively). Biased groups continued to be more confident during learning (*M* = 2.05, *SD* = .44) than Unbiased groups, (*M* = 1.70, *SD* = .41), $t(201) = 5.98$, $d = .08$. There were no significant difference between the Biased and Unbiased groups at test, $t(201) = .08$.

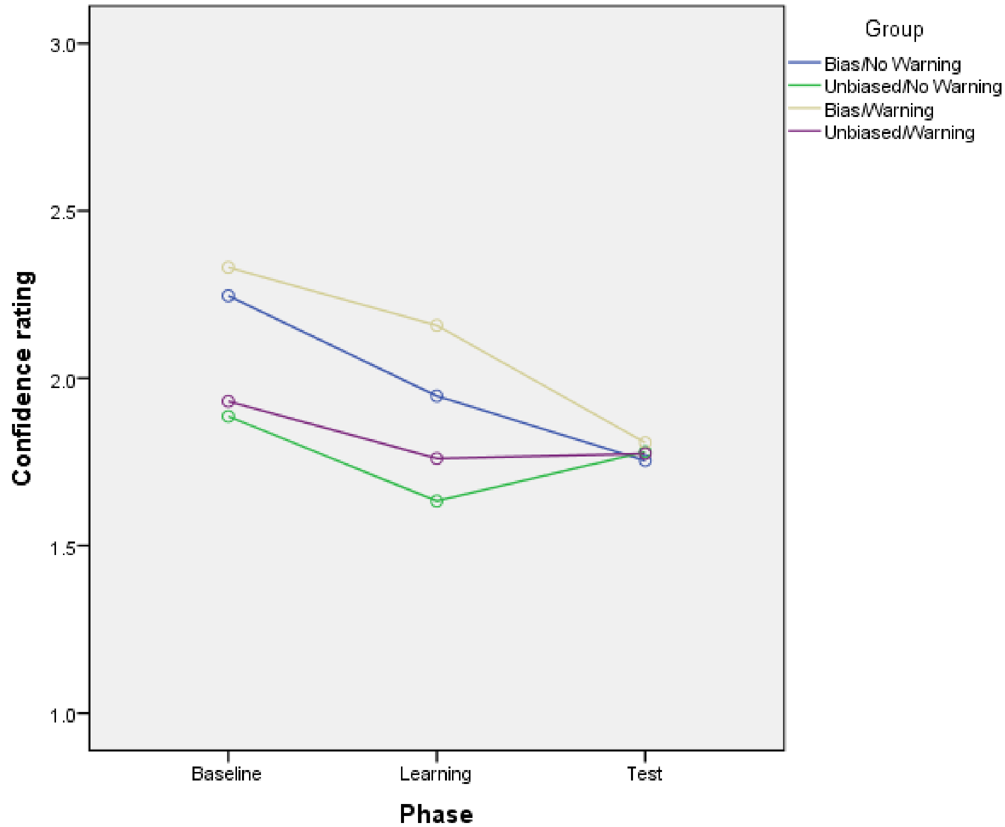


Figure 23. Confidence on bias congruent trials for Experiment 2.

A third 3 x 4 mixed-model ANOVA examined confidence on bias irrelevant trials for all groups at each phase of the task (see Figure 24). A main effect of phase was found, $F(1.79, 356.51) = 22.30$, $MSE = .11$, $\eta_p^2 = .10$. Bonferroni corrected pairwise comparisons revealed that confidence on bias irrelevant trials was higher at test ($M = 1.77$, $SE = .02$) than at baseline ($M = 1.60$, $SE = .02$) or learning ($M = 1.58$, $SE = .03$), but that baseline and learning confidence were not significantly different from one another.

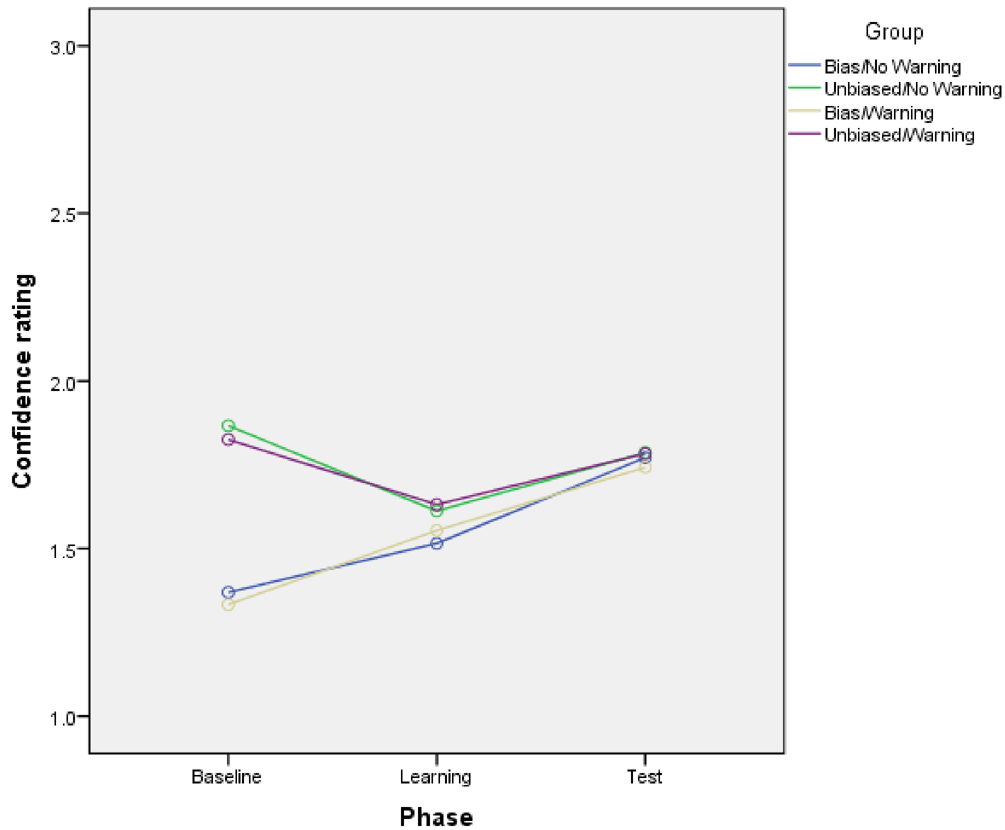


Figure 24. Confidence on bias irrelevant trials for Experiment 2.

The main effect of group was found to be significant, $F(1, 199) = 14.25$, $MSE = .05$, $\eta_p^2 = .18$. Pairwise comparisons with Bonferroni correction revealed that confidence was not significantly different for the Biased/No Warning ($M = 1.55$, $SE = .03$) and the Biased/Warning ($M = 1.54$, $SE = .03$) groups, nor for the Unbiased/No Warning ($M = 1.76$, $SE = .03$) and the Unbiased/Warning ($M = 1.75$, $SE = .03$) groups. The biased and unbiased groups were significantly different from each other, however. The interaction between phase and group was also found to be significant, $F(5.36, 356.51) = 11.00$, $MSE = .11$, $\eta_p^2 = .14$. Independent samples t-tests were conducted to explore this interaction, collapsing across biased and unbiased groups. The biased group was found to be significantly less confident on bias irrelevant trials at baseline

(*M* = 1.35, *SD* = .31) compared to the unbiased groups (*M* = 1.85, *SD* = .36), *t*(201) = -10.54, *d* = -1.49. These differences were not found to exist at learning, *t*(201) = -1.55, or test, *t*(201) = -.72.

Response Time. In order to determine how response time varied throughout the task, individual 3 (phase: baseline, learning, test) x 4 (group: Biased/No Warning, Unbiased/No Warning, Biased/Warning, Unbiased/Warning) mixed-model ANOVAs were carried out on response time for each of the three trial types.

A first 3 x 4 mixed-model ANOVA examined response times on bias incongruent trials for all groups at each phase of the task (see Figure 25).

A main effect of phase was found, *F*(1.61, 319.48) = 74.31, *MSE* = 19663.57, η_p^2 = .27, indicating that response times differed during the task phases. Bonferroni corrected pairwise comparisons indicated that baseline response times (*M* = 5.63, *SE* = .15) were significantly different than both learning (*M* = 4.50, *SE* = .12) and test (*M* = 4.19, *SE* = .13) response times, and that learning and test response times were significantly different from each other.

The main effect of group was not found to be significant, *F*(1, 199) = .97, *MSE* = 19663.57, η_p^2 = .06. The interaction between phase and group was significant, *F*(4.85, 319.48) = 4.55, suggesting that the groups differed in their response times at different phases of the task. This interaction was explored with four repeated measures ANOVAs examining response time at each phase of the task for each group independently.

An ANOVA for the Biased/No Warning group found the main effect of phase to be significant, *F*(2, 102) = 7.59, *MSE* = 1477.38, η_p^2 = .13. Bonferroni corrected pairwise comparisons revealed that response time was significantly slower at baseline (*M* = 5.57, *SE* = .30) than at learning (*M* = 4.67, *SE* = .24), but that no significant differences in response time existed between learning and test (*M* = 4.92, *SE* = .29) or baseline and test.

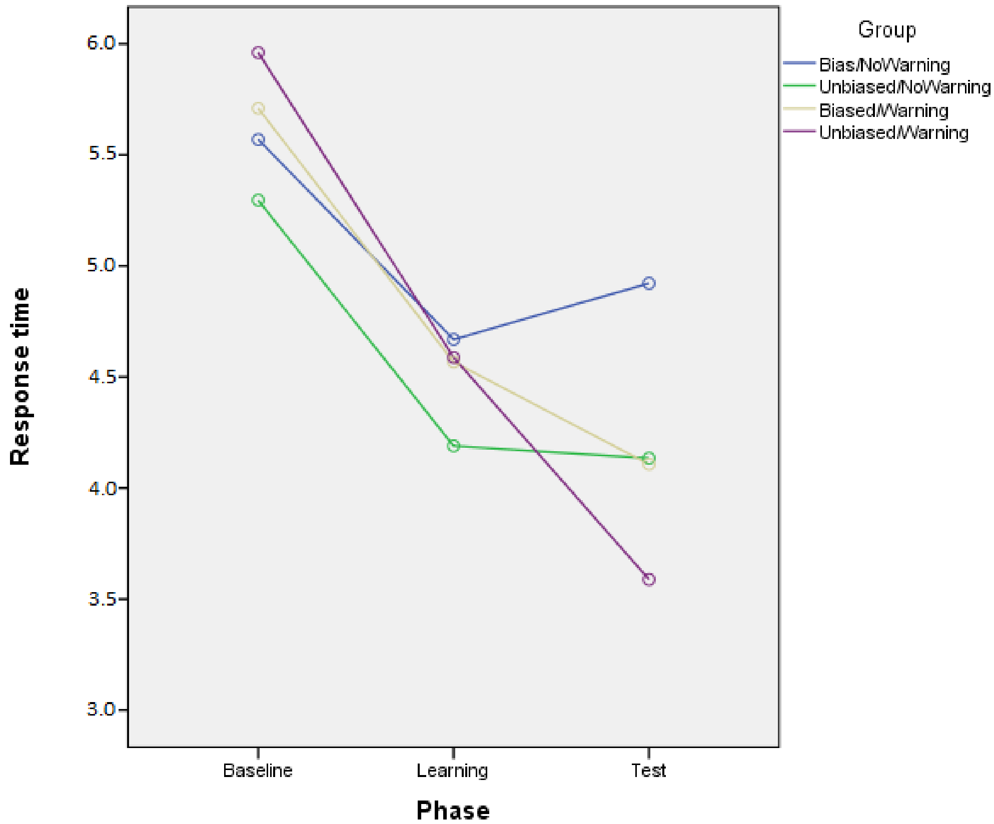


Figure 25. Response time in seconds on bias incongruent trials for Experiment 2.

An ANOVA for the Biased/No Warning group found the main effect of phase to be significant, $F(2, 102) = 7.59$, $MSE = 1477.38$, $\eta_p^2 = .13$. Bonferroni corrected pairwise comparisons revealed that response time was significantly slower at baseline ($M = 5.57$, $SE = .30$) than at learning ($M = 4.67$, $SE = .24$), but that no significant differences in response time existed between learning and test ($M = 4.92$, $SE = .29$) or baseline and test. An ANOVA for the Unbiased/No Warning groups was performed. Mauchly's test indicated that the assumption of sphericity had been violated from the main effect of phase, $\chi^2(2) = .75$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .80$). The main effect of phase was found to be significant for the Unbiased/No Warning group, $F(1.60, 79.81) = 15.78$, $MSE = 1740.11$, $\eta_p^2 = .24$. Bonferroni corrected pairwise comparisons revealed that baseline response time (*M* = 5.30, *SE* = .25) was significantly slower than at learning (*M* = 4.19, *SE* = .26) or test (*M* = 4.13, *SE* = .26), but that there were no significant differences between learning and test. An ANOVA was performed for the Biased/Warning group. Mauchly's test indicated that the assumption of sphericity had been violated from the main effect of phase, $\chi^2(2)$ = .55. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .69). A main effects of phase was found, *F*(1.38, 70.36) = 20.43, *MSE* = 2514.45, η_p^2 = .29. Bonferroni corrected pairwise comparisons indicated that baseline response time (*M* = 5.371, *SE* = .35) was significantly slower than at learning (*M* = 4.57, *SE* = .25) or test (*M* = 4.11, *SE* = .29), but that there were no significant differences between learning and test. Lastly, an ANOVA was performed for the Unbiased/Warned group. Mauchly's test indicated that the assumption of sphericity had been violated from the main effect of phase, $\chi^2(2)$ = .74. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .80). A main effect of phase was found, *F*(1.59, 74.83) = 39.60, *MSE* = 2161.46, η_p^2 = 46. Bonferroni pairwise comparisons revealed that response times differed significantly from one another at each phase of the task, decreasing from baseline (*M* = 5.96, *SE* = .31) to learning (*M* = 4.59, *SE* = .23), and again from learning to test (*M* = 3.59, *SE* = .20).

A second 3 x 4 mixed-model ANOVA examined response times on bias congruent trials for all groups at each phase of the task (see Figure 26). A main effect of phase was found, *F*(1.67, 332) = 104.81. *MSE* = 229141.47, η_p^2 = .35, indicating that the groups were exhibiting different response times on bias congruent trials at different phases of the task. Bonferroni corrected pairwise comparisons revealed that groups were significantly slower at baseline (*M* = 5.89, *SE* = .15) than at learning (*M* = 4.44, *SE* = .12) and test (*M* = 4.05, *SE* = .12).

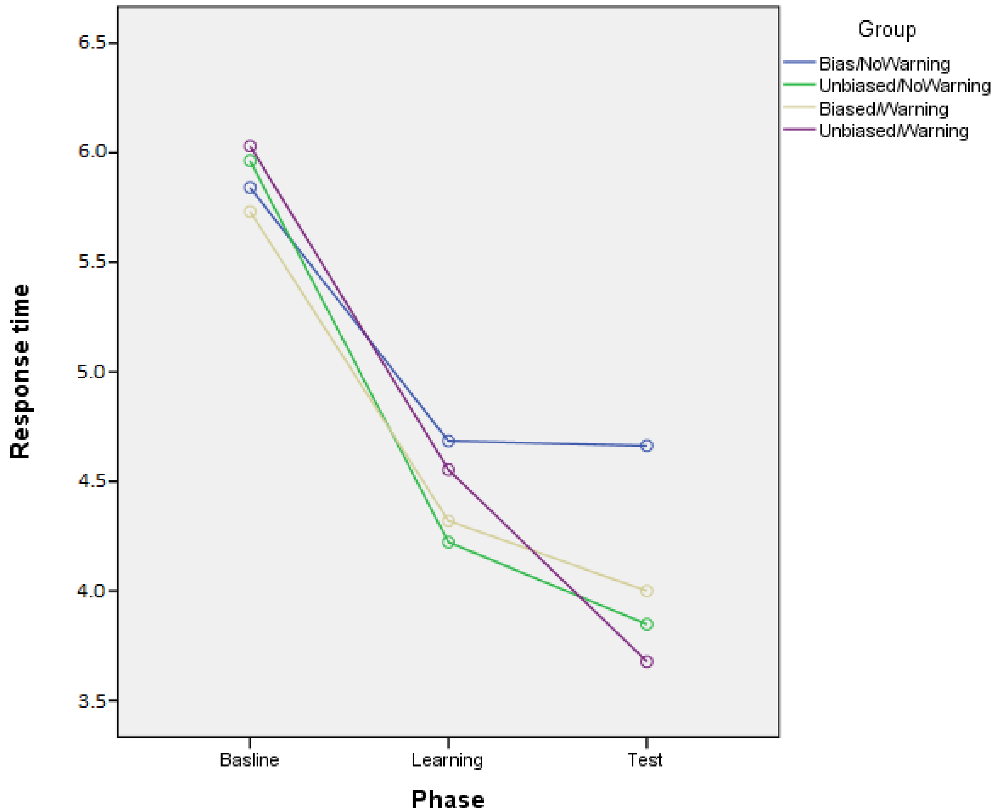


Figure 26. Response time in seconds on bias congruent trials for Experiment 2.

Response times were also found to be significantly slower at learning than at test. There was no main effect of group, $F(1, 199) = .77$, and no significant interaction between phase and group, $F(1.67, 322) = 2.05$.

A third 3 x 4 mixed-model ANOVA examined response times on bias irrelevant trials for all groups at each phase of the task (see Figure 27). A main effect of phase was found, $F(1.78, 353.16) = 96.18$, $MSE = 240100.49$, $\eta_p^2 = .33$, indicating that the groups were exhibiting different response times for bias irrelevant trials at different phases of the task. Bonferroni corrected pairwise comparisons revealed that baseline response times ($M = 6.28$, $SE = .17$) were significantly slower than learning ($M = 4.83$, $SE = .14$) and test ($M = 4.30$, $SE = .14$) response times, and that learning and test response times were significantly different from each other.

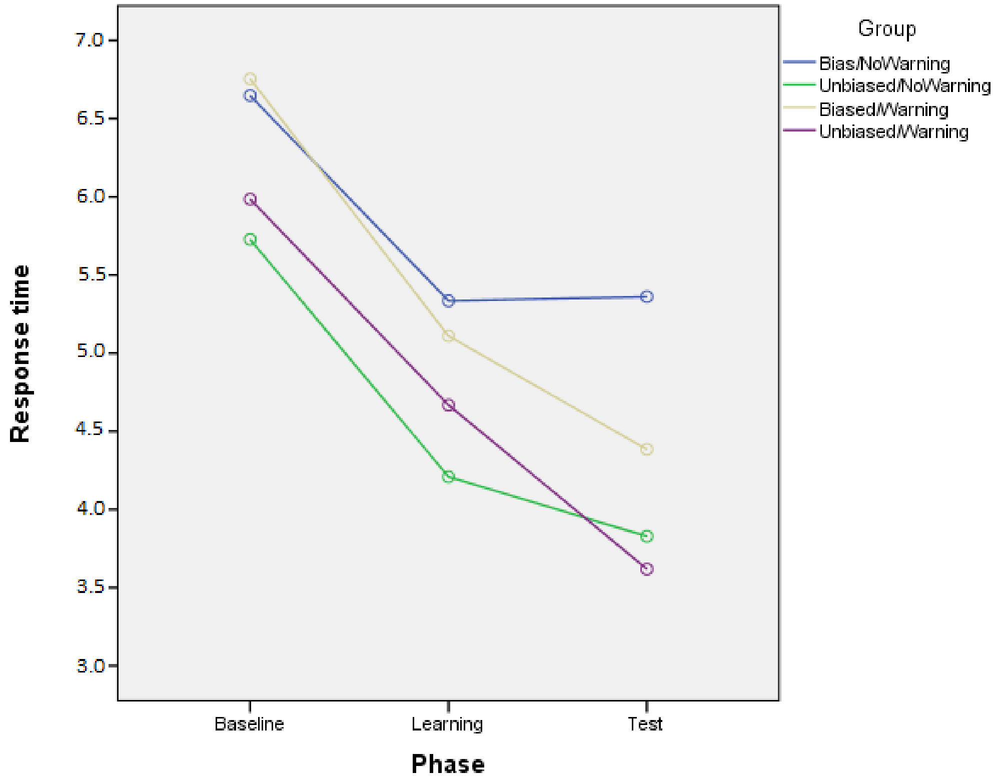


Figure 27. Response time in seconds on bias irrelevant trials for Experiment 2.

A main effect of group was also found, $F(1, 199) = 5.17$, $MSE = 3087.21$, $\eta_p^2 = .07$, indicating the groups were exhibiting different response times over all phases of the task. Bonferroni corrected pairwise comparisons revealed no differences between Bias/No Warning ($M = 5.78$, $SE = .24$) and Biased/Warning ($M = 5.42$, $SE = .24$) groups and no differences between Unbiased/No Warning ($M = 4.59$, $SE = .25$) and Unbiased/Warning ($M = 4.76$, $SE = .25$) groups. However, both Biased groups were found to be significantly slower in their response times than both of the Unbiased groups. The interaction between phase and group was not significant, $F(1.78, 353.16) = 1.92$.

Bias Strength. In order to determine how bias strength varied throughout the task, a 3 (phase: baseline, learning, test) x 4 (group: Biased/No Warning, Unbiased/No Warning, Bias/Warning, Unbiased/Warning) mixed-model ANOVA was conducted on the computed difference score (see Figure 28).

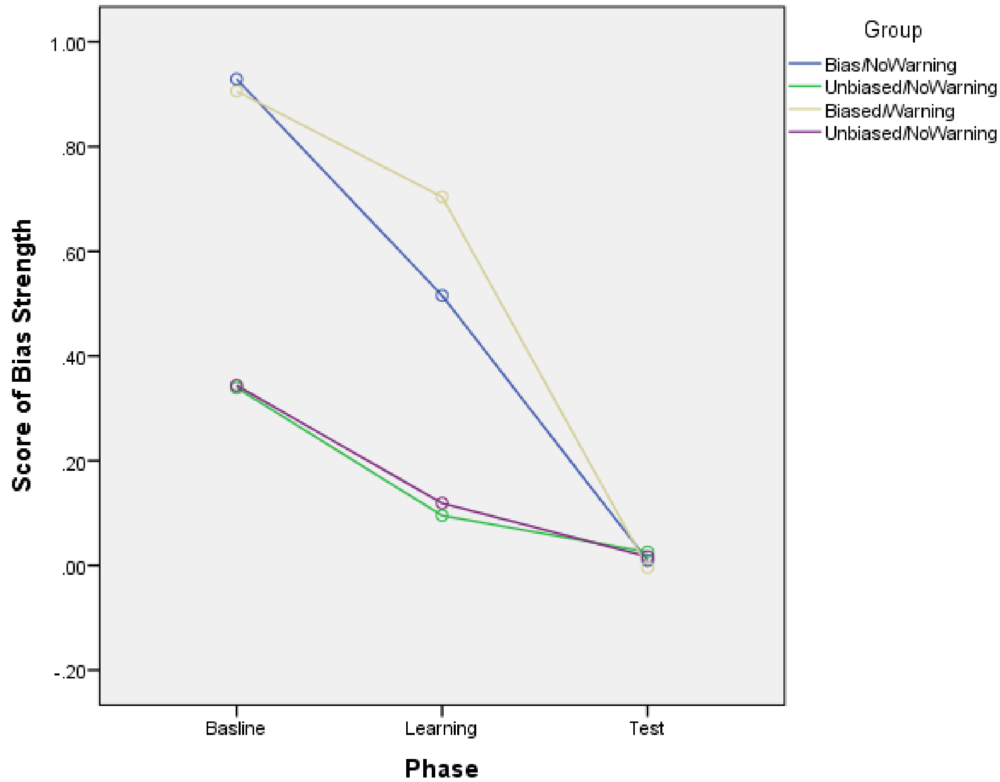


Figure 28. Bias strength score throughout task for Experiment 2.

A significant main effect of phase was found on bias strength, $F(1.93, 383.83) = 280.16$, $MSE = .07$, $\eta_p^2 = .59$. Bonferroni corrected pairwise comparisons indicated that bias strength at baseline ($M = .63$, $SE = .02$) was significantly greater than at learning ($M = .36$, $SE = .02$) and test ($M = .01$, $SE = .02$), and that bias strength at learning was significantly greater than at test.

A significant main effect of group was also found for bias strength, $F(3, 199) = 58.66$, $MSE = .04$, $\eta_p^2 = .47$. Pairwise comparisons were adjusted for multiple comparisons using the Bonferroni correction. These comparisons indicated that there were no significant differences in accuracy performance on bias incongruent trial between Biased/No Warning ($M = .49$, $SE = .03$) and Biased/Warning ($M = .54$, $SE = .03$) groups, nor did differences in accuracy exist between

Unbiased/No Warning (*M* = .15, *SE* = .03) and Unbiased/Warning (*M* = .16, *SE* = .03) groups. Bias strength was significantly different between both Biased and Unbiased groups, however.

The interaction between phase and group was also found to be significant, $F(5.79, 383.83) = 27.44$, $MSE = .07$, $\eta_p^2 = .29$, indicating that bias strength at each phase of the task differed between groups. Independent samples t-tests further explored this interaction by collapsing across biased and unbiased groups to examine changes in bias strength throughout the task. These tests revealed that, at baseline, Biased groups demonstrated a significantly stronger bias (*M* = .92, *SD* = .19) than Unbiased groups (*M* = .34, *SD* = .45), $t(129.50) = 11.69$, $d = 2.05$. Levene's test indicated unequal variances and so degrees of freedom were adjusted from 201 to 129.50. Bias strength decreased for both Biased (*M* = .61, *SD* = .30) and Unbiased groups (*M* = .11, *SD* = .27) during the learning phase, with the Biased group continuing to exhibit significantly stronger bias, $t(201) = 12.65$, $d = 1.78$. Both groups showed equal bias at test, $t(201) = -.58$, $d = -.08$, with Biased and Unbiased group demonstrating virtually no bias strength (*M* = .00, *SD* = .22 and *M* = .02, *SD* = .23 respectively).

Explicit Measure of Bias at Test. In order to compare participants' implicit bias as indicated by the above performance measures on the computerized matching task to their explicit indication of bias, analyses were performed on ratings of compatibility factor importance provided by participants in writing at baseline and test. Composite scores of compatibility factor importance were calculated for each compatibility factor by averaging the importance ratings for both bachelors. Of particular interest are the compatibility factors of Entertainment Preference (biased factor) and Hair Color (critical factor).

A 2 (phase: baseline, test) x 4 (group: Biased/No Warning, Unbiased/No Warning, Bias/Warning, Unbiased/Warning) mixed-model ANOVA was conducted for importance ratings of Entertainment Preference at baseline and test for all groups (see Figure 29).

There was a significant main effect of phase on ratings of Entertainment Preference, $F(1, 184) = 15.78$, $MSE = .80$, $\eta_p^2 = .08$. Pairwise comparisons with Bonferroni correction revealed that this factor was rated as more important at baseline ($M = 3.93$, $SE = .07$) than at test ($M = 3.56$, $SE = .08$). A significant main effect was found for group, $F(3, 184) = 32.12$, $MSE = 1.60$, $\eta_p^2 = .08$, indicating groups were providing different ratings of Entertainment Preference importance across both task phases. Bonferroni corrected pairwise comparisons indicated that there were no significant differences in ratings of Entertainment Preference importance between Biased/No Warning ($M = 4.27$, $SE = .11$) and Biased/Warning ($M = 4.35$, $SE = .12$) groups, nor did differences in accuracy exist between Unbiased/No Warning ($M = 3.18$, $SE = .11$) and Unbiased/Warning groups ($M = 3.20$, $SE = .12$). Importance ratings were significantly different between the Biased and Unbiased groups.

The interaction between phase and group was also found to be significant, $F(1, 184) = 8.44$, $MSE = .80$, $\eta_p^2 = .12$, indicating that the importance of Entertainment Preference changes at each phase and is different across groups. Post-hoc independent samples t-tests for Biased and Unbiased groups revealed that importance ratings for the biased factor of Entertainment Preference was significantly different for biased groups at baseline and test, $t(93) = 6.91$, $d = 1.43$, with bias decreasing from baseline ($M = 4.7$, $SD = .59$) to test ($M = 3.9$, $SD = 1.03$).

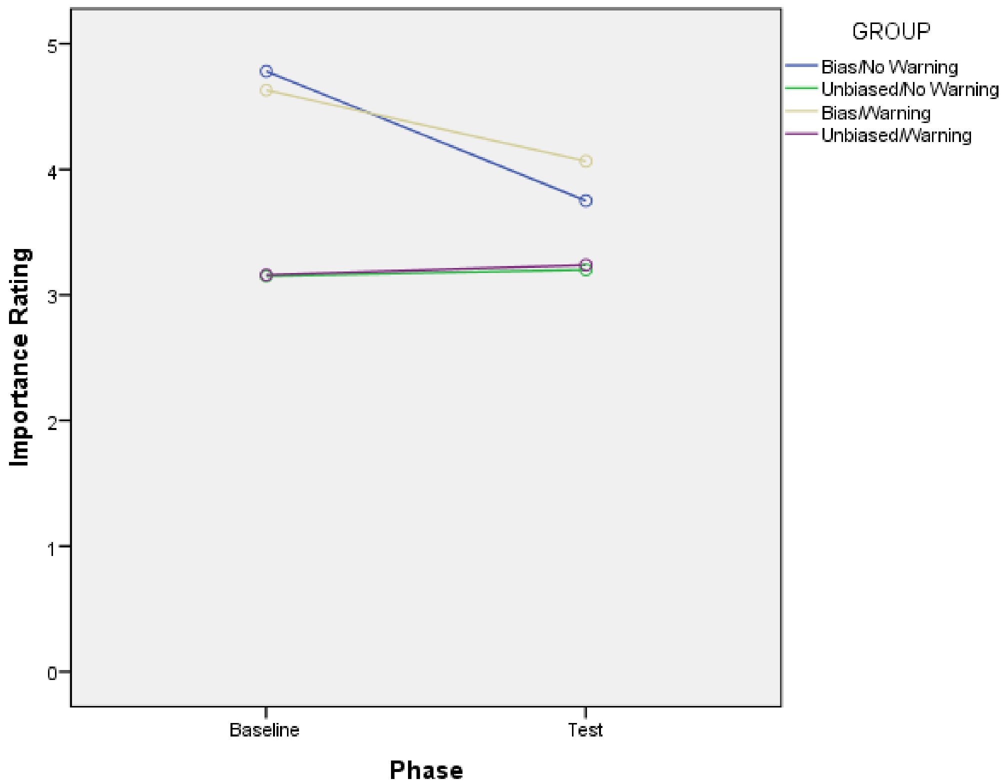


Figure 29. Importance rating of entertainment preference for Experiment 2.

Because Entertainment Preference was the biased compatibility factor, these results suggest that after completing the learning phase, Biased participants became significantly less biased. Importance ratings of Entertainment Preference were not significantly different between the two time points for Unbiased groups, $t(93) = -.443$, with ratings of Entertainment Preference importance comparable at baseline ($M = 3.15$, $SD = 1.10$) and test ($M = 3.22$, $SD = 1.20$). This suggests that Unbiased participants view this compatibility factor as only moderately important throughout the task. Overall, these results indicate that while implicit measures of bias strength as derived from the computerized matching task suggest that bias is near eliminated at the time of test, all groups still rate the biased compatibility factor of Entertainment Preference as at least moderately important when explicitly asked.

A 2 (phase: baseline, test) x 4 (group: Biased/No Warning, Unbiased/No Warning, Biased/Warning, Unbiased/Warning) mixed-model ANOVA was also conducted for ratings of Hair Color importance at baseline and test for all groups (see Figure 30).

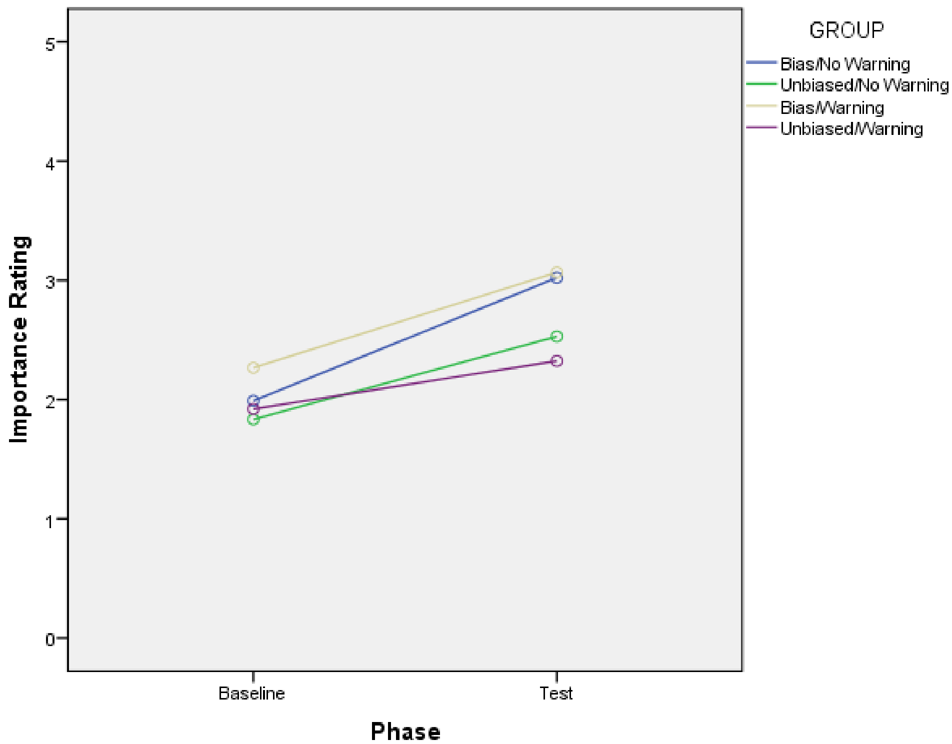


Figure 30. Importance rating of hair color for Experiment 2.

A main effect of phase was found, $F(1, 183) = 45.23$, $MSE = 1.11$, $\eta_p^2 = .20$. Bonferroni corrected pairwise comparisons revealed that Hair Color was rated as significantly less important across all groups at baseline ($M = 2.0$, $SE = .09$) than at test ($M = 2.74$, $SE = .11$). There was no main effect of group, $F(1, 183) = 1.40$, and no interaction between phase and group, $F(3, 182) = 1.40$. These results suggest that after completing the learning phase, participants in all groups had become more aware of the importance of the critical compatibility factor of Hair Color, although they are still rating this factor as slightly less than moderately important and less important than the biased compatibility factor of Entertainment Preference.

Although participants rated the biased compatibility factor lower at test than at baseline and believed that the critical factor was more important at test than at baseline, descriptive measures show that the biased factor was still seen as the most important factor in making a good match at test (see Figure 31). The number of participants that endorsed the biased factor important (rating it 4 or 5 on written tests after task completion) was greater for Entertainment Preference than any other compatibility factor.

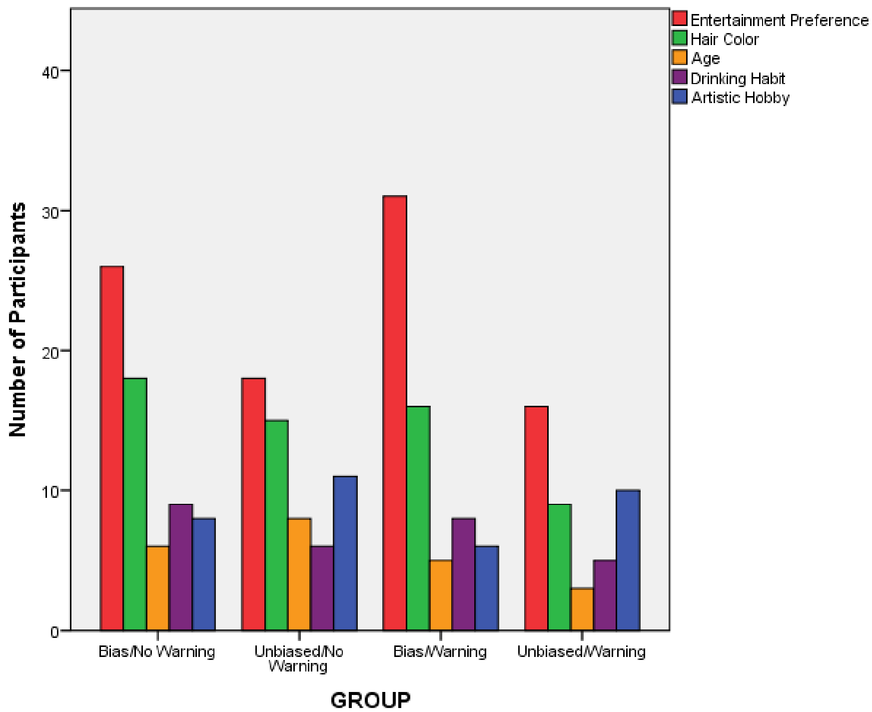


Figure 31. Compatibility factors endorsed as important at test.

DISCUSSION

Warning participants of error in feedback did not enable the participants in those groups to achieve better accuracy than their unwarned counterparts. Rather, the accuracy performance of those in the warned groups mimicked the performance on their unwarned counterparts on both bias congruent and bias incongruent trials. Because warning in this experiment had no effect on decision making, these data cannot speak to the theories posited by Williams, et al. (2013) or Bodenhausen (1988).

While warning did not impact rule learning in this experiment, clear differences were found to exist between the biased and unbiased groups, particularly during the baseline and learning phases, with biased groups tending to achieve higher accuracy on bias congruent trials and lower accuracy on bias incongruent trials at these times. Interestingly, all groups were achieving comparable levels of accuracy on both trial types at the time of test. This suggests that, as in Experiment 1, progressing through the task and receiving constant feedback is sufficient for improving task performance. Importantly, bias strength in biased groups appears to have vanished by the time of test as indicated by choice performance on the Matchmaker task. These findings are striking when one considers the explicit rating of important compatibility factors indicated at test. Although the critical factor is rated as being more important at test than at baseline, it is still not rated as highly as the biased factor, which was overwhelming indicated most frequently as an important factor by all participants. Clearly there is a disparity between biased decision making during task performance and when asked explicitly about important factors.

Confidence ratings again show that biased groups become less confident on bias incongruent trials as the task progresses, which is to be expected as bias strength diminished and is no longer relied upon to make decisions when performing the Matchmaker task. All groups displayed quicker response times as the task progressed, likely due to increased familiarity with the procedure and rating system.

GENERAL DISCUSSION AND CONCLUSIONS

The experiments described in this paper aimed to determine if cognitive biases could be overcome by utilizing particular strategies: considering opposite hypotheses and warning of feedback error. These experiments also hoped to elucidate the cognitive processes that underlie decision making when operating under bias.

In the first experiment, encouraging participants to consider opposite hypothesis and thus make negative feedback more salient did not lead to better rule learning. Despite considering both good and bad matches for each bachelor in the Matchmaker task, participants in this group did not demonstrate better accuracy on any of the trial types throughout the task compared to those not making alternating matches. This could be due to the fact that those in the task-switch group were not effectively prompted to test opposite hypotheses by the instruction to make a bad match, but rather simply used their bias to assign matches to opposite bachelor. For example, instead of assigning Frank a match based on one of the other compatibility factors when told to make a bad match, participants likely noted Entertainment Preference and assigned matches to the bachelor who appeared biased toward the other Entertainment Preference option.

Both groups showed marked improvements in accuracy on bias incongruent trials from baseline to test and diminished accuracy on bias congruent trials, which indicates that the influence of bias was lessened as participants received feedback throughout the task. While the task-switch group was slower to select a match during the learning phase, this was likely due to the effort involved in reorienting to the presented instructions (good or bad match), rather than any attempt to formulate hypotheses that countered their initial bias.

Interestingly, those in the task-switch group were found to be more accurate in their matchmaking when making bad matches on bias incongruent trials, but not on other trial types.

This, combined with their low performance on bias congruent trials during the learning phase, suggests that the task-switch manipulation influenced feedback processing in some way. Exactly how this might have occurred remains unclear as performance during the learning phase did not influence final test outcomes.

In Experiment 1, bias strength diminished for both groups in similar ways as the task progressed, suggesting that simply proceeding through the task and receiving feedback was enough to lessen dependence on the bias initially created when first introduced to each bachelor. In this experiment, bias was not completely eliminated for either group, although it was reduced by nearly half for both groups. This reduction in bias strength is reflected in the explicit responses on the provided test forms, which required participants to rate the most important compatibility factors at baseline and test. Both groups rated the biased factor of Entertainment Preference lower at test than baseline and stated that Hair Color was more important at test than at baseline. Although bias strength diminished, participants still reported the biased compatibility factor of Entertainment Preference as being the most important compatibility factor in determining a best match when explicitly asked. Taken together, these results indicate that progression through the task while receiving probabilistic feedback lead to a reduction in bias at an implicit level as reflected in task performance, but this bias was still influencing the explicit determination of the most important compatibility factor.

A second experiment examined the ways in which warning participants of errors in feedback might impact their decision making on the Matchmaker task. Warning appeared to have no effect on the decision making process, as warned biased and unbiased groups achieved the same levels of accuracy as their unwarned counterparts. It is possible that participants in both groups had developed strong ideas of bachelor preferences after the baseline phase, and these ideas about what makes a good match persisted in spite of warnings of feedback error. It may also be the case that since participants were unable to draw strong conclusions about when feedback was inaccurate, the warning was ignored altogether.

All groups showed significant improvement in their accuracy performance on bias incongruent tasks and decreased accuracy on bias congruent trials, with accuracy on both at moderately above chance by the time of final test. These results suggest that after 60 learning phase trials, participants in the biased groups had learned that the original bias primed factor of Entertainment Preference was not predictive in making good matches and may have been testing other alternatives. Similarly, those in unbiased groups may have been testing hypotheses during the learning phase which lead to a better understanding of good matchmaking.

The measure of bias strength in biased groups reached zero by the end of the task, which indicates that reliance on Entertainment Preference had been eliminated by the time of test. However, explicit ratings of the importance of each compatibility factor on hypothesis tests at the test phase demonstrated that participants were still rating the biased factor of Entertainment Preference as the most important determinant of a good match. As with the first experiment, a discrepancy is found to exist between performance on the Matchmaker task, which indicates that participants no longer relying on bias to make good matches, and their explicit indicators of important factors. When asked directly, participants still express an allegiance to their initial bias.

One finding of this experiment that is difficult to explain is why participants in the unbiased groups demonstrated some preference for Entertainment Preference at baseline when this factor was never made specifically salient to them. Although performance on the Matchmaker task indicates that reliance on this bias had been eliminated at test, these participants also rated Entertainment Preference as the most important factor in assigning a good match when explicitly asked to do so. A potential explanation is that a match's description always listed Entertainment Preference first, and so participants paid more attention to this information than to the other compatibility factors that followed. Future iterations of this task should randomize the presentation of compatibility factors during the learning phase to test this hypothesis.

The salient finding from this set of these experiments reveals the power of feedback on decision making in probabilistic learning. Neither of the manipulations meant to encourage effortful, systematic processing of information lead to improvements in performance, suggesting that these cognitive processes are not useful for learning in these types of tasks. Similar outcomes were found in previous iterations of the Matchmaker task. Ledet (2013) reported that, in highlighting negative feedback through the use of an unpleasant buzzer sound and red feedback screen, accuracy was increased from baseline to test during the task but explicit statements of relevant factors still revealed a reliance on bias. In both of these experiments, it appears as though feedback alone was sufficient to improve overall accuracy. That this improvement in accuracy was revealed only through task performance suggests that this learning relies on automatic processes operating below the threshold of explicit awareness. When asked to explicitly state how decisions are being made, participants still reported using their biased reasoning, even though their task performance suggested otherwise. Thus it appears as though subconscious rule learning is occurring, but that this information has not been consciously integrated into conscious, deliberative reasoning processes.

These findings speak to recent critiques that dual process models, which advocate a quick and intuitive cognitive system and a slow and deliberate cognitive system, are overly simplistic

(De Neys, 2012). Rather than one system being purely intuitive and another system acting as purely rational, fMRI evidence has indicated that there is some subconscious awareness that fast, intuitive reasoning may be incorrect, allowing fast, rational processes to be utilized efficiently. Studies using conditional reasoning statements as stimuli have been used to demonstrate that when asked to make fast judgments about the validity of such statements, participants are able to use quick and intuitive cognitive processes to provide accurate answers (Thompson, 2014; Handley, Newstead, & Trippas, 2010). It was suggested that these errors in biased reasoning may have occurred because responses were based off of beliefs, which could be more compelling than simple logic.

Feedback in these experiments may have powerfully influenced intuitive reasoning due to being present completely and consistently. In the Matchmaker task, participants were given a correct or incorrect message after assigning every match, which lead to bias strength reduction. In the real world, this type of constant feedback in relatively uncommon. We might not always receive feedback or the feedback we chose to pay attention to might only serve to complement our already existing biases. For example, if one holds certain political beliefs, they are more likely to consume information from news sources that support those beliefs and ignore those counter them, or not be aware of every instance in which information regarding those beliefs is present so that they may adjust beliefs accordingly. In this case, feedback serves to support existing beliefs due to selectivity and inconsistency in presence. Consistency and completeness of feedback have been shown to be important in previous versions of the Matchmaker task, with feedback being reported intermittently (every 5 trials) was not as effective in improving learning as when it was present at every trial (Ledet, 2013).

In conclusion, the experiments presented in this study served to demonstrate the power of bias in making accurate decisions. Enlisting strategies meant to encourage systematic processing of information did not appear to be beneficial; rather, feedback alone served to reduce implicit reliance on bias and improve decision making. That this accuracy on task performance did not translate to explicit recognition of bias reduction suggests that conscious processes were not involved in eliminating reliance on bias. Future studies may be interested in investigating if prolonged time in the learning phase would encourage the involvement of explicit cognitive processes, leading to conscious awareness of correct rule learning.

REFERENCES

Arkes, H.R., Faust, D, Guilmette, T.J., & Hart, K. (1988). Eliminating the hindsight bias. *Applied Psychology, 73,* 305 – 307.

Biros, D.P., George, J.F., & Zmud, R.W. (2002). Inducing sensitivity to deception in order to improve decision making performance: A field study. *MIS Quarterly, 26,* 119 -144.

Bodenhausen, G.V. (1988). Stereotypic bias in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology, 55,* 726 – 737.

Bodenhausen, G.V., & Wyer, R.S. (1985). Effects of stereotypes on decision making and information-processing strategies. *Journal of Personality and Social Psychology, 44,* 267 – 282.

Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment and uncertainty. *Cognition, 58,* 1 – 73.

Cox, J.L., & Popken, D.A. (2008). Overcoming confirmation bias in causal attribution: A case study of antibiotic resistance risks. *Risk and Analysis, 28,* 1155 – 1171.

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science, 7,* 28 – 38.

Downar, J., Bhatt, M., & Montague, P.R. (2011).Neural correlates of effective learning in experienced medical decision-makers. *PLoS One, 6.*

Gilbert, D.T. (1991). How mental systems believe. *American Psychologist, 46,* 107 – 116.

Gladwell, M. (2005). *Blink.* New York: Little, Brown and Company.

Graesser, A.C., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology, 10,* 17 – 31.

Handley, S.J., Newstead, S.E, & Trippas, D. (2010). Logic, beliefs, and instruction: A tests of the default interventionist account of belief bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 28 – 43.

Hoch, S.J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 719 – 731.

Kahneman, D. (2011). *Thinking, Fast and Slow.* New York: Farrar, Straus, and Giroux.

Klein, G. (1999). *Sources of Power: How People Make Decisions.* MIT Press.

Koriat, A., Lichtenstein, S., & Fischoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107 – 118.

Kray, L. J., & Galinsky, A.D. (2003). The debiasing effect of counterfactual mind-sets: Increasing the search for disconfirmatory information in group decisions. *Organizational Behavior and Human Decision Processes, 91,* 69 – 81.

Ledet, P. (2013). *Influencing the confirmation bias on a Matchmaker task through manipulation of the feeling of rightness* (Unpublished doctoral dissertation). Louisiana State University, Baton Rouge, LA, USA.

Lilienfield, S.O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives of Psychological Science, 4,* 390 – 398.

Lord, C.G., Lepper, M.R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 47,* 1231 – 1243.

Melcher, J.M. & Schooler, J.W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language, 35*, 231 – 245.

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin, 26,* 1142 – 1150.

Natter, H.M. & Berry, D.C. (2005). Effects of active information processing on the understanding of risk information. *Applied Cognitive Psychology, 19,* 123 – 135.

Nickerson, R.S. (1998). Confirmation Bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175 – 220.

Sanna, L.J., Schwarz, N., & Stocker, S.L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 497 – 502.

Schwarz, et al. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61,* 195 – 202.

Silverman, B.G. (1992). Modeling and critiquing the confirmation bias in human reasoning. *IEEE Transactions on Systems, Man, and Cybernetics, 22,* 972 – 982.

Stern, E., Aprea, C., & Ebner, H.G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13,* 191 – 203.

Stiff, J.B., Kim, H.J., & Ramesh, C.N. (1992). Truth biases and aroused suspicion in relational deception. *Communication Research, 19,* 326 – 345.

Thompson, V., & Morsanyi, K. (2012). Analytic thinking: Do you feel like it? *Mind & Society, 11,* 93 – 105.

Thompson, V.A., Gibb, M., & Newman, I. (2014, November). Thinking fast and slow: Analytic beliefs and intuitive logic? Paper presented at the meeting of the Psychonomic Society, Long Beach, CA.

Thompson, V.A., Turner, J.A., & Penncook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology, 63,* 107 – 140.

Topolinksi, S., Likowski, K.U., Weyers, P., & Strack, F. (2009). The face of fluency: Semantic coherence automatically elicits a specific pattern of facial muscle reactions. *Cognition and Emotion, 23,* 260 – 271.

Topolinski, S., & Reber, R. (2010). Gaining insight into the “aha” experience. *Current Directions in Psychological Science, 19,* 402 – 405.

Toris, C. & DePaulo, B.M. (1985). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology, 47,* 163 – 1073.

Wason, P.C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129 – 140.

Williams, J.J, Lombrozo, T., & Rehder, B. (2013). The hazard of explanation: Overgeneralization in the face of exceptions. *Journal of Experimental Psychology, 142,* 1006 – 1014.

Zacharakis, A.L.& Shepard, D.A. (2001). The nature of overconfidence on venture capitalists’ decision making, *Journal of Business Venturing, 16,* 311 – 332.

VITA

Dina Marie Acklin was born in Pittsburgh, Pennsylvania to parents Rochelle Staley and Theodore E. Acklin, Jr. She graduated with honors in 2008 from the University of Pittsburgh in Pittsburgh, Pennsylvania earning a B.S. in Psychology. After working at the University of Pittsburgh Medical Center and Colorado State University, she entered graduate school at Louisiana State University in 2012 under the supervision of Dr. Robert Mathews. She will continue to pursue her Ph.D. at the same university.