PHILIP THOMAS DUNWOODY, II
The Use of Base Rate Information as a Function of Experienced Consistency and Utility
(Under the Direction of Robert P. Mahan)

The use of base rate information has been widely studied in decision making with the conclusion that people underweight or ignore base rate information when compared to a normative standard. This work extends the current body of research by demonstrating that base rate usage is moderated by the statistical characteristics of the base rate information. Two studies demonstrated that experienced base rate consistency and utility both affect base rate usage. Experiment 1 showed that participants use base rate information more often when it is consistent than when it is inconsistent. Experiment 2 showed that when base rate consistency and utility are manipulated separately, participants decisions are mostly influenced by the utility of the base rates and not the consistency. These studies demonstrate that base rate usage can be an adaptive response to environmental contingencies.

INDEX WORDS: Base Rate Neglect, Consistency, Utility, Decision Making

THE USE OF BASE RATE INFORMATION AS A FUNCTION OF EXPERIENCED CONSISTENCY AND UTILITY

by

PHILIP THOMAS DUNWOODY, II

B.A., Richard Stockton College of New Jersey, 1994

M.S., The University of Georgia, 1998

A Dissertation Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2000

THE USE OF BASE RATE INFORMATION AS A FUNCTION OF EXPERIENCED CONSISTENCY AND UTILITY

by

PHILIP THOMAS DUNWOODY, II

Approved:

Major Professor: Robert P. Mahan

Committee: Steven Beach
Adam Goodie
Richard Marsh
Leonard Martin

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
December 2000

ACKNOWLEDGMENTS

A dissertation is ideally the apex of one's academic studies. Consequently, everyone who has been involved in my educational experience deserves an acknowledgment. However, I will settle for a shorter list. My Dissertation Advisory Committee, consisting of Robert Mahan, Leonard Martin, Richard Marsh, Adam Goodie and Steven Beach, has pushed me to think critically and improve my work. My lab mates, Christopher Marino and Eric Haarbauer have always been there to help me fumble through and issue or two; as well as push me to do just one more shot of tequila. My family has done more than I can possible express and I am eternally grateful. My wife Marlene is the light of my life and is always there to shine upon my day no matter how wet and soggy I come home. I would also like to thank my in-laws for help making Georgia a home to me.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

Base rate neglect is one of the most widely studied biases within the area of judgment and decision making. Despite this fact, the validity of the claim that people ignore or largely underweight base rate information is ambiguous. Koehler (1996) provides a recent review of the base rate fallacy and argues that the conditions under which this phenomenon occurs have been largely ignored. He states there is little understanding of "how the ambiguous, unreliable, and unstable base rates of the real world are and should be used" (p. 1). This research explores the adaptive nature of base rate usage under conditions of direct experience with unstable and unreliable base rates.

When people are given specific (or individuating) information about a case, plus information about the population distribution from which the case was drawn (or base rate information) they often underweight the base rate information when judging the likelihood of an event. For example, the following question is a commonly used base rate problem (Kahneman & Tversky, 1972; Bar-Hillel, 1980; Tversky & Kahneman, 1980):

> A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:
>
> (a) 85% of the cabs in the city are Green and 15% are Blue.
>
> (b) A witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident

> and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time. What is the probability that the cab involved in the accident was Blue rather than Green? (Tversky & Kahneman, 1982, p. 157)

Bayes's Theorem is the normative rule usually used to obtain the criterion against which participants' judgments are compared. Using the odds form of Bayes's Theorem where B and G represent the hypotheses that the cab was Blue or Green and W represents the witness's report, the correct answer may be obtained by the following formulas:

$$\frac{P(B \mid W)}{P(G \mid W)} = \frac{P(W \mid B)P(B)}{P(W \mid G)P(G)} = \frac{(.8)(.15)}{(.2)(.85)} = \frac{.12}{.17}$$

$$P(B \mid W) = \frac{.12}{.12 + .17} = .41$$

In this example, the base rate is more extreme than the witness is credible and therefore the cab is more likely to be Green (.59) than Blue (.41). Despite this fact, the typical response (.80) corresponds with the individuating information and is unaffected by the base rate (Tversky & Kahneman, 1982, p. 157). There are many other examples in the literature of participants ignoring the base rates and instead relying predominantly on the specific information (Kahneman, Slovic, & Tversky, 1982; Koehler, 1996). This finding was considered so robust that Bar-Hillel (1980) stated, "The genuineness, the robustness, and the generality of the base-rate fallacy are matters of established fact" (p. 215).

However, this strong view about the robustness of base rate neglect has not lasted. Koehler (1996) argues, “We have been oversold on the base rate fallacy in probabilistic judgment from an empirical, normative, and methodological standpoint” (p. 1). He argued that few studies demonstrate a total neglect of base rates. More common are studies where base rate usage is less than is prescribed by Bayes’s Theorem. He also questioned the appropriateness of Bayes’s Theorem as an unambiguous judgment criterion, arguing that there are few examples in the real world where Bayes's Theorem can unambiguously be mapped to the problem space.

Similarly, Hammond (1996) points out that there is a distinction between the coherence of a judgment (e.g., how well it matches with Bayes’s Theorem) and the correspondence of a judgment (how well it predicts events in the world). Research on base rate usage that emphasizes the rationality, or coherence, of a judgment typically evaluates judgment performance against a normative criterion (i.e., Bayse's Theorem). An alternative approach to emphasizing the coherence of a judgment is to emphasize the correspondence of a judgment. Researchers that emphasize the correspondence of a judgment evaluate how well decision makers adapt to different environments and argue that there is no single normative criterion but rather base rate usage should be adaptive from a performance standpoint.

This correspondence/coherence distinction is clarified when coupled with Popper’s (1972) three-world framework. Bjorkman (1984), in describing Popper's three-world framework, notes that World 1 is that of physical objects, World 2 is that of the subjective experience, and World 3 is that of scientific concepts. Evaluating the accuracy of a person’s perceptions (World 2) with events in the world (World 1) is an assessment

of the corresponding quality of a judgment. That is, how well does the judgment correspond to the real world? Evaluating a match between a person's perceptions (World 2) with a scientific formula (World 3) is an assessment of the coherence, or rational quality, of a judgment. However, the mathematical models of World 3 do not necessarily match with the reality of World 1. For example, Newtonian physics (a model from World 3) has been a useful scientific tool despite the fact that it does not perfectly predict events in the world (World 1).

These distinctions are important because the vast majority of research showing base rate neglect has been of the coherence variety. That is, judgments (World 2) have been assessed against a mathematical model (World 3, i.e., Bayes's Theorem) as the criterion. Little consideration has been given to how well Bayes's Theorem maps onto events in the real world.

The assumption of many researchers is that the world is Bayesian and therefore we ought to be as well. This assumption is often tested by using word problems with summary statistics (as shown above in the cab problem) and comparing peoples' answers with that obtained by Bayes' Theorem. Lyon and Slovic (1976) state, "Since the world operates according to Bayes's Theorem, experience should confirm the importance of base rates" (p. 296-297). Most research of the coherence approach (using summary statistics in word problems and comparing the observed answers with that given by Bayes' Theorem) has demonstrated that Bayesian reasoning is not part of our judgmental repertoire. Rather, the opposite conclusion has been reached. Kahneman and Tversky (1973) state, "In his evaluation of evidence, man is apparently not a conservative Bayesian: he is not Bayesian at all." (p. 237). This failure to reason in a Bayesian manner

has been largely attributed to processing limitations and inherent biases (Lichtenstein, Fischoff & Phillips, 1982; Tversky & Khaneman, Chapter 10, 1982).

However, Bjorkman (1984) and Hammond (1996) note that people do not encounter Bayes's Theorem, a portion of World 3. What they encounter are events in World 1. These events may be summed up as descriptive statistics or turned into formulae but what people usually experience are events. Bjorkman states, "World 1 is not in itself Bayesian, it is the Baysians' representation of it that is Bayesian" (p. 409). Despite this view, many researchers implicitly assume that base rate information is unambiguous and should be considered a veridical representation of the world.

A similar representation argument is made by Gigerenzer and Hoffrage (1995). They argue that "organisms did not acquire information in terms of probabilities and percentages until recently" (p. 686) and we should therefore not expect people to naturally make appropriate use of this information. They argue that the information representation (percentages and probabilities) may not match with people's internal representations. It is like presenting an Arabic math problem to someone who only learned Roman Numerals and then concluding that they have no understanding of mathematics when they fail to solve the problems. While the two representations are mathematically equivalent, they are not psychologically equivalent. Hence, Gigerenzer and Hoffrage propose that a more natural format would increase people's Bayesian responses. What would a more natural information representation look like?

"We assume that as humans evolved, the 'natural' format was frequencies as actually experienced in a series of events, rather than probabilities or percentages" (Gigerenzer & Hoffrage, 1995, p. 686). While this is a compelling argument, Gigerenzer

and Hoffrage do not utilize directly experienced events. Instead, they provide participants with a series of word problems using summary statistics in frequency terms. For example, “Five out of 100 tests are positive” is the frequency format of “5% of the tests are positive.” By using the same word problems and changing only the numerical representation, they demonstrated that participant Bayesian responses increased significantly when a frequency, rather than a percentage, representation was used.

While Gigerenzer and Hoffrage (1995) offer evidence that a frequency representation increases Bayesian responses, they do not directly test conditions of directly experienced base rates. Most studies examining base rate usage under direct experience are of the coherence variety (Estes, Campbell, Hatsopoulos, & Hurwitz, 1989; Gluck & Bower, 1988; Goodie & Fantino, 1995). Goodie and Fantino (1999a) state,

> Surprisingly…studies that demonstrate base-rate neglect under direct experience have not manipulated the base rates that, the researchers conclude, subjects neglect. The inference is instead drawn from data points that deviate from an optimal standard in a direction consistent with base-rate neglect. (p. 159).

Goodie and Fantino (1999a) manipulated base rates and cue accuracies to see whether decisions would vary in a manner consistent with Bayes's Theorem. They found that decisions were affected by differing base rates but less than Bayes's Theorem would suggest. Despite this fact, Goodie and Fantino argued that in some environments this could be an adaptive strategy. They state,

> In a world where base rates and cue accuracy can change, sometimes without warning, what is the best way for an organism to proceed? More specifically, when base rates change relatively often and cue accuracy relatively seldom, isn't it

> just as well to underweight base rates, since they're liable to change at any moment? (1999b, p. 327).

Similarly, Brunswik (1956, 1957) argued that researchers ought not to implicitly assume participants should treat stimuli as perfectly reliable information sources. He argued that humans evolved to function in environments filled with redundant and less than perfect information sources. Consequently, it would not be functionally adaptive to rely solely on one information source. It would be more evolutionarily adaptive to integrate many information sources as a function of their individual utilities. Such an integration style is robust whereas sole reliance on an individual information source could potentially be catastrophic.

Goodie and Todd (in preparation) made this same argument and tested it in a Monte Carlo simulation. They found that an irrational strategy, e.g. base rate neglect, could rival Bayesian integration in performance given the right environment. In one environment tested where base rates changed more often than cue accuracies, simulated participants who neglected base rates gave correct responses 72% of the time. Simulated participants who used Bayesian integration gave correct responses 74% of the time. Although the Bayesian integration was more successful, the difference was very small.

These views suggest it can be adaptive to neglect base rates given a particular environment and that humans, as adaptive organisms will be responsive to such environments. It is possible that people use base rate information to the extent they believe it is adaptive to do so. Although there is no empirical evidence that people will differentially use base rates as a function of base rate consistency or utility, indirect evidence suggests this to be true.

Studies have shown that as the perceived credibility of base rate information increases, so does its usage. For example, Ginossar and Trope (1987) found that base rate information usage increased as the credibility of the individuating information source decreased (e.g., the individuating information source was either a palm reader or a psychologist). While this example involves the use of summary statistics and a similar manipulation of credibility, this finding is likely to generalize to conditions of direct experience. For example, one is likely to use base rate information more when one has experienced reliable base rates than when one has experienced unreliable base rates. While there are studies that examine base rate usage under direct experience (e.g., Goodie & Fantino, 1995; 1999a; 1999b), none have examined base-rate use as a function of experienced base rate consistency.

In two experiments reported in this paper participants directly experienced base rates. In both experiments participants were shown a cue that was designed to correspond to the individuating information used in previous base rate research. In the previously discussed cab example a witness testified that the cab was blue. Similarly, in the present experiments participants received a cue that was either a blue or a green square. In the cab example participants are told the credibility of the witness whereas in the present experiment they experience the credibility of the cue (e.g., the probability of the cue given the outcome). In the cab example participants were told the base rate proportions of green and blue cabs. In the present experiments participants experience the base rates by seeing that a blue or green square is the correct answer for each trial. In such a trial-by-trial situation, the base rate is the total proportion of times that blue or green is the

correct answer. This direct experience design allows for examination of the adaptive nature of base rate usage as a function of base rate consistency.

CHAPTER 2

EXPERIMENT 1

Experiment 1 addressed the issue of information use as a function of information consistency. I hypothesized that people would use consistent information more than inconsistent information. This hypothesis follows from the idea that people will use information more when they believe it to be credible. Reliable base rates are likely to be viewed as more credible than unreliable base rates. Specifically, I predicted people would use base rates more when base rates are consistent than when base rates are inconsistent.

This usage pattern should also result in people using the specific information, or cue accuracy, more when it is consistent than when it is inconsistent. There is no performance advantage for choosing one strategy over the other. That is, even though the information sources vary in their consistency, they are, in the long run, equal in their utilities. Consequently, participants can choose to use the consistent information or always use the cue, regardless of its consistency, without detrimental performance.

Method

Participant. Forty-four undergraduates participated for course credit. Upon arrival participants were randomly assigned to either the consistent base rate/inconsistent cue accuracy or the inconsistent base rate/consistent cue accuracy condition. Please note that when the base rate is consistent or inconsistent, the cue accuracy is always the opposite.

Twenty-two people participated in each condition. The mean age was 19.4 (s=1.44), 9 were male, 31 were female, and 4 participants failed to indicate their gender on the questionnaire.

Materials. The surface characteristics of the stimuli were identical for both groups while the depth characteristics (i.e., base rates and cue accuracies) varied between conditions. Base rate was defined as the proportional reinforcement for a particular response. For example, the green cue was reinforced 70% of the time. Cue-accuracy was defined as the probability of the cue given the outcome, or $p(c|o)$. This probability was also set at 70%. The probability of the outcome given the cue, or $p(o|c)$ was held constant.

The first 20 trials were identical in both conditions with a base rate of 70% green and a cue accuracy of 70%. After the first 20 trials either the base rate or cue-accuracy fluctuated every subsequent twenty trials. Participants performed 200 trials with either the cue accuracy or the base rate fluctuating every twenty trials. They then performed 100 additional trials in which the cue accuracy and base rate were held constant at 70% reinforcement. The initial 200 trials served only to give subjects experience in two unique environments. Consequently, all analyses were conducted on the last 100 trials. This design allowed for participants to learn the task in two separately designed environments even though the testing environment (the last 100 trials) was identical. This design is central to the hypothesis that information usage can be predicted as a function of previous experience.

The structure of the task is schematically depicted in Figure 1. After the inter-trial-interval (ITI) the green cue was presented 58% percent of the time with the blue cue

presented the other 42% of the time. When the green cue was presented, green decisions were reinforced 84% of the time with blue decisions being reinforced 16% of the time. However, when the blue cue was presented, green and blue judgments were reinforced equally (50%).

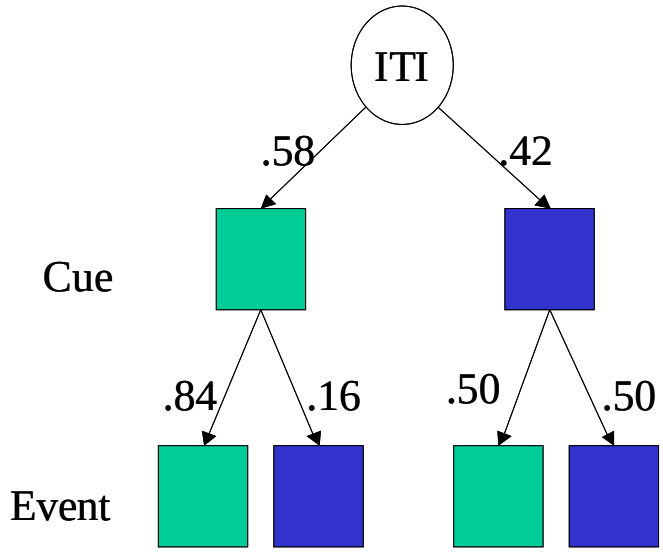

Figure 2. Schematic of Experiment 1 Reinforcement.

Since the experiment was designed such that the base rate favored green choices while matching the cue was reinforced 70% of the time, the informative decision block is the right-hand side of Figure 1, where the blue cue is presented. When the blue cue was presented, green and blue choices were rewarded equally. In this condition, if participants utilize a general base rate strategy they will choose green more than blue since the overall base rate is 70% green. However, if they utilize a general cue-matching strategy they will choose blue since cue-matching is reinforced 70% of the time.

Procedure. Participants were told that they would be presented with either a green or blue square as a cue (the specific information) and that they are to guess if the correct answer is a green or blue square. They were told that they would receive one point for every correct answer and it was goal to collect as many points as possible.

Participants were then seated at a computer and instructed to use the mouse to make their choices and proceed through the experiment. On each trial, a green or blue square appeared at the top of the screen. Below were both a green and a blue square and participants chose one of these as their response. Immediately following their choice they were given outcome feedback that said either; "Correct. You now have __ points" or "Sorry, that is incorrect. You still have __ points." See Figure 2 for a screen shot of one completed trial. Participants were asked if they understood the task and the feedback after the first 10 trials. If they did not, clarifications were made until they affirmed that the task and feedback were clear. They were instructed to proceed until the experiment ended and the computer instructed them to go outside the room and talk with the experimenter. At this time they were debriefed.

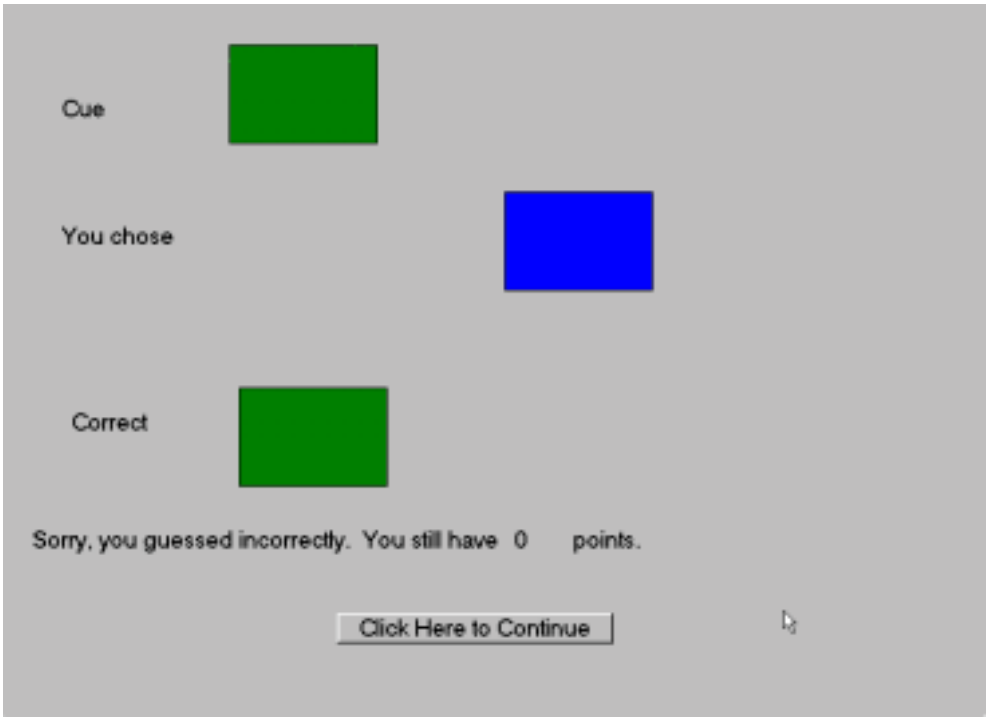

Figure 2. Screen capture of one completed trial.

Experimental Design. Participant decisions after blue-cue presentation were compared for the consistent base rate and consistent cue-accuracy conditions via a non-directional *t*-test.

Results And Discussion

Participants who experienced consistent base rates in the first 200 trials chose green 56% (*SEM*=0.047) of the last 100 trials, while participants who experienced consistent cue accuracies (and fluctuating base rates) in the first 200 trials chose green 37% (*SEM*=0.057) of the last 100 trials. This difference was statistically significant in the predicted direction ($t_{obs}(42)=2.631$, $p<0.05$). While the hypothesis was specified as directional a priori, evaluating the difference with a non-directional test still yields a statistically significant difference ($p<0.05$).

An examination within each group indicates that there may be differential sensitivity for cue accuracies and base rates. The group that experienced consistent base rates and inconsistent cue accuracies selected base rate responses more than they matched the cue (56%/44%). However, this difference was not statistically significant ($t_{obs}(21)=1.067$, $p>0.05$). Conversely, the group that experienced inconsistent base rates and consistent cue accuracies matched the cue more than they chose the base rate (63%/37%). This difference was significant ($t_{obs}(21)=2.835$, $p<0.01$). At first glance this appears to indicate that participants are more sensitive to consistent cue accuracies than consistent base rates. However, one must consider that people have a strong tendency to match the cue (Goodie & Fantino, 1996). Given that people have a tendency to match the cue, the 56% of base rate choices that go against cue matching may be viewed as evidence that people are sensitive to consistency in base rates and not just cue accuracies. However,

since this experiment was designed primarily as a between groups comparison, a decisive answer to this question will have to await future research. Future research that addresses this issue should be choose cues and responses that avoid the possibility of cue matching. Previous research has shown that unrelated cue-response stimuli do result in different choice patterns than when related cue-response stimuli, which allow for cue matching, are used (Goodie & Fantino, 1996).

These results support the hypothesis that experienced information consistency influences the likelihood that participants will utilize that information source. In short, participants are more likely to use base rate information when it is consistent than when it is inconsistent. This finding raises the possibility that some of the base rate neglect observed in the literature may be due to people's prior experience with base rates. This experiment demonstrates that base rate usage (or neglect) is influenced by past experience with base rates.

CHAPTER 3

EXPERIMENT 2

Experiment 1 demonstrated that base rate consistency affects base rate usage, Experiment 2 was designed to test a likely moderator of this effect. In Experiment 1 the groups differed in consistency of the information but not in the utility of the information. Experiment 2 was designed with two main goals; first, to replicate the findings from Experiment 1 and second, to evaluate how choice judgments are affected when information consistency and utility are pitted against each other. That is, when the consistent information is less useful than the inconsistent information, which will people utilize?

As in Experiment 1, information consistency was manipulated. However, the two consistency levels were crossed with two utility levels for a total of four conditions. Consistency was defined as it is in Experiment 1, with each parameter varying every twenty trials or remaining constant. Utility was defined as the percent correct a participant could obtain by solely utilizing that information source. For example, high base rate utility is defined as reinforcing green decisions 70% of the time and low base rate utility is defined as reinforcing green decision 60% of the time. Likewise, high cue accuracy utility is defined as reinforcing cue-matching 70% of the time and low cue accuracy utility is defined as reinforcing cue-matching 60% of the time. This 2 (consistency) by 2 (utility) designed is represented in Figure 3.

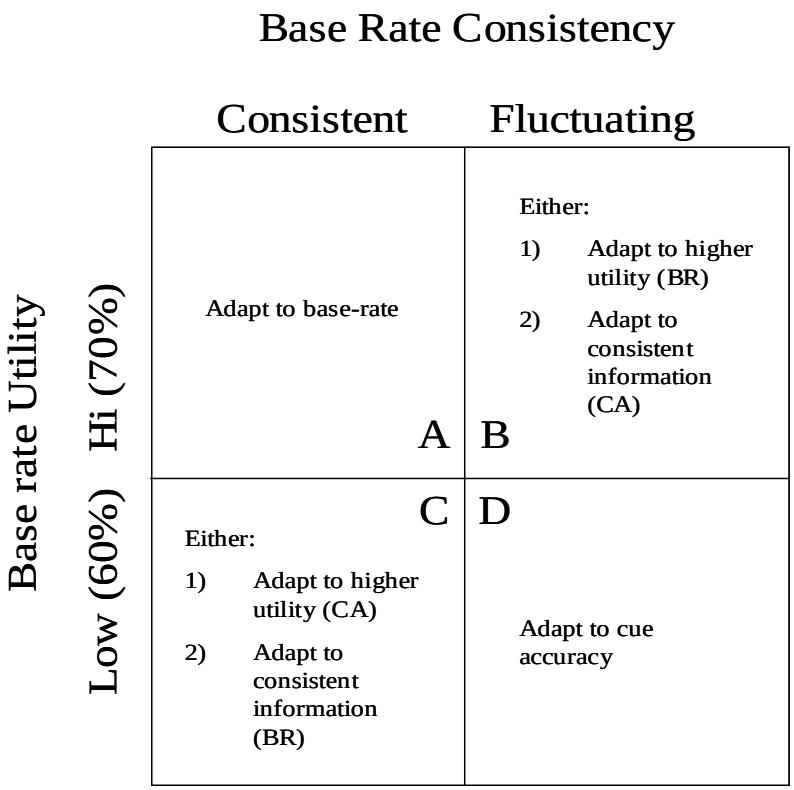

Figure 3. Two (base rate consistency) by two (base rate utility) design for Study 2.

Two planned comparisons were planned. The first planned contrast compared participants' green choices after a blue cue for cells A and D of Figure 3. This comparison was designed to replicate the findings from Experiment 1. In cell A of Figure 3 the base rate is of higher utility and is more consistent than the cue-accuracy. Conversely, in cell D the cue accuracy is of higher utility and consistency than the base rate. Because high consistency is coupled with high utility, it was expected that this difference will replicate that of Experiment 1. It is likely that the raw effect size will be larger since utility and consistency are being simultaneously manipulated.

The second planned contrast compared the proportion of participants' green choices after a blue cue for cells C and D. This comparison was designed to evaluate how

choice judgments are affected when information consistency and utility are pitted against each other.

In cell C the base rate is of lower utility than the cue accuracy but the base rate is constant. Participants may favor either the consistent information (the base rate) as they did in Experiment 1 or the information higher in utility (the cue accuracy). In cell B the base rate is of higher utility than the cue accuracy but the base rate fluctuates. Again, participants may favor either the consistent information (the cue accuracy) as they did in Experiment 1 or the information higher in utility (the base rate).

There are five possible outcomes that may be observed when participants' proportions of green choices after seeing a blue cue are compared between cells C and B. The first is the null hypothesis, that there would be no difference between green choice proportions when examining cells C and B. The other four outcomes are addressed below.

Outcome 1: Participants in cells C and B could adapt to the consistent information despite the utility. This outcome is plausible from a satisficing (Simon, 1957). perspective. That is, even though performance can be improved by using the information highest in utility, participants may be satisfied with their level of performance using the consistent information. However, previous research utilizing a similar design found that providing a monetary incentive did not change participant behavior (Goodie & Fantino, 1995, Experiment 2). Although Experiment 1 has already demonstrated that base rate usage is influenced by information consistency, it is unclear if this consistency adaptation will continue despite the utility of the information.

Outcome 2: Participants in cells C and B could adapt to the information highest in utility despite the consistency. This outcome is plausible from the "human as intuitive statistician" perspective (Brunswik, 1956). According to this perspective, we are inherently adaptive organisms and therefore will be intuitively sensitive to the statistical utility of the information. Such a finding would mean that participants were sensitive to a utility difference of only 10% (60% vs. 70%). This finding would provide an obstacle for theories that claim base rate neglect is an inherent bias built into our information processing (Lichtenstein, Fischoff & Phillips, 1982; Tversky & Khaneman, 1982).

Outcome 3: Participants could neglect the base rates and simply match the cue. Specifically, they will utilize the cue-accuracy when it is of higher utility and fluctuating in cell C and they will utilize the cue-accuracy when it is of lower utility and consistent in cell B. This outcome is plausible from the inherent bias base rate neglect perspective (Kahneman, Slovic, & Tversky, 1982).

Outcome 4: Participants could neglect the cue-accuracy and use the base rates. Specifically, they could use the base rate when it is consistent but of lower utility in cell C and they could use the base rate when it is of higher utility but fluctuating in cell B. There is little evidence to suggest that this outcome will occur.

Two of the possible outcomes (outcomes 1 and 2) require that green choices after the blue presentation in cells C and B will be significantly different. Two of the possible outcomes (outcomes 3 and 4) predict that green choices after the blue presentation will be approximately equal for cells C and B. To discriminate between these two groups of outcomes a planned comparison was conducted. If a significant difference is found, outcomes 3 and 4 will be ruled out. If the planned comparison shows that green choice

proportions in cell C are significantly greater than green choice proportions in cell B, then outcome 1 will be supported. This observed outcome pattern would indicate that participants utilize consistent information despite its lower utility. Conversely, if green choice proportions in cell C are significantly smaller than green choice proportions in cell B, then outcome 2 will be supported. This observed outcome pattern would indicate that participants utilize information high in utility even when that information fluctuates in its utility.

If no significant difference is found between cells C and B then outcomes 3 and 4 remain. Although outcomes 3 and 4 do not predict that green choice proportions will be different between cells B and C, they do predict that green choice proportions will either be below .5 in both cells (outcome 3) or above .5 in both cells (outcome 4).

Consequently, if the planned comparison shows no significant difference between cells C and B, then the choice proportions for both cells will be combined and compared against .5 via a non-directional *t* test. If the observed choice proportion is significantly lower than .5 then outcome 3 will be supported. If the observed choice proportion is significantly greater than .5 then outcome 4 will be supported.

Method

Participants. One-hundred undergraduates participated for course credit. Upon arrival participants were randomly assigned to the four conditions resulting in 25 participants per cell. The mean age was 19.9 (*s*=3.33), 25 were male, 73 were female, and 2 participants failed to indicate their gender on the questionnaire.

Materials. The surface characteristics of the stimuli were identical to that in Experiment one. The depth characteristics of the stimuli include the consistency

manipulation in Experiment 1 with the addition of a utility manipulation. Participants in high utility conditions experienced 70% reinforcement and participants in low utility conditions experienced 60% reinforcement. The four conditions, two levels of consistency and utility, are represented in Figure 3. Part way through Experiment 2 a post-experimental questionnaire was added to assess the participants' retrospective performance strategy and awareness of base rates and cue accuracies. The specifics of this questionnaire are addressed in Tables 1, 2, and 3 of the following results and discussion session.

Procedure. Procedures were identical to those used in Experiment 1.

Experimental design. The four-cell design (see Figure 3) was analyzed via the planned contrasts discussed in the previous section.

Results And Discussion

On average participants who experienced base rates that were high in utility and consistent (cell A of Figure 3) in the first 200 trials chose green 58% (*SEM*=0.047) of the last 100 trials. Conversely, participants who experienced base rates that were low in utility and fluctuating (cell D in Figure 3) in the first 200 trials chose green 31% (*SEM*=0.038) of the last 100 trials. This difference was statistically significant and in the predicted direction ($F_{obs}(1,96)=21.41$, $p<0.001$). This planned contrast replicates the findings of experiment 1. The mean difference for the consistency comparison in Experiment 1 is 19% while the mean difference for the consistency/utility comparison in Experiment 2 is 27%. The larger mean difference in experiment 2 implies that the utility manipulation had an effect in addition to the consistency manipulation.

An examination within each group again indicates that there may be differential sensitivity for cue accuracies and base rates. The group that experienced consistent base rates of high utility selected base rate responses more than they matched the cue (58%/42%). However, this difference was not statistically significant (t_{obs} (24)=1.712, p>0.05). Conversely, the group that experienced consistent cue accuracies of high utility matched the cue more than they chose the base rate (69%/31%). This difference was significant (t_{obs} (24)=5.045, p<0.01).

This finding strengthens the claim made in Experiment 1 that participant use of base rate information is influenced by past experience. However, it does not disambiguate the effects of consistency and utility. The second planned contrast is aimed at disambiguating the effects of consistency and utility.

On average participants who experienced base rates that were high in utility and fluctuating (cell B of Figure 3) in the first 200 trials chose green 69% (*SEM*=0.045) of the last 100 trials. Conversely, participants who experienced base rates that were low in utility and consistent (cell C in Figure 3) in the first 200 trials chose green 26% (*SEM*=0.037) of the last 100 trials. This difference was statistically significant (F_{obs}(1,96)=64.94, p<0.001).

The group that experienced fluctuating cue accuracies of high utility matched the cue more than they chose the base rate (69%/31%). This difference was statistically significant (t_{obs} (24)=6.583, p>0.01). Conversely, the group that experienced fluctuating base rates of high utility chose the base rate more than they matched the cue (69%/31%). This difference was significant (t_{obs} (24)=4.324, p<0.01).

The results of the second planned contrast of Experiment 2 provide a clear answer to the question, "when the consistent information is less useful than the inconsistent information, which will people utilize?" Participants overwhelming utilized the information that was highest in utility despite information consistency. Of the 50 participants that were in cells C and B, 43 (86%) had a choice proportion favoring the information highest in utility, 5 (10%) had a choice proportion favoring consistent information, and 2 (4%) had choice proportions that showed no strong preference (between 0.45 and 0.55).

When base rates were of higher utility than the cue accuracy, participants chose base rate responses an average of 69% of the time. When cue accuracy was higher in utility than base rates participants matched the cue 74% of the time. This is a large difference favoring information that is most useful; particularly when one considers that the utility difference was only 10 percentage points (60% vs. 70%).

Part way through experiment 2 a post-experimental questionnaire was added. Thirty-nine of the 100 participants in Experiment 2 completed the questionnaire. Each question is listed below with tables indicating group descriptives.

Question 1 (forced choice)

Would you describe your decision making strategy as primarily…

a) choosing the same color as the cue color

b) choosing the opposite color of the cue color

c) primarily choosing one color regardless of the cue color

Table 1. Answers to question 1.

	High Utility and Consistent Base Rates (Group A)	Low Utility and Fluctuating Base Rates (Group D)	High Utility and Fluctuating Base Rates (Group B)	Low Utility and Consistent Base Rates (Group C)
a) choosing the same color as the cue color	36% (n=4)	18% (4)	29% (2)	50% (5)
b) choosing the opposite color of the cue color	9% (1)	32% (7)	0% (0)	0% (0)
c) primarily choosing one color regardless of the cue color	55% (6)	50% (11)	71% (5)	50% (5)

Cells indicate percentage and number of participants in each group that chose each answer.

Question 2 (forced choice)

If you chose answer 'c' above, what color did you primarily choose?

Table 2. Answers to question 2

	High Utility and Consistent Base Rates (Group A)	Low Utility and Fluctuating Base Rates (Group D)	High Utility and Fluctuating Base Rates (Group B)	Low Utility and Consistent Base Rates (Group C)
Green	86% (n=6)	86% (6)	100% (5)	83% (5)
Blue	14% (1)	14% (1)	0% (0)	17% (1)

Cells indicate percentage and number of participants in each group that chose each answer.

Most participants (55%) who experienced consistent base rates of high utility reported primarily choosing one color regardless of the cue color (that is, choosing base rates). However, 50% of participants who experienced fluctuating base rates of low utility

also reported primarily choosing one color regardless of the cue color. Of the participants who reported primarily choosing one color regardless of the cue color, the vast majority reported choosing green (see Table 2).

Questions 3 and 4 (fill in the blank)

Choosing a color that matched the cue was correct ___% of the time.

Regardless of the cue color, green was the correct answer ___% of the time.

Table 3. Answers to questions 3 and 4.

	High Utility and Consistent Base Rates (Group A)	Low Utility and Fluctuating Base Rates (Group D)	High Utility and Fluctuating Base Rates (Group B)	Low Utility and Consistent Base Rates (Group C)
Choosing a color that matched the cue was correct ___% of the time.	M=59.55 S=13.13	M=64.09 S=20.10	M=57.86 S=23.43	M=65.00 S=17.00
Regardless of the cue color, green was the correct answer ___% of the time.	M=69.55 S=10.60	M=63.64 S=13.98	M=73.57 S=16.50	M=56.00 S=13.50

Cells indicate mean response with standard deviation for each question.

Participants who experienced base rates that were of lower utility than the cue accuracy estimated the percent of times that cue matching was correct as higher than participants who experienced base rates that were of higher utility than the cue accuracy (64%/65% vs. 60%/58%). Likewise, participants who experienced base rates that were of higher utility than the cue accuracy estimated that green was the correct answer 69% and 74% of the time. Participants who experienced base rates that were of lower utility than the cue accuracy estimated that green was correct 64% and 56% of the time.

Overall participant self-reports were fairly accurate. Participant responses to the above questions were assessed against their actual choice behavior to determine self-report accuracy. Seventy-two percent of the participants had post experimental questionnaire results that matched with their actual choice behavior. Twenty-eight percent reported a decision strategy that did not match with their choice behavior. For example, if a participant reported primarily matching the cue but their choice proportions showed that they primarily chose one color regardless of the cue, they were considered inaccurate in their self-report.

The post experimental questionnaire showed that participants were somewhat aware that utility is affecting their use of base rate information. Participant estimates of how often cue matching was correct (question 3) and how often green was the correct answer regardless of the cue (question 4) are a result of whether they experienced high or low base rate utility and not a function of consistency. However, this is a between groups comparison and 28% of the participants gave self-report descriptions that did not match their behavior. This difference indicates that some of the influence of base rate information on choice behavior may be automatic and only partially available to retrospective reports.

CHAPTER 4

GENERAL DISCUSSION

The prime focus of this study was to identify factors that influence base rate usage under direct experience. Two experiments were designed to test the effects of information consistency and utility on the usage of base rate information. The experiments confirmed the general hypothesis that base rate usage is dependent upon the statistical characteristics of the base rate information. In short, base rate usage was found to be a function of past experience.

Perhaps previous research has shown that participants underweight base rate information because participants have a long history of fluctuating base rates. The conservative decision maker, not knowing when base rates are high or low, would likely use them cautiously. Cautious usage may be a functionally adaptive bias. Although this possibility was suggested by Goodie and Fantino (1999a) and tested in a Monte Carl simulation by Goodie and Todd (in preparation) no previous empirical studies have demonstrated adaptive use of base rates based on base rate consistency and utility. Experiment 1 demonstrated that participants who experienced consistent base rates utilized base rate information more than participants who experienced inconsistent base rates. This shift in base rate usage occurred even though the utility of the base rate information was identical between conditions. While a strategy that relied solely on base rates (and neglected cue accuracies) would provide the same long-term average performance as relying solely on cue accuracies, such a strategy would provide for more

erratic short-term performance. A more cautious approach would be to use base rate information more when it appears to be a stable source of information. This strategy was observed in Experiment 1 and makes sense when one considers the error potential in relying on an erratic information source.

Experiment 2 extended these findings by demonstrating that when consistency and utility are pitted against each other, the vast majority of participants (86%) used the information source highest in utility. This finding is consistent with that of Ginossar and Trope (1987) who found that base rate usage varied as function of the credibility of the individuating information. Ginossar and Trope manipulated the source of the individuating information (palm reader vs. a psychologist) to manipulate credibility. The utility manipulation in the present study had the effect of making the base rate more or less credible than the cue accuracy. Different participants experienced different levels of utility and used the base rate information as a function of its utility.

Brunswik (1956,1957) argued that experimenters should not assume that participants would use cues as though the cues were perfect since humans evolved to cope in uncertain circumstances where the reliability and utility of information sources varied. He argued instead that researchers ought to fluctuate the utility (or ecological validity in Brunswik's terms) of cues and evaluate whether participant decisions reflect a parallel shift in utilization. This approach proved advantageous in the current study. Clearly participants demonstrated sensitivity to the statistical characteristics of the base rate information.

The results of this study provide new implications for the phenomenon base rate neglect. Most studies finding base rate neglect have presented base rates as summary

statistics. When participants do not utilize the base rates in a manner consistent with Bayes's Theorem, they are said to neglect the base rates. However, this approach often assumes that participants treat the base rate information as though it were a stable and useful information source. It is likely that participants' often experience unreliable base rate information outside of the laboratory. This study provides empirical evidence that participants will differentially utilize base rate information as a function of the consistency and utility of the information.

CHAPTER 5

REFERENCES

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44,* 211-233.

Bjorkman, M. (1984). From dualism to pluralism: The third world in judgment under uncertainty. In K.M.J. Lagerspetz & P. Niemi (Eds.), *Psychology in the 1990s*. Amsterdam: Elsevier.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.

Brunswik, E. (1957). Scope and aspects of the cognitive problem. In H. Gruber, K. R. Hammond, & R. Jessor (Eds.), *Contemporary approaches to cognition* (pp. 5-31). Cambridge: Harvard University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684-704.

Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology, 52*(3), 464-474.

Gluck, M. A, & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117*(3), 227-247.

Goodie, A. S, & Fantino, E. (1995). An experientially derived base-rate error in humans. *Psychological Science, 6*(2), 101-106.

Goodie, A. S, & Fantino, E. (1999a). Base rates versus sample accuracy: Competition for control in human matching to sample. *Journal of the Experimental Analysis of Behavior, 71*(2), 155-169.

Goodie, A. S., & Fantino, E. (1999b). What does and does not alleviate base-rate neglect under direct experience. *Journal of Behavioral Decision Making, 12,* 307-335.

Goodie,A. S., & Fantino, E. (1996). Learning to commit or avoid the base-rate error. Nature, 380, 247-249.

Goodie, A. S., & Todd, P. M. (in preparation). The ecological rationality of base rate neglect.

Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237-251.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral & Brain Sciences, 19*(1), 1-53.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). Cambridge, England: Cambridge University Press.

Lyon, D., & Slovic, P. (1976). Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologica, 40,* 287-298.

Popper, K. R. (1972). *Objective knowledge; an evolutionary approach*: Clarendon Press.

Simon, H. A. (1957). *Models of man: social and rational; mathematical essays on rational human behavior in a society setting*. New York, New York: Wiley.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 49-72). Hillsdale, NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 153-162). Cambridge: Cambridge University Press.