GIGERENZER’S ECLECTIC NORMATIVISM

ASLI SELİM

BOĞAZİÇİ UNIVERSITY

2009

GIGERENZER’S ECLECTIC NORMATIVISM

Thesis submitted to the

Institute for Graduate Studies in the Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Cognitive Science

by

Aslı Selim

Boğaziçi University

2009

Thesis Abstract

Aslı Selim, "Gigerenzer's Eclectic Normativism"

This thesis examines Gigerenzer's criticism of classical rationality and evaluates the adequacy of the ecological rationality view that he offers in its place. Classical rationality assumes that normative standards are determined by formal logic, probability theory, and decision theory. Several studies have demonstrated that people usually fail to conform to the norms of classical rationality and concluded that people are subject to various cognitive biases and fallacies. Gigerenzer rejects this view, claiming that classical rationality is not suitable for the study of human reasoning. First, I analyze Gigerenzer's criticism of the cognitive fallacy studies and the normative benchmarks of classical rationality. I argue that rational norms need not be descriptively correct and that formal logic, probability theory and decision theory should be retained as the normative benchmarks of rationality. Secondly, I discuss Gigerenzer's ecological rationality view, in which it is assumed that instead of formal logic, probability theory, and decision theory, psychologically plausible heuristics can be used for describing human reasoning and prescribing rational norms. I argue that the heuristics that have been proven to be effective and are suitable for prescription are not psychologically plausible and the ones that are psychologically plausible do not perform well consistently enough to be suitable for prescriptive purposes. I conclude that the study of psychologically plausible heuristics should be confined to the description of human behaviour and that the heuristics that are suitable for prescription need not be psychologically plausible.

Tez Özeti

Aslı Selim, “Gigerenzer’ın Seçmeci Normativizmi”

Bu tez Gigerenzer'ın klasik akılcılık normlarına karşı getirdiği eleştiriler ile klasik akılcılık normlarına alternatif olarak öne sürdüğü ekolojik akılcılık görüşünü incelemektedir. Klasik akılcılıkta normların mantık, olasılık ve karar teorilerince belirlendiği görüşü benimsenmiştir. Ancak, yapılan pek çok araştırma insanların klasik akılcılık normlarına uymadığı belirlemiş ve insanların pek çok bilişsel yanılgılara düştüğü sonucuna varmıştır. Gigerenzer insan akılcılığına yapılan bu yaklaşımı reddetmekte ve de klasik akılcılığın insan aklıyla uyuşmadığını savunmaktadır. Bu tezde ilk olarak Gigerenzer'ın bilişsel yanılgı çalışmalarına ve klasik akılcılığa karşı yaptığı eleştiriler incelenmektedir. Gigerenzer'ın aksine, akılcılık normlarının betimsel olma yükümlülüğü olmadığı ve de mantık, olasılık ve karar teorilerinin akılcılık normları olarak korunması gerektikleri savunulmaktadır. İkinci olarak ise Gigerenzer'ın psikolojik olarak makul olan buluşsalların hem insan aklını betimleme de hem de akılcılık normlarını belirlemede kullanılabileceğini öne süren ekolojik akılcılık görüşü incelenmektedir. Bu tez, akılcılık normlarını belirlemeye uygun olan başarılı buluşsalların psikolojik olarak makul olmadıklarını, psikolojik olarak makul olanların ise başarıları yeterli sürekliliği göstermediği için akılcılık normlarını tayin etmeye elverişli olmayacaklarını öne sürmektedir. Psikolojik olarak makul olan buluşsalların yalnızca insan aklını betimlenmesi yönünde kullanılması gerektiğine, akılcılık normlarını tayin edebilecek yetkinlikteki buluşsallarda ise psikolojik olarak makul olma şartının aranmaması gerektiği sonucuna varmaktadır.

CONTENTS

FIGURES

CHAPTER 1

INTRODUCTION

The classical view of rationality that is widely accepted and used in economics and psychology sets the normative standards based on the laws of probability theory, logic and decision theory. According to classical rationality, conforming to the rules that are legitimated by these theories will lead us to the decisions that are in our best interest. However, the practical applications of these rules to real life decisions prove to be quite difficult, or in some cases impossible, which lead many people to question their validity .

Some people argue that following the laws of logic, probability theory, and decision theory requires unbounded computational power, time, and resources and that in real life we almost never have the time and the energy to make consistent and optimal judgments or decisions (Over, 2004). Herbert Simon (1957, 1983) introduced the term *bounded rationality* to express the view that models of rationality should take into account the fact that people make decisions in life with limited knowledge, computational powers and in time. Simon argues that because of these limitations people strive to *satisfice* rather than optimize their decisions. He used the word *satisfice*, which is a blend of the words *suffice* and *satisfy*, to express the view that people do not consider or have access to all the alternative solutions to a problem; instead they set an aspiration level and stop their research once they find an alternative that satisfies it. Thus, in the bounded rationality approach the emphasis is on making decisions fast and with limited resources to reach good-enough results rather than achieving the optimal one.

Following Simon's theory of bounded rationality, it has been suggested that people do not make decisions according to the rules set by logic or probability theory, but that they rely on *heuristics*, or rules-of-thumb. Picking up on this idea, two psychologists, Amos Tversky and Daniel Kahneman launched the heuristics-and-biases research program which aims to discover what heuristics people use for making decisions and what sorts of errors (biases) they can lead to. Heuristics-and-biases program assumes that logic, probability theory, and decision theory provide the normative benchmarks and that the decisions that deviate from the norms that they appoint are irrational. They tested the descriptive accuracy of economics theories about human rationality, and their research revealed that people make many reasoning errors and rarely follow the rules of rational decision making (Kahneman, Slovic, & Tversky, 1982).

The results of their research made a huge impact on economics and psychology and led the view that humans are irrational and heuristics lead to wrong decisions to be widely accepted. However, some researchers reject the findings of Kahneman and Tversky and argue that not only humans are rational, but also the heuristics which they rely on are very efficient and lead to successful decisions. Psychologist Gerd Gigerenzer and the Adaptive Behavior and Cognition (ABC) Group launched a research program called "bounded rationality" dedicated to the study of heuristics. Gigerenzer and his colleagues argue that Kahneman and Tversky merely studied the discrepancies between the norms and human judgment and labeled these discrepancies as fallacies (e.g., Gigerenzer (2004). They criticize Kahneman and Tversky for putting “the blame … on the human mind rather than on the norm” (Gigerenzer & Selten, 2001b, p.5) and claim that setting unrealistic standards for rationality is mistaken. Furthermore, they assert that heuristics-and-biases program misinterpreted Simon’s concept of bounded rationality. They maintain that Simon did not use the term bounded rationality to refer to people’s flawed decision making skills, but to emphasize the importance of the match between the environment and the decision maker and cite the following quote from Simon(1990): “Human rational behavior … is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor” (p.7). Taking on this view, the proponents of the bounded rationality program adopt the idea that human rationality cannot be studied without paying attention to the environment because “Studying only one blade is not enough; it takes both for the scissors to cut” (Gigerenzer & Selten, 2001b, p.4).

Bounded rationality program adopts the evolutionary perspective on the study of cognition; this research program is built on the idea that human cognition is made up of heuristics that have evolved to exploit the specific structure of the environment. By combining this evolutionary perspective with Simon’s bounded rationality, the proponents of the bounded rationality program conceived a new kind of rationality for their research: *ecological rationality.* They claim that rationality of a heuristic cannot be evaluated by comparing it to the dictation of probability calculus, logic or decision theory, but by studying how successfully it has adapted to the structure of an environment (Gigerenzer, Todd & the ABC Research Group, 1999; Gigerenzer & Selten, 2001b). With this definition in mind, Gigerenzer (2008) reevaluates the studies of the heuristics-and-biases research and claims that when the requirements of the structure of the environment are taken into consideration "many so-called cognitive illusions largely disappear" (p.18).

Bounded rationality program aims to build a descriptively accurate and realistic model of human cognition. The proponents of this program have two objections to the current models of rationality. First, they state that these models are lack any practical value. For example, Gigerenzer (2004) cites the following anecdote to point out impracticality of the expected utility theory: [1]

> A decision theorist from Columbia University was struggling whether to accept an offer from a rival university or to stay. His colleague took him aside and said, "Just maximize your expected utility – you always write about doing this." Exasperated, the decision theorist responded, "Come on, this is serious." (p.62)

Although expected utility theory is an eloquent theory, its application to real life decisions can be quite difficult because it is not easy to accurately predict the probability of events or their outcomes. Furthermore, in cases where "incommensurable goals exist" (Gigerenzer, 2001a, p.40) it is not possible to express the desirability or the "goodness"

[1] Decision theory posits that the decision that would assure maximization of goal satisfaction can be reached by calculating expected utilities. Expected utility theory appoints the normative choice by comparing the alternatives in terms of the desirability of the outcomes and their consequences, and the probability of their occurrences. For example, let us say there is a bet based on the roll of a regular die: you get 10$ if the outcome is "6", and 5$ if it is "not 6". Which one should you bet on? Expected utility theory tells us that we should compute the expected utility of each option and pick the one with the highest expected utility. In this case, the probability of "6" is 1/6, and the utility of betting on it is 10$. Hence the expected utility of betting on "6" is 10* 1/6= 10/6. Similarly, the expected utility of betting on "not 6" is 5/6*5= 25/6. Since the expected utility of betting on "not 6" is higher than that of betting on "6", a rational decision maker should bet on "not 6."

of an outcome in numbers. To illustrate how the process of assigning utilities can be difficult, Gigerenzer (2008) gives the following example:

> A few years after his voyage on the Beagle, the 29-year-old Charles Darwin divided a scrap of paper (titled "This Is the Question") into two columns with the headings "Marry" and "Not Marry" and listed supporting reasons for each of the two possible courses of action, such as "nice soft wife on a sofa with good fire" opposed to "conversation of clever men at clubs." Darwin concluded that he should marry. (p.30)

He contends that expected utility theory would be of no use to Darwin, because it would not be possible for him to assign utilities to having a nice wife or a stimulating conversation with friends. Most things in life do not come with price tags, and thus, he concludes, expected utility theory is inadequate for solving real life problems (Gigerenzer, 2001a).

Fast-and-frugal Heuristics

The second objection that the researchers from the bounded rationality program have against optimization methods and the classical theories of rationality is that these theories "treat the mind as a Laplacean superintelligence equipped with unlimited resources of time, information, and computational might" (Gigerenzer, 2001a, p.37). Their aim is to dispense with the 'Laplacean demon view,' and embrace real world problems with all the constraints on time, knowledge, and computational capacity associated with them. They declare that optimization is a fiction, that it is "often based on uncertain assumptions" (Gigerenzer & Selten, 2001b, p.4) and in most cases "is computationally intractable in any implementation" (Gigerenzer, 2008, p.82). [2] What distinguishes the bounded rationality program from other research programs on heuristics is that it claims that heuristics perform as well as the optimization algorithms. Thus, they are taking the idea of heuristics one step forward by using them not only for descriptive purposes, but also for prescribing decisions. Gigerenzer(2008) often gives the following example to explain how heuristics can be quite successful thanks to their ability to exploit the structure of the environment:

> Imagine you want to build a robot that can catch balls – fly balls, as in baseball and cricket. (It's a thought experiment – no such robots exist yet.) For the sake of simplicity, consider situations where a ball is already high up in the air and will land in front of or behind the player. How would you build such a robot? One vision is omniscience: you aim at giving your robot a complete representation of its environment and the most sophisticated computational machinery. First, you might feed your robot the family of parabolas, because, in theory, balls have parabolic trajectories. In order to select the right parabola, the robot needs to be equipped with instruments that can measure the ball's initial distance, initial velocity, and projection angle. Yet in the real world, balls do not fly in parabolas, due to air resistance, wind, and spin. Thus, the robot would need further instruments that can measure the speed and direction of the wind at each point of the ball's flight, in order to compute the resulting path and the point where the ball will land, and to then run there. All this would have to be completed within a few seconds the time a ball is in the air. An alternative vision exists, which does not aim at complete representation and information. It poses the question: Is there a smart heuristic that can solve the problem? One way to discover heuristics is to study experienced players. Experimental studies have shown that players actually use several heuristics. One of these is the gaze heuristic. When a fly ball approaches, the player fixates the ball and starts running. The heuristic is to adjust the running speed so that the angle of gaze remains constant (or within a certain range; see McLeod & Dienes, 1996). (p.30)

[2] Computational intractability (also referred to as NP-hard) means that an algorithm solves the problem in polynomial-time. This means that as the size of the problem increases, the time required for the algorithm to solve the problem increases very quickly. Thus, when the problem size is big, the solution becomes impossible to compute in a reasonable amount of time. Gigerenzer claims that Bayesian networks developed by Judea Pearl (2000) is not a reasonable model of human cognition because it is NP-hard (Gigerenzer, 2008).

Gigerenzer believes that human cognition can be explained with heuristics which are fast, simple and effective just like the gaze heuristic in the example. Bounded rationality envisions the human mind as an *adaptive toolbox* : "the collection of specialized cognitive mechanisms that evolution has built into the human mind for specific domains of inference and reasoning, including fast and frugal heuristics" (Gigerenzer & Todd, 1999, p.30). Instead of a "general-purpose-decision-making-algorithm," such as the expected utility theory, the adaptive toolbox solves problems by using a variety of simple heuristics that are adapted to specific environments. The common characteristics of heuristics are that they are all "fast-and-frugal," meaning they are computationally cheap, fast and require little information. They are successful because they "are able to exploit the structure of environments" (Gigerenzer,2008, p.90).

Although few people object to the psychological reality of heuristics, most people are skeptic about their success. For example, in their evaluation of a heuristic which is observed to be widely used in practice, Keeney and Raiffa (1993) acknowledge that it offers "administrative ease," but also note that "it is rarely appropriate," claiming it "will rarely pass a test of 'reasonableness'" (p.78). [3] To prove that the study of ecological rationality can be prescriptive as well as descriptive, researchers from the bounded rationality program have carried out several studies to assess the effectiveness of heuristics compared to complex optimization methods that are commonly used in decision analysis, such as multiple regression and Bayesian networks (e.g., Gigerenzer

[3] Interestingly, Howard Raiffa is 'the decision theorist from Columbia University' that Gigerenzer talks about in the previous example. Raiffa (2002) states that this rumor is not true and denies that such a conversation has taken place: "Nope, this is not true. I never said such a thing. Actually, my wife and I subjected this decision to a primitive, multiple-value analysis involving 10 objectives" (p.180).

& Goldstein, 1996; Martignon, 2001) These studies produced quite surprising results that showed that simple heuristics can actually be very effective, leading many people to change their minds about heuristics.

Recognition heuristic

Recognition heuristic is one of the first heuristics that the researchers from the bounded rationality program have come up with. Goldstein and Gigerenzer (1999) carried out an experiment in which they asked the undergraduates at the University of Chicago "Which city has more inhabitants: San Diego or San Antonio?" Only two thirds of the students could answer this question correctly. However, when this study was replicated in Germany, all of the German students who participated in the study gave the correct answer "San Diego." Although the Americans are more knowledgeable about these two cities, more Germans answered the question correctly (Goldstein & Gigerenzer, 1999). Goldstein and Gigerenzer concluded that the German students answered this question using the fast-and-frugal recognition heuristic. The German students did not know much about either of the cities; however, they had previously heard about San Diego but not about San Antonio. Since one of the cities was familiar to them and the other one was not, they inferred that San Diego must be a bigger city than San Antonio. Goldstein and Gigerenzer(2002) formulated the definition of recognition heuristic as follows : "If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion" (p.76). They claim that this heuristic was not available to the Americans because simply "they know too much" (Goldstein & Gigerenzer, 2002, p.76).

In another study, 50 Turkish students who had very little knowledge or interest in English soccer teams and 54 British students who were quite knowledgeable on the subject were asked to forecast the outcomes of the 32 English FA Cup third-round soccer matches (Ayton and Önkal,1997). The Turkish students predicted almost as accurately as the British students did; 63% of their forecasts were correct, whereas British students got 65% of their forecasts correct. Goldstein and Gigerenzer (1999) state that the Turkish students were successful, because they could rely on recognition heuristic thanks to their limited knowledge about the British football league. In 95% of the cases where they had to choose between a team that they had heard of before and an unfamiliar one, they chose the team that they recognized. Thus, the Turkish students were advantageous because they knew less, just like the German students in the previous study; Goldstein and Gigerenzer(1999) named this "the less-is-more effect." They concluded that the ecological rationality for this heuristic is that, it is successful when recognition is strongly correlated with the criterion (Goldstein & Gigerenzer, 2002).

Goldstein and Gigerenzer (2002) also add that the recognition heuristic reflects an evolved capacity. They state that the studies on memory have revealed that the capacity for recognition is distinct from that for recalling. For example, they point out that people are capable of recognizing more items than they can recall and that the people who have lost their ability to recall memories retain their capacity for recognition

(Warrington & McCarthy, 1988). Furthermore, they contend that recognition is also commonly used in the animal kingdom; for instance, rats rely on a recognition mechanism to decide whether a food is poisonous or not (Galef, 1987). In light of these findings, they conclude that recognition is a "primordial psychological mechanism" (Goldstein & Gigerenzer, 2002, p.77).

Borges, Goldstein, Ortman and Gigerenzer (1999) tested the success of the recognition heuristic by applying it to the stock market. They chose this medium because "Financial markets are notoriously unpredictable" (Borges et al.,1999, p.60) and the large number of technical trading models and expertise would make up good rivals for testing the performance of the recognition heuristic. They asked German and American pedestrians (laypeople) and stock market experts to pick the companies that they recognized in both the German and the American stock markets and constructed investment portfolios based on these surveys. Their results showed that the investment portfolio based on the recognition of the German laypeople was the one which performed most successfully in the American stock market, followed by the portfolios based on the recognition of the German experts, American laypeople and American experts. Similarly, the portfolio based on the recognition of the American laypeople proved to be the most profitable one in the German stock exchange, followed by those of the American experts, German lay people and German experts. In other words, the Germans outperformed the Americans in the American stock markets, whereas the Americans outperformed the Germans in the German stock market. Moreover, in all of the cases the portfolios based on the recognition of the laypeople outperformed those that are based on the recognition of the experts. In short, people who knew less predicted better with the recognition heuristic. Borges et al. also state that the portfolios based on recognition outperformed stock market indices and mutual funds as well. For example, the portfolio based on the recognition of the American laypeople outperformed the German stock market by 23%. Borges et al. interpret the result of their study as proof that recognition heuristic can help people make successful inferences in real-world domains.

Take-the-Best

Take-the-Best is probably the most successful, well-renowned and controversial heuristic offered by the bounded rationality program so far. Take-the-Best promotes the idea that making decisions based on only one reason can be as successful as (or sometimes even more successful than) the decisions that are made by considering all the relevant criteria (Gigerenzer & Goldstein, 1996). It simply states that when you are trying to make an inference between two alternatives, order your criteria based on their validities and compare the alternatives on the basis of these criteria starting from the top of the list. The moment a criterion differentiates between the two alternatives, make your decision and stop. Stated more formally:

> Take-the-Best
>
> 1. Search by validity: Search through cues in order of their validity Look up the cue values of the cue with the highest validity first.
>
> 2. One-reason stopping rule: If one object has a positive cue value (1) and the other does not (0 or unknown), then stop search and proceed to Step 3. Otherwise exclude this cue and return to Step 1. If no more cues are found, guess.
>
> 3. One-reason decision making: Predict that the object with the positive cue value (1) has the higher value on the criterion.
>
> The validity of a cue *i* is defined as $v_i = R_i / P_i$ where R_i = number of correct predictions by cue *i* and P_i = number of pairs where the values of cue *i* differ between objects. (Gigerenzer, 2008, p.32)

For example, imagine that you are going to buy a new house and you need to make a decision among two options. You might have several criteria in your mind that would affect your decision, such as the proximity of the house to your work, its price, size and so forth. Instead of assessing the alternatives by all of these criteria, Take-the-Best tells you to rank your criteria based on their validities. Let us say that price is the most important criterion for you, followed by the proximity to your work, and its size. You first compare the houses based on their prices. If they are both expensive, you conclude that this criterion will not be useful and skip to the next criterion. If you find out that one of the alternatives is very close to your work whereas the other one is quite far, you stop your decision process and infer that you should buy the alternative that is close to your work. Take-the-Best is frugal, because it does not require an exhaustive information search among all the criteria or a complex computation to integrate all the available information.

Although common sense tells us that such a simplistic method cannot actually lead to good decisions, Goldstein and Gigerenzer (1996) showed with computer simulations that Take-the-Best can be quite a powerful tool for inference. They devised a two-alternative choice task environment in which they compared the performance of the Take-the-Best with those of complicated optimization algorithms, such as multiple regression and weighted linear model. The task was to answer two-alternative choice problem: "Which city has a larger population? (a) Hamburg (b) Cologne" (Goldstein & Gigerenzer, 1996, p. 651). They suggested nine criteria (cues) that can help in this inference problem, such as whether the city is a state capital, was once home to an exposition site or has a football team in the German football league Bundesliga. The results of the simulations showed that the Take-the-Best was the fastest algorithm, and it drew as many correct inferences as one of the competitors and outperformed all the others (Gigerenzer & Goldstein, 1996).They report that two other variations of the Take-the-Best algorithm, namely the Take-the-Last heuristic and the Minimalist heuristic, also performed remarkably well. [4]

Take-the-Best algorithm has a wide area of applications. For example, it has been used to design a classification heuristic tree to help the doctors at the University of Michigan Hospital decide which patients to send to the coronary care unit (see Figure 1). The doctors simply have to follow the steps of the tree and stop once a decision is made. For instance, if a patient has an anomaly in her electro diagram (depicted in the ST

[4] Instead of a static cue ranking like the Take-the-Best, Take-the-Last selects the criterion that discriminated between the alternatives in the last question. If that criterion does not discriminate between the current alternatives, then it tries the one that worked before that. The Minimalist algorithm picks the cue (i.e. criterion)for inference randomly (Goldstein & Gigerenzer,1996).

segment branch) then he should be sent to the case unit right away. If this is not the case, then the doctor should check if the patient has any chest pains. If he does not, then the doctor should conclude that the patient runs a low risk of developing a heart attack (Green & Mehr, 1997). According to Green and Mehr, this fast and frugal tree proved to be more accurate than the physicians' decisions and the expert system which carried out complex computations over some 50 probabilities and a logistic regression.

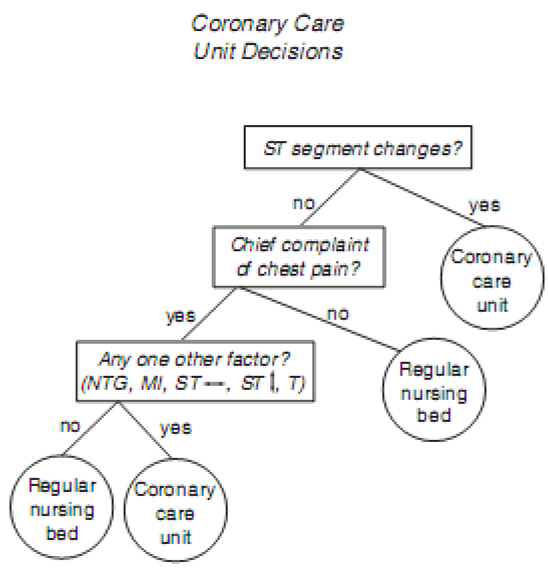

Figure 1. Application of the Take-the-Best to coronary care unit allocation problem (Gigerenzer, 2008, p.43).

Intelligence of the Gut Feeling

Gigerenzer (2007) asserts that heuristics are powerful tools for decision making and complains that in most situations people are discouraged from following their gut feelings. For example, he states that his studies on the police officers working at the Los Angeles airport have revealed that through years of experience they develop an expertise in detecting drug couriers, and that these officers successfully detect drug couriers by following their hunches. When the police officers find illegal drugs on the person, in court they are required to explain why they stopped to search him in the first place. However, since police officers are unable provide a reason why they had such a hunch, and gut feelings do not constitute admissible proof in courts police officers usually refrain from relying on their guts. Gigerenzer states that doctors also complain that they are afraid to follow their gut feelings out of fear of getting sued and claims that the intelligence of gut feelings should be officially acknowledged.

Cognitive Illusions

I will briefly summarize Gigerenzer’s analyses of the cognitive fallacy studies from the heuristics-and-biases program. This discussion will reveal why he is skeptical about the classical rationality and give clues about how the theory of fast-and-frugal heuristics were conceived.

Conjunction fallacy

Conjunction fallacy refers to the error of assigning a higher probability to a specific event than the general event that contains it. Let A and B be two events. Since the set A&B is a subset of A, it follows that $p(A\&B) \leq p(A)$. Tversky and Kahneman (1983) demonstrated that people's probability judgments violate the conjunction rule. The problem that they used in their study is as follows:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and social justice and also participated in anti-nuclear demonstrations.
>
> Which of two alternatives is more probable:
>
> Linda is a bank teller (T)
>
> Linda is a bank teller and is active in the feminist movement (T&F)
> (Tversky &Kahneman, 1983, p.299)

Since the second alternative is a subset of the first alternative, the probability of the first alternative is higher than the second one. However, their study revealed that 85% of the participants judged the probability of Linda being a bank teller and an active feminist to be more likely than her being a bank teller. Tversky and Kahneman named this reasoning error as the conjunction fallacy.

Gigerenzer (1994, 2000) objects to the results of this study on grounds that the normative answer to this question is disputable. He states that there are several different interpretations of probability and that for some of them judging that T&F is more probable than T in this question would not constitute a violation of the probability theory.

He explains that frequentists define probabilities for long-run frequencies and not for single events or outcomes: “ For a frequentist like the mathematician Richard von Mises, the term ‘probability,’ when it refers to a *single event*, ‘has no meaning at all for us’” (Gigerenzer,2000, p. 246). Gigerenzer asserts that since Linda being a bank teller poses a single-event probability, von Mises would argue that assigning a probability to it would be meaningless. When a probability cannot be defined, one cannot impose conjunction rule to it either. Furthermore, he states the statistician George Alfred Barnard is also of the opinion that a person’s subjective probability assignments cannot be objected to the rules of probability. Thus, if von Mises’ or Barnard’s view is adopted as Gigerenzer suggests, then the violation of the conjunction rule cannot be labeled as a fallacy.

In lines with this view, Gigerenzer(2000) changes phrasing of the question and asks the participants to make frequency assessments instead of single-event probabilities:

> There are 100 persons who fit the description above (i.e., Linda’s). How many of them are :

(a) Bank tellers
(b) Bank tellers and active in the feminist movement (p.250):

He reports that when the problem is presented this way, the percentage of conjunction violations drops from 80-90% to 10-20%.

Gigerenzer (2001b) also criticizes Tversky and Kahneman for applying the rules of probability and logic “content-blindly”; he claims that rules of logic and probability cannot be imposed without paying attention to the context of the problem. For example, he asserts that the ambiguity of the words *probable* and *and* in the Linda problem creates confusion and leads to different interpretations of the statements. He explains that the word *probable* in English does not only refer to a mathematical possibility, but is also used express that something is "conceivable." Similarly, he points out that the word *and* is not used or understood in daily language in the same manner as it is used in logic: "noun-noun phrases often refer to the disjunction, not the conjunction, of two elements or classes. For instance, the announcement "We invited friends and colleagues" does not refer to the intersection between the two groups, but to the joint set of both groups" (Gigerenzer, 2008, p.72). He maintains that this analysis is supported by the study in which rephrasing the question as "bank teller as well as active feminists" caused the conjunction fallacy to largely disappear (Mellers, Hertwig, & Kahneman, 2001).

Base-rate fallacy

Base rate fallacy refers to the error that stems from neglecting the prior probability of a hypothesis when assessing the conditional probability of hypothesis based on some evidence. This fallacy has been studied with the problems of the following type:

> If a test to detect a disease whose prevalence is 1/1000 has false positive rate of 5%, what is the chance that a person found to have a positive result actually has

the disease, assuming you know nothing about the person's symptoms or signs? (Tversky & Kahneman, 1982, p.154)[5]

When people are answering this question, they typically ignore the low base-rate of the disease (1/1000) and base their probability assignment only on the reliability of the test. These types of conditional probability problems are solved using Bayes' theorem. Bayes's theorem can be stated as follows:

$$p(H \| E) = \frac{p(H)p(E \| H)}{p(H)p(E \| H) + p(-H)p(E \| -H)}$$

Where *H* stands for hypothesis being true, *-H* for hypothesis being false and *E* for evidence. For instance, for the problem given above *p(H)* stands for the probability of the patient having the disease, *p(H/E)* stands for the probability of the patient having the disease given that he tested positive, *p(E/H)* stands for the probability of the patient testing positive given that he actually has the disease and so forth. When the numbers are inserted into the formula, we get the following result:

$$p(H \| E) = \frac{(0.001)(1)}{(0.001)(1) + (0.999)(0.05)} \approx 0.02$$

Thus, even if a patient has tested positive for the disease, the probability of him actually having the disease is as low as 2%. However, when this question was asked to the staff and the students at Harvard Medical School, almost half of the participants assigned 0.95 to the probability of the patient actually having the disease, and only 18% of the participants could come up with the answer 0.02. The researchers from the heuristics-

[5] It is assumed that if a person has the disease then the test accurately detects it 100% of the time (Tversky & Kahneman, 1982, p.154).

and-biases program interpreted the results of this study as evidence that people are bad at Bayesian reasoning.

Gigerenzer has two objections to this study. First, he claims that there is not a single normative answer to the question. Similar to his discussion of the conjunction fallacy, he claims that this question posits a single-event probability judgment, which is assumed to be meaningless by the frequentists. Thus, he argues, since probability is not defined for this event, Bayes' theorem cannot be applied either (Gigerenzer, 2000, p.252).

The second objection that Gigerenzer (2000, chap. 6) concerns the information format of the problem. He claims that people cannot solve this problem, because there is a mismatch between the information format (structure of the environment) and the mental algorithm humans use for Bayesian reasoning. He argues that our statistical reasoning skills have not evolved to work with percentages or other types of normalized probabilities and therefore when the information is presented in these formats the algorithm fails to work. He likens the working of our statistical reasoning skills to that of a pocket calculator. He explains that our pocket calculators can function correctly only if the input is in decimals, that is, base-10. If we enter the input in binary numerals, the calculator would not be able carry out the correct calculation. Like the pocket calculator which fails to function properly when the input is entered in the wrong format, he claims that we cannot reason well when the information format used to express probabilities is incompatible with our mental processes. He states that the normalized probability formats, such as percentages, have entered our lives very recently and were not present in the environment that we have evolved in. Instead, he contends that our statistical reasoning has evolved to operate on *natural frequencies*, which are "raw" counts of events. For example, the statement "of the 8 people I have met today, only 2 wore glasses" uses natural frequencies, whereas in standard probability format the same statement would be expressed as "only 25% of the people I have met today wore glasses." Gigerenzer believes that our mental processes have adapted to natural frequencies, because they are compatible with the information format that *natural sampling* yields (Gigerenzer, 2000, p.94-96). Natural sampling is defined as the "sequential process of updating event frequencies from experience" (Gigerenzer, 2000, p.63); it is the mechanism that we use to acquire statistical information. The term "natural" has been picked to highlight the idea that natural sampling occurs through direct observations in a natural environment.

Gigerenzer (2000) states natural sampling is the means with which a foraging animal in the wilderness acquires statistical information about its environment (p.62). To be able to predict where food resources might be, the animal would need to learn the cues (such as the existence of other foraging animals around the site) that indicate the existence of food at a certain site. He claims that the animal would learn the predictive power of a cue by updating the frequency with which it encounters this cue near a food resource. Thus, each time this animal visits, that is samples, a potential food resource, it will make note of whether a certain cue is present. If the cue is present, then it will increase its count by one. Another example of natural sampling could be, he states, a physician from an illiterate society who does not have any access to medicine books or statistical research. This physician would rely on her encounters with past patients to have an understanding of the rate of occurrence of a disease and its symptoms (Gigerenzer, 2000, p.62).

Gigerenzer (1994) claims that experimental studies support the natural sampling theory. For example, he calls our attention to the fact that children are capable of counting events starting from a young age; however an understanding of fractions and "other types of normalized counts" are acquired only in the much later stages of their development (Dehaene, 1997). He reports that a number of studies have demonstrated that people can keep track of frequencies quite accurately (Barsalou & Ross, 1986; Hintzman &Black, 1972; Jonides & Jones, 1992) and that Hashner and Zacks's (1979) experiments revealed that "frequencies are one of the few kinds of information (the others being word meaning and spatial and temporal location) that are monitored *automatically*- that is, without intention or much attention, and without interfering with other tasks" (Gigerenzer,1994, p.139). He claims that "what is now called *automatic frequency processing* seems to be generally accurate" (Gigerenzer,1994, p.139).

Gigerenzer (2000) states that during the Enlightment, mathematicians like Laplace and Condorcet had declared that probability theory was "the common sense of educated people, who were known then as 'hommes éclairés'" (p.92). Then, with the onset of heuristics-and-biases program this idea was replaced with the view that human reasoning simply cannot comply with the dictates of probability theory or logic because of their "limited information-processing abilities" (Lichtenstein, Fischhoff, Phillips, 1982, p.333). However, with the theory of natural sampling and automatic frequency processing he claims that “the reasonable man is back … less élite (everyone is a reasonable intuitive statistician)” (Gigerenzer, 1994, p.139).

Gigerenzer (2000) asserts that changing the information format in the base-rate problems to natural frequencies drastically improves people’s Bayesian reasoning performance. To test this theory, Gigerenzer and Ulrich Hoffrage (1995) asked their subjects to answer a classical problem used in the base rate fallacy studies presented in three different information formats. In standard normalized probability formats the question was expressed as follows:

> The probability of breast cancer is 1% for a woman at age forty who participates in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.69% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer? (Gigerenzer & Hoffrage, 1995, p.688)

The question expressed in natural frequencies is:

> 10 out of every 1,000 women at age forty who participate in routine screening have breast cancer. 8 out of every 10 women with breast cancer will get a positive mammogram. 95 out of every 990 women without breast cancer will also get a positive mammogram. Here is a new representative sample of women at age forty who got a positive mammogram in routine screening. How many of these women do you expect to actually have breast cancer? (Gigerenzer & Hoffrage, 1995, p. 693)

Gigerenzer and Hoffrage (1995) report that when the problem was stated in standard probability format, on average 24% of the participants gave the correct answer, whereas for the natural frequencies this percentage raised 48%. Gigerenzer (2000) interprets this result as proof that natural frequencies facilitate Bayesian reasoning; he claims that natural frequencies make Bayesian inference simpler because base rates are "implicit in the frequencies" (p.63) and the Bayes's rule becomes computationally simpler.[6]

Overconfidence Bias

Overconfidence bias is one of the most intensively investigated biases in the decision making literature. Overconfidence bias refers to "an individual's overvaluation of her or his own skills, knowledge, or judgment" (Sternberg, p.453). In a typical experiment setting for overconfidence bias, the participants are asked several general-knowledge questions such as the following (Lichtenstein, Fischhoff & Phillips, 1980, p. 307) :

---

[6] To demonstrate how natural frequency format would facilitate Bayesian reasoning, Gigerenzer (2000, chap. 6) rephrases the mammogram problem by setting the story around the physician from the illiterate society. A new disease has stricken her people, and she needs to refine her diagnosis skills based exclusively on her experience with past patients. She realizes that a certain symptom is a forerunner of the disease, but not in every case. Of the 1,000 patients she has treated during her life, only 10 actually had the disease. 8 of these patients with the disease had the symptoms, but 95 of the 990 people who did not have the disease also showed the symptom. Therefore, although she had encountered a total of 8+95=103 people with the symptom, only 8 of them actually had the disease. Gigerenzer argues that to assess the probability of patient with the symptom actually having the disease, she would simply need the number of cases where both the symptom and the disease were present (in this case, 8), and the number of cases where the symptom present without the disease (95). This physician could avoid the complex Bayesian computation by simply carrying out the following calculation (Gigerenzer, 2000, p.98):

$$p(H \backslash D) = \frac{d\&h}{d\&h + d\&-h} = \frac{8}{8+95}$$

Gigerenzer (2000) states that this equation is the formula of "Bayes' rule for natural frequencies", where *d&h* stands for the number of cases where both the disease and the symptom were present, and *d&-h* for the number of cases where the patients showed the symptom, without actually having the disease (p.62) .

What is absinthe?

a) a precious stone
b) a liqueur

Participants are asked to select the correct answer and afterwards to state the probability (on a 0 to 1 scale) that their answers were correct. The experimenter evaluates the results by grouping together the answers with similar probability assessments and computing the percentage of correct answers within each group. If the mean of the probabilities assigned by the participant exceeds the proportion of her correct answers, then her assessments are said to be "badly calibrated" reflecting "overconfidence" (Lichtenstein et al., 1980, p.308). Most of the studies on the overconfidence bias report that the participants' probability assessments, that is, their confidences are higher than their accuracy of their answers. For example, when the participants state their confidence level to be 90%, only 75% of their answers were actually correct. (Kahneman & Tversky, 1996)

Gigerenzer (2000, chap 12) claims that this experiment does not show that people are overconfident by arguing that this bias is merely a result of confusion over the norms. He argues that when people are rating their levels of confidence in having given the correct answer on a question, they are actually expressing their subjective belief in a single-event probability. However, the measure of their accuracy, which is taken to be the proportion of their correct answers, is a relative frequency. Once again Gigerenzer points out that according to frequentism single event probabilities are meaningless and hence comparing single-event probabilities with relative frequencies is like "comparing apples with oranges" (Gigerenzer, 2000, p.246) . Consequently, he argues, the relative number of correct answers the subjects have cannot constitute a ‘norm’ to which the subjects’ confidence rates can be compared.

To distinguish between the single-event confidences and the frequencies, Gigerenzer, Hoffrage and Kleinbölting (1991) repeated the experiment in such a way that the subjects were asked to estimate the relative frequency of their correct answers. For example, instead of asking the subjects how confident they are in their answers, which would be a single-event probability, they asked them to estimate the relative frequency of the questions that they think they got right. Gigerenzer(2000) reports that when their subjects express their confidences as relative frequencies their estimates predict the actual relative frequency of their correct answers quite accurately and that overconfidence bias “disappears” (p.150). He interprets this as proof that what has been recognized as a cognitive illusion in the psychology literature was simply the outcome of a confusion over probability norms.

Probabilistic Mental Model Theory

Probabilistic Mental Model (PMM) Theory was developed by Gigerenzer et al. (1991) to explain the mental processes that give rise to the overconfidence effect. Although initially this theory dealt only with modeling confidence, it later served as the basis for the fast-and-frugal heuristics such as the Take-the-Best .

Gigerenzer et al.(1991) claim that when a participant is asked to answer a two-alternative general-knowledge question, as is typical in the overconfidence questionnaires, she will first try to construct a “local mental model (local MM) of the task” (p.506). Local MM involves consulting the long-term memory for the available information and using elementary logical operations to make inference on them. If the necessary information is not available in the long-term memory, then a PMM is constructed that will make inductive inference based on “probabilistic information from a natural environment” (Gigerenzer et al., 1991, p.506).

Gigerenzer et al. explain that the construction of a local MM is based on the following three factors : “(a) precise figures can be retrieved from memory for both alternatives, (b) intervals that do not overlap can be retrieved, or (c) elementary logical operations, such as the method of exclusion, can compensate for the missing knowledge” (p.507).

For example, if the general-knowledge question asks the participant which city has the bigger population, Bonn or Heidelberg, the participant will first consult her memory. In case she knows from memory that the population of Heidelberg is between 100,000 and 200,000 and that the population of Bonn exceeds 290,000, then without constructing a PMM, the participant will select Bonn and assign a confidence level of 100%. As can be seen, in this example the participant’s local MM satisfies all three requirements listed above. She could retrieve precise figures or intervals from her memory for both of the alternatives: 100,000 – 200,000 for Heidelberg and 290,000 for Bonn. The figures that she recalled did not specify overlapping intervals, for 290,000 is bigger than the highest figure she put for the population of Heidelberg, that allowed for a simple comparison between the populations and judgment (See Figure 2).

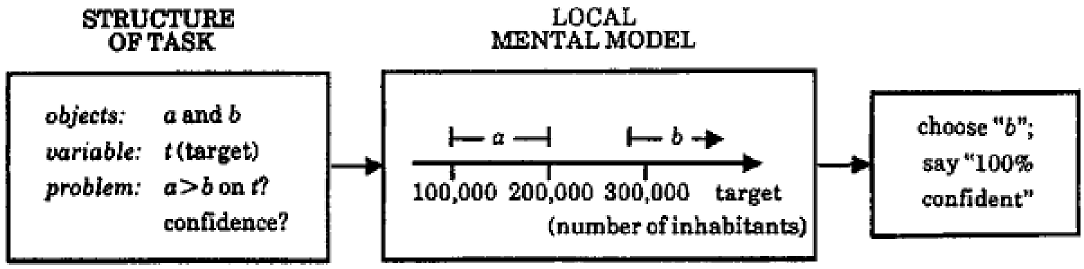

Figure 2. Local MM for the population question (Gigerenzer et al., 1991, p.508**).**

The second example they provide illustrates how elementary logical operations could be of help to compensate for the lack of information. This time the participant is asked if when sees the nationality letter "P" on a car whether she would infer that the car is from Poland or Portugal. She might know what abbreviation is used for Poland, (PL), but not for Portugal. However, this would be sufficient for her to deduce that the answer to the question is Portugal by means of exclusion, since she knows that Poland has a different nationality letter (Allwood & Montgomery, 1987).

Gigerenzer et al.(1991) note that the outcome of a local MM is "evaluated as certain" and thus the participant states that she is 100% confident in her choice. They are aware of the fact that memory is susceptible to failure and claim that these failures are one of the reasons behind the overconfidence observed in the questions that the participant assigns 100% confidence.

For the cases where a local MM cannot be constructed, Gigerenzer et al. (1991) claim that a PMM is used to solve the task. They emphasize that PMM's are sensitive to the structure of the task and they use a probability structure that matches well to the natural environment. For example, the natural environment corresponding to the previous example is the class of all cities in Germany, and the target of the question, which was the size of the population, is set as the variable. They distinguish a PMM from a local MM by noting that it requires " a *reference class* of objects that includes the objects *a* and *b* …. [and] it uses a network of variables in addition to the target variable for indirect inference" (Gigerenzer et al., 1991, p.507).

The reference class is the key to the functioning of the PMM, as it is the determinant factor in deciding which probability cues should be used and assessing their cue validities. For the population question mentioned above, Gigerenzer et al.(1991) suggest that "all cities in Germany" could be taken as the reference class and a corresponding probability cue could be existence of a soccer team in the German soccer league Bundesliga. They claim that this cue could be useful because a city with a big population is more likely to have its soccer team compete in the Bundesliga. Therefore, when two cities are compared, if one of them has a soccer team in the Bundesliga and the other one does not, then this would indicate that the first one probably has the larger population. They regard a variable C_i as a probability cue, if the alternatives *a* and *b* having different values on this cue leads to a change in the probability $p(a)$ of *a* being correct. Stated more formally :

$$p(a) \neq p(a|aC_ib; R)$$

where aC_ib denotes the relation of a and b on the cue $C_{i.}$. They define the cue validity of the cue C_i in the reference class R as $p(a|aC_ib; R)$ (Gigerenzer et al., 1991, p.508).

Cue validities are a part of a person's PMM, reflecting her own knowledge of the environment and so forth. The *ecological validity*, on the other hand, is "the true relative frequency of any city having more inhabitants than any other one in R if aC_ib" (Gigerenzer et al.,1991, p.508). For example, they state that the ecological validity of the cue 'having a team in the Bundesliga' is 0.91, if the reference class is taken to the be the set of all German cities with population larger than 100,000. They explain that this information reflects the fact that 91% of the comparisons between all possible pairs of cities in which one city has a soccer team in the Bundesliga and the other one does not, yield the result that the city with a team in the Bundesliga has the larger population.

Cues can be activated, that is used for inference, only if the value on the cue is known for both of the alternatives and these values are different. For example, when comparing two cities, if one knows that one of the cities has or does not have a team in Bundesliga, but does not know such information about the soccer team of the other city, she cannot use this cue for inference. Likewise, if both of the cities have (or do not have) soccer teams in Bundesliga, then again the cue would not be helpful for inference and thus would not be activated. If such conditions prevent a cue from being activated, then another cue is generated and goes through a similar testing cycle. If the values of both alternatives on the cue is known and they are different from each other, then an inference is made and the search is over. The choice rule is stated as follows (Gigerenzer et al., 1991, p.508):

$$\text{Choose } a \text{ if } p(a|aC_ib; R) > p(b|aC_ib; R)$$

Furthermore, according to the PMM theory when the participant chooses *a*, her confidence in her answer being correct should be equivalent to the cue validity, that is $p(a|aC_ib; R)$.

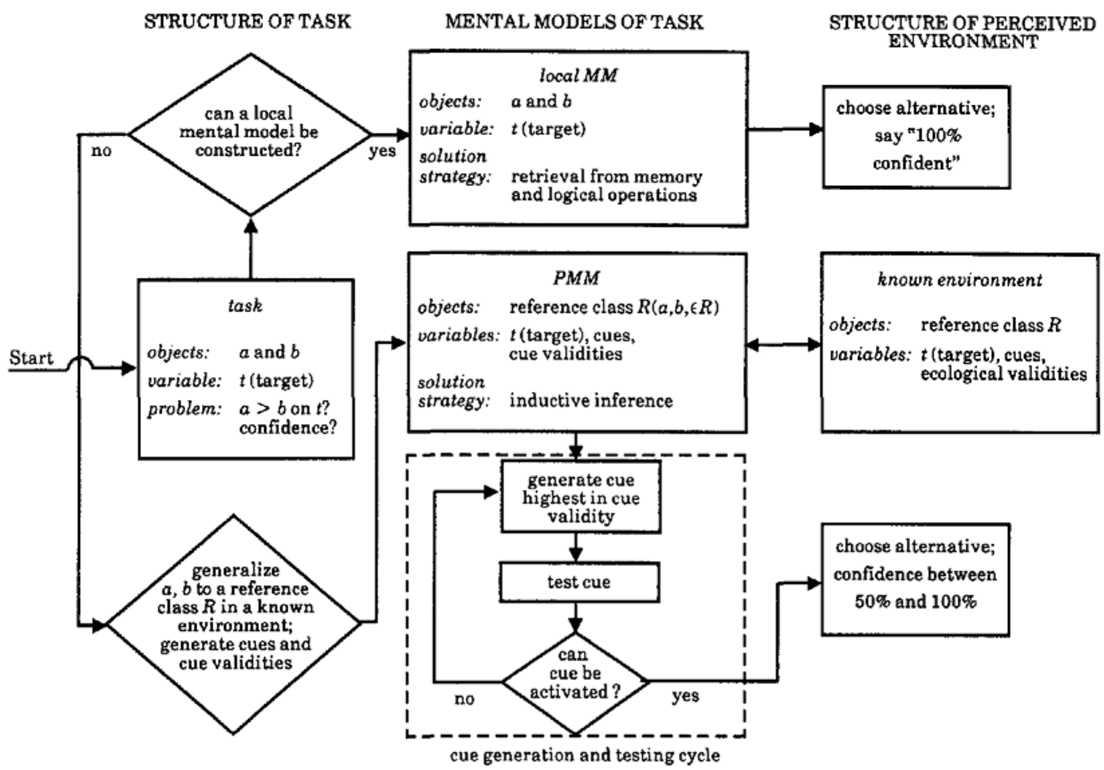

Figure 3. Probabilistic mental model. (Gigerenzer et al., 1991, p.508).

CHAPTER 2

ON RATIONALITY

In his criticism of the studies on the cognitive fallacies and biases, Gigerenzer (e.g., 1994, 2000) asserts that the rational norms that the researchers from the heuristics-and-biases program use to evaluate human performance are disputable. He draws our attention to the fact that there are different interpretations of probability and that the same reasoning that is seen as a violation of probability theory by some school of probability can be sanctioned as rational by another. Thus, he argues, the problems used in the cognitive fallacy studies do "not have one and only one answer" (Gigerenzer, 2000, p.244). Since these problems do not have a single correct answer, Gigerenzer claims that using them for demonstrating reasoning errors does not make sense.

I will argue that Gigerenzer is too hasty in dismissing the findings of the heuristics-and-biases program based on the debates on probability theory and that his stance on this issue is self-contradictory. I will first contend that the probability interpretations that he cites do not pose a real challenge for the findings of the heuristics-and-biases program, unless he provides reasons why we should adopt these particular views for our analysis of the norms. I will suggest that he should formulate a set of criteria on the basis of which he should discuss whether these probability views are suitable for the interpretation of the norms. Then I will argue that the probability interpretations that he draws on in his argument are in contradiction with each other and with his own theories and conclude that he is not justified in using these views in his arguments. I will discuss the adequacy of the alternative probability theories that he offers for the interpretation of the norms and claim that they are not admissible. Next, although he claims that he does not take sides on the issue of which probability interpretation is correct, I will argue that this is not the case. I will claim that he is a finite-frequentist because he builds his own theories on finite-frequentism and that consequently he should analyze the norms from the perspective of finite-frequentism. Lastly, I will contend that his stance on the norms lacks a theoretical basis and is merely dependent on human performance (description).

One and Only One answer

In his analyses of the Linda problem, the base rate neglect and the overconfidence effect, Gigerenzer(2000) states that different interpretations of probability yield different results on these studies; he claims that by switching from one interpretation to another one can make the cognitive illusions “disappear, reappear, or even invert” (p.243). His argument is based on the false belief that each probability interpretation is equally admissible and valid and that no lines can be drawn between a good and a bad one. This leaves the false impression that each probability interpretation is applicable to the problems used in the studies and that consequently it is not possible to assign uniquely correct answers to the problems.

Contrary to Gigerenzer's claims, how probability is interpreted is not simply a matter of faith and exempt from a rational discussion; through an analysis of its weaknesses and strengths it is possible to accept or reject a theory. However, in his refutation of the norms used in the studies, Gigerenzer is content with merely pointing out that if certain different interpretations of probability are adopted, then one would arrive at different conclusions on the studies. For example, as has been mentioned in the first chapter, he claims that if von Mises' interpretation is adopted then it would not be possible to call the violations of the laws of probability as fallacies since von Mises believes that probability cannot be assigned to single-events. Yet, he fails to address the most important question pertinent to his criticism: why should we base our norms on von Mises' view of probability?

Failure to provide a justification for adopting different views of probability is the main weakness of Gigerenzer's arguments against the norms. Merely pointing out that there are conflicting views on a subject does not constitute a strong argument, because it is possible to find conflicting views about literally every subject in the world. If we are to lift the obligation of providing good reasons for adopting a certain view, then any theory can be easily refuted and any nonsense theory can be accepted. For instance, the diagnosis of a person who thinks that he is Napoleon with schizophrenia can be said to be disputable, simply because his schizophrenic inmates agree with him that he is indeed Napoleon. Similarly, one can easily dismiss Gigerenzer's own theory of ecological rationality as nonsense just by citing the creationist view, which does not acknowledge that evolution has taken place. However, I do not think that Gigerenzer would take such a criticism of his theory seriously, because he most probably does not think that creationism is a credible theory to begin with.

As I have stated above, it is possible to evaluate the soundness of a probability theory by analyzing its weaknesses and strengths. For instance, Wesley Salmon (1966) put forth a set of criteria, such as applicability and ascertainability, and assessed the adequacy of each interpretation of probability based on them. In a similar fashion Gigerenzer could have come up with his own set of criteria and based on these criteria discussed why the probability interpretations that he suggests provide an eligible basis for the norms. For instance, because Gigerenzer is interested in developing a theory for solving real life decision problems, we can speculate that applicability would be an important criterion for him, as it is for Salmon. Likewise, one of his criteria could be inspecting what advantages adopting a probability interpretation could offer on evolutionary terms, as his theory of ecological rationality is built on evolution. For instance, instead of dismissing single-event probabilities as meaningless and claiming that they are not subject to the laws of probability, he can discuss why adopting a probability interpretation that enforces the laws of probability for past frequencies but allows them to be violated for single-events is advantageous on evolutionary terms.

Different Interpretations of Probability

Even though Gigerenzer (1994) avoids discussing the admissibility of the probability theories that he bases his criticisms on, stating that his aim is "not to champion one side over another – frequentism over Bayesianism, or vice versa" (p.141), we can nevertheless claim that he does not and cannot agree with them, because they are in conflict with his own theories

Von Mises' Interpretation of Probability

On several occasions Gigerenzer(e.g., 1994, 2000) refers to the frequentist Richard von Mises in his arguments against the application of frequencies to single-case probabilities. Von Mises (1957) believes that probabilities are long-run frequencies, and he defines probability with limiting relative frequencies in infinite reference classes. He invents the term 'collective', which refers to "hypothetical infinite sequences of attributes (possible outcomes) of specified experiments that meet certain requirements" (Hajek, 2007, p.567). Then, the limiting relative frequency of an attribute *A* in a collective ω is defined to be the probability of *A* relative to ω (Hajek, 2007). Von Mises asserts that probabilities must always be defined relative to a collective. He points out that this definition of probability necessitates probabilities be properties of classes and thus renders probabilities of individual instances meaningless. More specifically he states:

> We can say nothing about the probability of death of an individual even if we know his condition of life and health in detail. The phrase 'probability of death', when it refers to a single person, has no meaning at all for us. (Von Mises, 1957, p.11)

Gigerenzer frequently cites this quote from Von Mises to support his argument that for a frequentist single-event probabilities are meaningless. However, Gigerenzer does not realize that the same argument that von Mises has against the single-event probabilities holds just as well against any finite case probability. Hajek(2007) draws our attention to the fact that von Mises defines collectives to be infinite sequences, and thus, according to this definition finite reference classes are not qualified to be collectives, either. Therefore, he argues, for von Mises probability is undefined not only for the death a single individual, but even for the death of a billion people; for as big as this reference class is, it is nevertheless finite, and consequently does not qualify for being a collective. Hence, from von Mises' perspective, probabilities for finite reference classes are also meaningless.

Gigerenzer's own theory of probabilistic mental models exclusively relies on finite relative frequencies for probabilistic inference. For the Linda problem he claims that replacing single-event probabilities with relative frequencies makes 'the cognitive illusion disappear.' In his arguments against the overconfidence effect he claims that replacing a subject's confidence in the correctness of a single question, which is a single-event probability, with the relative frequency of correct answers eliminates the overconfidence effect. Even the execution of some of his fast-and-frugal heuristics, such as the Take-the-Best, depends on finite relative frequencies. If Gigerenzer believes that von Mises' stance on single-event probabilities is worthy of notice and should be taken seriously, then he should also accept that in light of this view his treatise of the fallacies and his PMM theory, which are all based on finite relative frequencies, are also meaningless.

More importantly, even if finite relative frequencies were legitimate probabilities, the use of cue validities for inference in the PMM framework contradicts Gigerenzer's previous attacks on the norms of the heuristics-and-biases program. Cue validities are relative frequencies, and thus when one uses a cue for inference based on its validity, one would actually be applying a relative frequency to a single-case probability. In fact, Gigerenzer (2000) himself states that in their account of the overconfidence effect, both the subject's confidence in having answered a single question correct and her judgment of the relative frequency of her correct answers are "explained by reference to experienced frequencies" (p.137). Although it is stated that single-event probabilities and relative frequencies "are not evaluated by the same cognitive process" (Gigerenzer, 2000, p.136-137), the only difference between these two probability assessments turns out to be the reference classes and the target variables they use. For example, when the question asks which German city among the alternatives has the largest population, confidence in having answered this question would use 'cities in Germany' as the reference class for inference. On the other hand, when the subject is asked for her estimate of the relative frequency of her correct answers, then her reference class will consist of "series of similar questions in similar testing situations" (Gigerenzer, 2000, p.136). Therefore, when a person is assessing her past performance, or answering a question, she makes an inference based on the relative frequencies. Thus, Gigerenzer's PMM theory legitimizes the use of relative frequencies for inference in single-case probabilities.

De Finetti's Interpretation of Probability

As I have explained in the previous chapter, in his criticism of the studies on the overconfidence effect Gigerenzer (2000) asserts that if the subjectivist interpretation of probability, as defended by Bruno de Finetti, is adopted, then the discrepancy that is observed between the relative frequency of correct answers and a participant's confidence in having answered a single question right cannot be called a bias. He states that for the subjectivists, relative frequencies are considered irrelevant for probability assessment. He cites the following quote from de Finetti(1931/1989):

> However an individual evaluates the probability of a particular event, no experience can prove him right, or wrong; nor, in general, could any conceivable criterion give any objective sense to the distinction one would like to draw, here, between right and wrong. (p.174)

This is a direct consequence of de Finetti's theory of probability in which he identifies probabilities with 'degrees of belief'. In his account, probability assessments are purely subjective and one does not need to draw on relative frequencies or, in fact, on anything at all for making probability judgments (Hacking, 2001; Hajek, 2007). What is interesting is that this probability theory that Gigerenzer cites to criticize the heuristics-and-biases program, is also in conflict with his own theory of probabilistic mental models. It is explained that the "PMM theory proposes a frequentist interpretation of degrees of belief: Both confidence and frequency judgments are based on memory about frequencies" (Gigerenzer et al., 1991, p.526). Thus, like Kahneman and Tversky, Gigerenzer does not permit unwarranted probability assignments. Therefore, if de

Finetti's extremely permissive probability theory is to be adopted for the norms as he suggests, then the legitimacy of his own theory of probability judgments will be in jeopardy.

This is not the only point that de Finetti's probability theory is conflict with Gigerenzer's views. Although de Finetti allows probability judgments to be made solely on the basis of subjective beliefs, he does not allow them to be completely unrestrained. He imposes the requirement that a person's probability judgments be "coherent." A probability judgment is said to be coherent if it cannot be exposed to a "Dutch book," that is, a sure-loss contract. Conforming to the axioms of probability protects one against a Dutch book (Hajek, 2009). As I have explained in the first chapter, Gigerenzer claims that the laws of probability cannot be imposed on the Linda problem, because frequentism renders single-event probability meaningless and thus exempt them from the laws of probability. However, De Finetti would interpret the violation of conjunction rule in the Linda problem as a fallacy, because noncompliance to this rule would render one's probabilities incoherent, and thus vulnerable to a Dutch book. [7]

While Gigerenzer quotes de Finetti as a reliable authority in his analysis of the overconfidence effect, he prefers to ignore him in the Linda problem and opts for the completely opposite view. He cites the following quotation from the statistician George Alfred Barnard (1979):

---

[7] It can be easily seen that when one does not follow the conjunction rule, a Dutch Book can be played against them ensuring sure loss. In fact, Kahneman & Tversky (1996) report that several studies revealed that the violation of conjunction rule is prevalent in most people's betting behaviour as well.

> If we accept it as important that a person's subjective probability assessments should be made coherent, our reading should concentrate on the works of Freud and perhaps Jung rather than Fisher and Neyman(p.171).

When Gigerenzer's own personal views are so profoundly in disagreement with de Finetti's theories, it is not reasonable for him to suggest that de Finetti's subjectivist view can serve as a good interpretations of the norms. Gigerenzer's inconsistent stance is proof that he refers to different interpretations of probability not because of theoretical reasons, but simply to save his arguments against the verdicts of the heuristics-and-biases program.

The Adequacy of the Probability Theories

In the previous section I have discussed the admissibility of von Mises and de Finetti's interpretations of probability vis-à-vis Gigerenzer's own standpoint. However if we take Gigerenzer's theory out of the picture, can we argue that these probability interpretations comprise admissible basis for the norms and thus pose a real objection for the studies done on cognitive fallacies? I will assess the adequacy of the above mentioned views of probability and argue that they do not provide an admissible basis for the norms used in real life decisions.

One of the three criteria that Salmon (1964) puts forth for assessing the adequacy of probability interpretations is "applicability." I agree with Salmon that applicability is a very important criterion and believe that a probability theory should provide a good guide for us to help deal with the uncertainties we face in life. Since rationality norms concern real life decisions, a probability theory that does not account for the probability judgments we make in our daily lives cannot provide an admissible interpretation for the norms.

Most of the decisions that we have to make in life pertain to single events; when a doctor is diagnosing a patient, a weather forecaster is deciding whether it is going to rain today, or a banker is deciding whether he should grant a loan to a customer, they all make single event probability assessments. Therefore, the adequacy of a probability interpretation that dismisses single event probabilities as meaningless is quite questionable. Von Mises' theory is especially problematic because it requires infinite sequences, which are simply not available to us in real life. For these reasons, I do not believe that the probability interpretations suggested by von Mises and other radical frequentists would provide a acceptable basis for the norms.

The adequacy of de Finetti's subjectivist probability theory is also questionable. As has been stated before, de Finetti does not impose any restriction on probability assessments, except for them to be coherent. Therefore, according to this radical subjectivist view a person is justified in assigning almost any probability to any event without providing any evidence or rationale for it. This stance on probability assessments draws a lot of criticism due to its extreme permissiveness and raises doubts about taking it as a benchmark for rationality. The permissive nature of this theory results in allowing even the most ridiculous probability judgments to be approved of as rational; for example, Hajek (2007) states that according to the radical subjectivists, a person would be considered rational even if she assigns a probability of 0.999 to the event that George Bush turns into a prairie dog, as long as she assigns 0.001 probability to this event not happening.[8] Hajek contends that such a theory that permits high probabilities be assigned to even false propositions is a "no-theory theory of probability" that becomes an "autobiography rather than epistemology" (Hajek, 2007, p.577) .

I think the criticisms that are raised against de Finetti's subjectivist theory are quite compelling; a person that makes his decisions based on whim, rather than evidence can hardly be said to be rational. For example, I do not think that Gigerenzer would trust the judgment of a doctor who diagnoses his patients without even finding out what their symptoms are. Therefore, I do not believe that radical subjectivist view that sanctions unrestrained probability assignments can be taken as a serious criticism of the rationality norms used in the cognitive fallacy studies.

As for the coherence condition, I do not agree with the view expressed in the quote taken from Barnard and believe that coherence is an important requirement for rationality. Several examples can be given to show how incoherence leads to irrationality. For example, a doctor that assigns '1', that is expresses certainty, to both the probability of a patient's recovery from a sickness and his dying from it could not be argued to be rational; an event and its complement cannot co-occur.

[8] Thus, the probability of the event "George Bush turns into a dog" and its compliment would add up to 1. This would assure that the probability assignment conforms to the axioms of probability, and fulfils de Finetti's coherence condition.

On Taking Sides

Gigerenzer (1994, 2001b) states that his aim is not to take sides with any of the probability interpretations, but to merely bring to our attention that the norms heuristics-and-biases program rests on are open to discussion. Kahneman and Tversky(1996) describe his stance on this issue as “normative agnosticism” and criticize it for being “unreasonably permissive” (p.586). However, I do not think that “normative agnosticism” is a correct description of Gigerenzer’s position because he is not actually undecided or neutral about which interpretation of probability is the correct one. His criticism of subjective probability and the fact that his own theories are built around frequentism clearly indicate that he does take sides on this issue. Consequently, I believe that his stance on the meaning of probability is not simply agnostic, but self-contradictory and unwarranted, and since he believes that humans are frequentist intuitive statisticians and that frequentism provides a good guide for making decisions in real life, his analysis of the studies of the heuristics-and-biases program should be done from the perspective of frequentism and not just from the angle of any interpretation that serves his argument.

Furthermore, within the frequentist school, his pronouncements indicate that Gigerenzer is in line with finite-frequentism.[9] Thus, his criticism would be acceptable only if he bases them on finite-frequentism. For example, although von Mises is a

[9] This follows from the fact that Gigerenzer’s PMM theory uses relative frequencies for calculating cue validities and dictates that single event probability judgments are based on relative frequencies.

frequentist, he is a hypothetical frequentist and that is not in accord with finite-frequentism. Therefore, Gigerenzer would not be justified in relying on von Mises' views in his arguments.

Ecological Rationality

I agree with Gigerenzer that laws of logic and probability should not be applied to problems content-blindly. The examples that he gives clearly illustrate how ambiguity in language and the intricacies of social relations can make the application of the principles of classical rationality difficult.[10] To avoid such confusions Gigerenzer(2001b) claims that the norms should be constructed by paying attention to the specific requirements of the environment. However, he never actually applies this principle in his studies. Instead of analyzing the environment to designate the appropriate norm, he judges the suitability of a norm based on its descriptive accuracy. When he finds out that people do not follow a norm, he focuses on proving that the norm is not suitable to the structure of the environment. If, on the other hand, he knows that people obey a norm, he tries construct an environment in which the norm would be

[10] For example, Gigerenzer(1996b) gives the following example to show how a social situation might make the violation of an axiom from decision theory look rational. Imagine you are at a dinner party, and you have been offered a bowl with a single apple in it. You have two choices: (A) you take the apple, or (B) you leave the apple. Out of consideration of others, you might be reluctant to take the last apple in the bowl and prefer to not take it. However, the host of party comes back and puts one more apple in the bowl(X). Now that there two apples in the bowl, you will probably choose to take the apple (A). Hence the introduction of a third element reversed your preferences about A and B. According to decision theory, this would be a violation of the independence of irrelevant alternatives axiom. Independence of irrelevant alternatives axiom of decision theory states that if A is preferred to B in the absence of X among the choices, then introducing X to the alternatives should not make B preferable to A. When social context is taken into consideration this violation seems like a reasonable behaviour.

ecologically rational. Unfortunately, adopting such a strategy makes it impossible to show that people can be irrational. The fact that Gigerenzer analyses the norms on a descriptive basis shows that his main concern is not actually constructing norms sensitive to the specific structure of the environment, but to prove that people are fundamentally rational.

That Gigerenzer (2008) is interested in setting norms descriptively is evident in his following statement :

> For instance, when Wason and Johnson-Laird (1972) criticized Piaget's logical theory of thinking as descriptively incorrect, they nevertheless retained the same logical standards as normatively correct for the behavior studied. When Tversky and Kahneman (1983) reported that people's reasoning violated a law of logic (the "conjunction rule"), they nevertheless retained logic as the norm for rational judgment. (p.6-7)

Thus, it seems that Gigerenzer would get rid of the norms, rather than accepting that people are irrational. However this reasoning is flawed. For example, the fact that a lot of elementary school students perform poorly in solving arithmetic and algebra problems does not indicate that algebra and arithmetic is wrong. Unfortunately, Gigerenzer's response to the studies on the base rate neglect is quite similar to this inference; to explain people's poor performance he claims that the Bayes' rule does not comprise a normative standard. He first claims that according to frequentism, no rules or laws can be imposed, because it posits a single-event probability and in frequentism single events do not have probabilities. Next, he contends that there are several different ways to solve the base rate problems which yield different solutions.[11] Thus, he claims there is not "just one 'correct' answer" (Gigerenzer,2000, p.262). Surprisingly, Gigerenzer's

[11] Such as the one suggested by Michael Birnbaum (1983) based on signal detection theory.

skepticism towards Bayes' rule disappears when studies show that people obey it when the problems are posed in natural frequencies. Furthermore, Gigerenzer takes pride in having trained hundreds of medical doctors to improve their decision making skills. Interestingly enough he has trained them to solve mammogram-type problems following Bayes's rule (Gigerenzer, 2008, chap. 8). Since he trusts the correctness of the Bayes' Rule to the degree that he would train doctors to use it in their decisions, his objection to its use in the heuristics-and-biases program is unjustified.

Furthermore, because his claims about the norms do not have a coherent theoretical basis and are adopted because they model human behaviour well, Gigerenzer cannot avoid making contradicting statements. For example, when experiments showed that relative frequency format does not facilitate Bayesian reasoning, but natural frequency format does, Gigerenzer (1994) comes up with an explanation to show that natural frequencies are more adaptive than relative frequencies are:

> Assume that some capacity of algorithm for statistical reasoning has been built up through evolution by natural selection. For what information representation would such an algorithm be designed? Certainly not for percentages and single-event probabilities (as is assumed in many experiments on human reasoning), since these took millennia of literacy and numeracy to evolve as tools for communication. Rather, in an illiterate and innumerate world, the representation would be frequencies of events, sequentially encoded as experienced – for example, 3 out of 20 as opposed to 15% or p=0.15 (p.142).

However, when his studies on the overconfidence effect demonstrate that people can actually reason well with relative frequencies, he contends that humans use relative frequencies for statistical reasoning. [12]

The alternative approach to rationality (i.e., ecological rationality) that the bounded rationality program introduced has made quite a big impact in the literature. For example, the Nobel laureate economist Reinhard Selten proposed developing an economic theory based on ecological rationality (Selten, 2001). Even the economist Vernon Smith, who shared the Nobel prize with Daniel Kahneman for his work on cognitive biases and fallacies, expressed his support for modeling economics on ecological rationality (Hammond, 2007, p. 234, p.266). Ecological rationality certainly offers quite an interesting and appealing view on rationality with its emphasis on the close relation between the environment and the mind. However, I believe that the concept and usage of ecological rationality remain a bit vague.

The ecological rationality view advocated by Gigerenzer and his colleagues is defined in the following way: a mental process or a heuristic is "ecologically rational to the degree that it is adapted to the structure of an environment" (Gigerenzer& Todd, 1999, p. 13). According to this definition, it is the process that leads to the decision rather than the decision itself that gets assessed. Therefore, this procedure checks whether a person is justified in using a certain strategy in an environment rather than whether he makes the decision that assures success. Then, a person can be said to be ecologically rational even if his solution to a problem is wrong. Inspecting the

[12] For example, the calculation of cue validities in the PMM and Take-the-Best all rely on relative frequencies.

adaptedness of human reasoning to a certain environment can only tell us why people do not follow a norm. However, it would not rule out the possibility that people are in violation of the norm. For example, in his discussion on the natural frequencies Gigerenzer (2000, chap. 6) only explains why people cannot be expected to perform well when the questions are presented in normalized probability format. However, the fact remains that they cannot provide the correct answer when the questions are in probability format.

The fact that our reasoning skills are not adapted to a certain environment does not mean that there are not right actions to be taken in that environment. For instance, if as Gigerenzer claims we have not evolved to work with normalized probabilities, does that mean that the questions posed in normalized probabilities do not have correct answers? The fact that we have not evolved to the modern world does not mean that we should give up trying to prescribe answers to problems. When our heuristics prove to be insufficient, we would necessarily need to rely on the classical tools of rationality for making our decisions. Gigerenzer is correct in his claim that their application to real life problems can be sometimes tricky; however I still believe that the classical benchmarks of rationality, such as logic and probability theory, are reliable tools for constructing the norms.

CHAPTER 3

ON HEURISTICS

Gigerenzer and the ABC Research group have launched the bounded rationality program to show that the heuristics that have been held responsible for making human reasoning prone to biases and fallacies are actually the pillars of good decision making. They have two important claims about the heuristics: First, they are intuitive and they are boundedly rational (i.e., computationally cheap, work fast, and require little information) and secondly they lead to correct decisions. In this section I will question the apparent simplicity of the fast-and-frugal heuristics and claim that their execution requires large amounts of information and cognitive resources, and that they cannot be reduced to simple intuitions. Additionally, I will contend that although intuitions and heuristics play an important role in our decisions, they do not necessarily lead us to correct answers and conclude that they cannot be used prescriptively.

Frequentism

I will start my discussion by questioning the adequacy of finite frequentism for probability judgments and decision making. There are a few reasons why a discussion of finite frequentism is necessary. First, probability judgments are an essential part of any decision theory. To account for the uncertainty in the environment and our incomplete knowledge, we need to employ probabilities while making our decisions. Therefore, how a decision theory addresses the question of how we make these probability judgments requires a close inspection. Secondly, without frequencies Gigerenzer cannot explain how probability judgments can be made intuitively, as his argument is contingent on the natural sampling and automatic frequency-registering processes. Lastly, frequentism constitutes the basis of the probabilistic mental model theory and the fast-and-frugal heuristics based on it, such as the Take-the-Best. Although it can be argued that there are many other heuristics that do not rely on frequencies, I believe that the Take-the-Best deserves special attention, since it was with this heuristic that Gigerenzer and his colleagues were able to show the efficiency of the fast-and-frugal heuristics compared to optimization methods.

The Shortcomings of Finite-frequentism

Solely relying on relative frequencies might mislead one to make inaccurate probability judgments. For example, a die could be perfectly fair, but nevertheless it is quite possible for us to observe the outcome “3” six times out of ten tosses, leading us to conclude that the die is biased and the probability of the outcome “3” is 0.6. Similarly, the toss of a fair coin might yield 8 tails out of 10 tosses. In other words, even a fair coin might not yield heads exactly half of the time, and thus a finite-frequentist who bases her probability judgments only on relative frequencies would wrongly judge a fair coin to be not fair. This problem has very important consequences for Gigerenzer’s theory. A fair coin that is incorrectly judged to be biased, translates into Gigerenzer’s theory as a cue that has actually no predictive value being taken as an important cue for the decision. For example, in his explanation of natural sampling, Gigerenzer talks about a foraging animal learning cues and their validities for finding food resources in the wilderness. Relying solely on frequencies might mislead this animal to judge something that is in fact irrelevant and not indicative of the existence of a food resource to be quite a valid cue for inference. For instance, let us say that this animal likes the leaves of a certain type of tree and is trying to learn the cues which might indicate its presence. Assume that this forest is inhabited by a particular type of bird which randomly picks the trees that it is going to build its nests on. Although the presence of the nests on a tree are random, it could happen that in most of the cases that our foraging animal found its favorite tree, the nests of this particular type of bird were present on its branches. This would lead the animal interpret this mere coincidence as an important factor and register the presence of the nests as an important cue.

Another problem with finite frequentism is that it does not help us assign probabilities to events if no observations have been made thus far. For example, no probability can be assigned to the outcome heads if the coin has not been tossed yet. Therefore, Gigerenzer's intuitive statistician would be clueless when he finds himself in a situation it has not 'sampled' before. This is a serious drawback for his theory, because in life we have to make decisions in novel circumstances all the time. Additionally, finite frequentism fails to assign probabilities to 'unrepeatable' events, such as the end of the first world war. This is also one of the criticisms that Kahneman and Tversky (1996) have against Gigerenzer's frequentist theory; they contend that in real life people have to make probability judgments and decisions about events that cannot be related to frequencies as they occur only once, such as a decision about whether to undergo an operation. They argue that such events call for belief-type probabilities.

Assigning probabilities to rare events is another problem finite frequentism runs into. Assume that we have tossed a coin only once, and the outcome was tail. Based on this observation, finite frequentism would assign 1 to the probability of tail and 0 to heads and this is a quite unwelcome result. For example, if a couple has only one child and it is girl, it would not be reasonable to think that the probability of their next child being a boy is 0. This problem brings up the question: how large should a reference class be? After how many samples can the intuitive statistician be confident about relying on her learned frequencies (i.e., cue validities) for making decisions? Can a doctor who has seen only a handful of patients be a reliable source for good diagnosis? Gigerenzer is fully aware of the importance of sample size. For example, he notes that a person should be more confident about making a judgment about the fairness of a die after 100 tosses than only 10 tosses (Gigerenzer,2001b). Nevertheless, in his theory Gigerenzer does not address this problem and does not discuss how big the sample size should be for ensuring successful inference; neither are his fast-and-frugal heuristics sensitive to the sample size, as their formulation also relies on ratios.

The Reference Class Problem

Although rare-events and single-events pose some difficulties for frequentism, the most serious problem that it runs into is the so-called reference class problem. Frequency judgments are made according to the relative frequency of an event or an attribute in a reference class; however which reference class should be used for the probability judgment is an unresolved issue, and frequency judgments can have extremely varying results based on which reference class is used. For example, let us imagine that a doctor who needs to assess the probability of her patient being positive for HIV. If she takes the entire population of Turkey as her reference class, she will assign a low probability, since the percentage of people who are positive for HIV is not that high. If, on the hand, she takes “heroin addicts in the poor districts of Istanbul” as her reference class, she will assign a high probability. Therefore, picking the “right” reference class is crucial for frequentistic probability judgments.

The most critical problem that reference class selection poses for Gigerenzer’s theory is that it undermines the apparent simplicity of his heuristics; selecting a relevant reference class to ensure good inference requires both a lot of information and good cognitive skills. A decision maker needs to have good knowledge about the environment, be aware of the causal relations between various factors, and be capable of performing complex reasoning. Therefore it is hard to argue that people can intuitively make frequency judgments. For example, in the example I gave above, it is necessary for our doctor to know that HIV is more common among heroin addicts, that HIV can be transmitted by sharing needles. Also, if she knows that this patient is very careful about his personal hygiene and comes from a wealthy family, she can deduce that it is unlikely that he would share needles with others and thus preclude using "heroin addicts in the poor districts of Istanbul" as her reference class.

Another reason why reference class selection is very important for his theory is that depending on which reference class is selected, the cue validities can greatly vary, leading to different inferences. A cue can have high predictive power in one reference class (i.e., a high relative frequency) and prove to be quite useless in another. Imagine that a doctor finds out that the patient has chest pains while examining his 20 year-old patient and. If he takes the "60 year-old patients" as his reference class, he might think that "chest pain" is a very important cue for his diagnosis (i.e., that it has a high cue validity) , as chest pain is an important precursor of heart attacks, which is quite common among elderly patients. However, if he takes "20 year-old patients" as his reference class, he might dismiss the chest pain as an important cue for his diagnosis, because heart attacks are rare among young people. As can be seen, depending on which reference class is taken, a cue might be judged to have a high or a low validity, which would have tremendous effects on the final decision.

Most of the discussions on the Take-the-Best and the PMM theory in the literature have focused on the computation of the cue validities and their ranking. Nevertheless, I believe that the selection of cue identities is a more important issue than the computation of their validities. Identification of cues relevant to the decision is a very complex process that requires domain knowledge and good reasoning skills, which again undermines the “fastness” and “frugality” of Gigerenzer’s heuristics. Additionally, it should be noted that identification of cues is also closely related to the reference class problem. When answering the questionnaire about the population of various German cities, if a decision maker does not select "the cities in Germany" as her reference class, she might not pick "being in the industrial belt" as one of her cues. The animal in the wilderness looking for its favorite tree might exclude "look for the specific flowers of the tree" as a cue, because it is not flowering season. Therefore, selecting a relevant reference class is necessary not only for the computation of cue validities, but also for the selection of relevant cues.

Gigerenzer (2000) is aware of the limitations of his theory; he maintains that decision makers should be sensitive to the changes in the world and careful about picking their reference classes, stating: “The intuitive statistician…must first check the structure of the environment (or of a problem) in order to decide whether to apply a statistical algorithm at all, and if so, which” (p.263). Then he gives the following example:

> You live in a jungle. Today you must choose between two alternatives: to let your child swim in the river, or to let it climb trees instead. You use only one criterion for that choice, your child's life expectancy. You have information that in the last 100 years there was only one accident in the river, in which a child was eaten by a crocodile, whereas a dozen children have been killed by falling from trees. Just yesterday your neighbor told you that her child was eaten by a crocodile. Where do you send your child? (Gigerenzer,2000, p.263)

Gigerenzer (2000) claims that instead of updating her cue validities in light of new information, this mother would think that the previous information, or reference class, is no longer relevant to the problem, because she might "suspect that the small river world has changed – crocodiles may now inhabit the river" (p.263). To avoid such mistakes, he states that the decision maker "has to check structural assumptions before entering into calculations" (Gigerenzer,2000, p.264). Unfortunately, when there are so many assumptions to check and causal relations to contemplate, it can no longer be claimed that the mother solves this decision problem merely by relying on her hunches.

My claim is not that we do not use frequencies for probability judgments at all, but that as Ian Hacking (2001) states, "Real-life application of the frequency principle requires a lot of judgment" (p.138). I do not believe that the intuitive statistician with a frequency-registering mechanism that works "without intention or much attention" presents a realistic model of human probability judgment, and I think that the underlying complexity of the frequentistic probability judgment robs the heuristics of their speed and frugality.

Empirical Findings

So far I have pointed out the theoretical problems with frequentism. Now I will briefly discuss the empirical studies that have investigated the psychological reality of Gigerenzer's theory.

Teigen, Brun and Frydenlund (1999) carried out a study that aimed to discover the role of frequencies in people's probability judgments. Their study revealed that people base their probability judgments more on considerations of the causal relations than frequencies. For example, people believe that a soccer team has a high probability of winning the next game not because it has won most of its past games but because the members of the team are highly qualified players (causal or dispositional interpretation). Their study showed that people ignore statistical evidence especially in the cases where the subjects were convinced about a causal explanation for the probability statement. For instance, to explain why 'car accidents when driving with summer tires in the winter' is highly probable, more than 90% of the subjects relied on non-frequentistic concepts, most of which emphasized causal conditions. (Teigen et al., 1999) The authors claim that the participants preferred to have causal proofs even in the cases where they were aware that statistical evidence could be provided. For example, they state that for evaluating the probability of a person successfully quitting smoking, the subjects preferred to know how motivated the smoker was about quitting, instead of the relevant statistical information. They concluded that frequentism by itself cannot account human probability judgments and the propensity interpretation of probability seems to offer a more accurate model for it.[13]

A study by Garcia-Retamero, Wallin and Dieckmann (2007) on the empirical validity of the Take-the-Best heuristic demonstrated that people look up cues that are causally linked to the problem more often than neutral cues, regardless of their validities. They report that the participants were more likely to make decisions based on the causal cues and they were more confident and faster in their decisions when they relied on the causal cues for inference. The results of this study agrees with Teigen et al. (1999)'s study that people rely on causal links more than statistical evidence.

Newell, Weston, and Shanks (2003) carried out experiments to test whether people actually use the Take-the-Best heuristic. Their experiment used a computer-based questionnaire which required the participants to answer a series of forced-choice questions in which they had to choose the most profitable one between the shares of two fictional companies. To assist them in their decision, participants were allowed to buy up to six pieces information (i.e., cues) about the companies, such as "Is the company listed on the FTSE?" or "Is the employee turnover low?", at the cost of 1 pence, while they got paid 7 pence for making a correct prediction. Participants were allowed to learn the validities of the cues through a trial period. Newell et al. kept records of the order in which each participant bought information and closely examined whether each individual adhered to the search, stopping and the decision rules of Take-the-Best. Their study revealed that 75% of the participants followed the search rule of Take-the-Best,

[13] Propensity interpretation of probability was developed by Karl Popper. According to this view, probability is the propensity of a certain physical set up to produce certain outcomes.

that is they looked up cues in order of their validity, but only 33% of the participants never bought unnecessary information. They report that most of the participants followed what they call ‘weight of evidence’ strategy, following which they bought unnecessary information in an attempt to seek confirmation for their decision and reduce risk. Lastly, 50% of the participants never violated the decision rule of Take-the-Best, in other words they chose the alternative that the most valid cue indicated. Newell et al. state that only 33% of the participants adhered to all three rules of Take-the-Best. Although Newell et al. are of the opinion that the results of their study challenges the psychological reality of Take-the-Best, they report that Gigerenzer does not agree with this view arguing that the environment of the experiment was not one in which Take-the-Best could thrive. Newell et al. warn that providing such post hoc explanations for the violations of heuristics runs the risk of rendering the theory of fast-and-frugal heuristics unfalsifiable. As they point out, when the circumstances under which people would be expected to follow a certain heuristic is not explicitly specified, empirically testing the psychological reality of this heuristic becomes impossible. Although initially I was planning to do an empirical research, similar to that of Newell et al., to test the psychological plausibility of fast-and-frugal heuristics, this discussion made me realize that carrying out more experiments that show that people do or do not follow a certain heuristic would not produce any conclusive results. For example, I have thought about testing the psychological plausibility of Take-the-Best on decision problems with varying pay-off structures. However, because it has not been specified when the trade-off between cost of information and pay-off would trigger the use of Take-the-Best, such an experiment would not be able to determine the descriptive accuracy of Take-the-Best.

I have run into a similar problem when I tried to construct a simulation environment to test the performance of recognition heuristic compared to an optimization method, such as a Bayesian network. The definition of the recognition heuristic is a tautology that does not lend itself to scientific testing: it is explained that the recognition heuristic is successful in environments where recognition is strongly correlated with the desired criterion (Goldstein & Gigerenzer, 2002). Therefore, even if I could successfully simulate an environment in which recognition heuristic performs poorly, it would not have been able to falsify the claim that recognition heuristic is a successful strategy. The poor performance of the heuristic would simply be blamed on the environment, in which, it would be argued, recognition is not highly correlated with the desired criterion.

Another problematic aspect of running empirical studies of fast-and-frugal heuristics is the difficulty of simulating real life decisions in experimental settings. Perhaps the most controversial point about the descriptive accuracy of fast-and-frugal heuristics is the claim that people trust such simple strategies while making decisions which have severe outcomes. For example, it sounds plausible to claim that people look at only one criterion while picking which brand of soap to buy; however, it is quite counter-intuitive to think that the same simple strategy would be assumed for a decision that involves expenditure of thousands of dollars, such as buying a new house. Unfortunately, it is not possible to test under experimental settings whether people actually use simple heuristics for such 'big' decisions which involve high risk. In a similar vein, Newell et al. (2003) note that quite a few of their participants preferred to guess the answers without even looking at any of the cues, suggesting that the pay-off structure of the experiment might not have been sufficient to motivate serious engagement in the task.

Most of the empirical studies on the fast-and-frugal heuristics have focused on testing people's ability to compute cue validities and ranking the cues. These studies revealed that the computation of cue validities and their ranking is a difficult task for the subjects (Newell, Rakow, Weston & Shanks, 2004; Dieckmann & Todd, 2004). Ironically, achieving the optimal ranking of the cues was shown to be computationally intractable (Martignon & Hoffrage, 2002), and some critics have pointed out that the automatic frequency counter theory has been refuted by the current memory models and claim that without the frequency counter process Gigerenzer can no longer explain how cue validities are learned (Dougherty, Franco-Watkins, Thomas,2008).

In reply to these criticisms, Gigerenzer, Hoffrage and Goldstein (2008) point out that automatic frequency counter is not an indispensable part of their theory; they state that in their theory they have never assumed a person's cue validities would perfectly match the ecological validities and that the Take-the-Best algorithm assumes that the decision maker subjectively ranks her cues. They acknowledge the fact that individual learning through sampling could be slow, "too dangerous (think of learning by feedback about which mushrooms are poisonous) or practically impossible when the events are rare or feedback absent or unreliable" (Gigerenzer et al., 2008, p.232). To compensate for the shortcomings of individual learning, they suggest two alternative ways of learning cue identities and validities : evolutionary and social learning (Gigerenzer et al., 2008).

Evolutionary learning refers to the biologically inbuilt preferences that we acquired through natural selection. Gigerenzer et al. (2008) claim that evolutionary learning is helpful about cue rankings especially in decisions regarding mate selection and food choice. As an example, they state that female sage grouses use the songs that the males sing as a cue for screening out the viable candidates for mating (Gibson, 1996). The influence of evolutionary learning on our decisions is undeniable, especially in biologically oriented decisions. Furthermore, since these biologically inherited preferences usually manifest themselves as instincts, evolutionary learning fits well with Gigerenzer's intuitive decision making paradigm.

Social learning, on the other hand, refers to learning the identities and validities of cues through exchanging information with other members of the populations. Gigerenzer et al. (2008) claim that social learning accounts for the largest part of learning cue identities and rankings among humans. They suggest that social learning is especially important in areas where individual learning would prove to be too slow or dangerous, such as in medicine. The existence of social learning is an indisputable fact, but unlike individual and evolutionary learning, it is not possible to explain the effects of social learning on our decisions by intuitions. A person who has been taught to make decisions in a particular way, can hardly be said to be an intuitive decision maker nor can his decision making skills be solely ascribed to evolutionary mechanisms. The fact that Gigerenzer opens up room for social learning in his theory shows that he admits that less-is-more effect is valid in only a limited number of cases and that comprehensive domain knowledge and training are essential for good decision making. Intuitions built in through evolution and individual learning alone would be insufficient or misleading. It should also be noted that this is at odds with his natural frequency theory, which was based on the claim that learning occurs exclusively through individual learning by sequential sampling.

Despite the fact that in some ways they are at odds with Gigerenzer's previous discussions, the alternative means of learning cues that they have come up with do compensate for much of the inadequacies their theory. However, while it solves the problem of cue learning, their solution introduces other difficulties to their theory. Gigerenzer et al. (2008) cite three different sources (evolutionary, social and individual learning) for acquiring cue information, but do not explain how a decision maker aggregates the information he acquires from these sources to build a single cue ranking. Assembling a single cue order based on three different resources, which probably yield conflicting information, is by itself a hard decision problem. In fact, it can be argued that how we combine conflicting cues and information lies at the heart of decision making. Telling from our own individual experiences, we all know how the information we gather from others, our natural instincts and personal experiences come into play in decision making. What makes decision making so hard, or probably distinguishes good decisions from bad ones, is how we combine these different factors to arrive at a single decision. Unless they can come up with a way to explain how we can aggregate the information from different sources and rank them in such a way that we can champion one criterion over all the others so as to base our decision on it, I do not think that the Take-the-Best algorithm can be considered complete. Gigerenzer et al. (2008) complain that cue learning and ranking problems concern only the Take-the-Best algorithm and that their critics usually ignore all the other heuristics which do not rely on cue orderings at all. As I have explained before, I think it is natural that the discussions mainly focus on the Take-the-Best; furthermore, I believe that the problems related to the Take-the-Best pose difficulties for their other heuristics as well. For example, as they argue, cue ordering problem is irrelevant to the Minimalist heuristic, because it selects cues randomly. However, the problem of learning cue identities does apply to the Minimalist.

Problems about simplicity and speed arise not only from learning cue identities and validities, but also from the execution of the heuristics themselves. For example, let us examine the "imitate-the-most-successful" heuristic (Garcia-Retamero, Takezawa & Gigerenzer,2006). Imitate-the-most-successful is a heuristic that Garcia-Retamero et al.(2006) developed to explain how social learning can be incorporated into the cue ranking. According to this heuristic, the decision maker imitates the most successful member of the group for making decisions by copying his cue orders. I think this is quite an accurate identification of a heuristic that we use in real life. No one can deny that copying the actions of a role model is one of the most widely used strategies people use in real life. However, I believe that the simplicity and success of this heuristic is questionable. Just as in the case of the Take-the-Best, the application of Imitate-the-most-successful algorithm requires a lot of judgment. First, it requires one to decide who the most successful member of the group is. Let us say that the problem at hand is choosing the right person to marry. Should we consider the person whose marriage lasted the longest, or the person whose marriage was relatively shorter but happier as the most successful member of the group? As can be seen, even the selection of the role model presents a decision problem. Let us assume that our decision maker has picked his role model and is ready to copy his decisions. But how do we know that the cues or decisions that led him to good decisions is going to help our decision maker as well? If our decision maker comes from a different background from his role model, with different needs and expectations, then following the footsteps of his role model might lead him to undesirable results. For example, his role model might have come from a low-income family and chose to select his spouse based on her income. If, on the other hand, our decision maker comes from a wealthy family than the income of his future wife could be irrelevant to his decision. Similarly, when the role model was making his decisions, maybe all of his alternatives were well-educated, thus the level of education was not a criterion he could use in his decision. However, the education level might vary among the alternatives that our decision maker has, and thus in his case the education level cue would be relevant to his decision. As is apparent in this example, a close examination of the success behind the role model is necessary before blindly following him. Thus, deliberate reasoning with a keen attention to cause and effect between decision factors and the outcomes is very important even for the application of the simplest heuristics.

The difficulty underlying the application of heuristics points to a more general problem in the adaptive toolbox theory: strategy selection. Gigerenzer (2000) explains that an adaptive toolbox is available to the decision maker from which he selects the heuristic that best suits the structure of the environment. For instance, in the example I gave above, the decision maker had to compare his background and desires with that of his role model before making the decision to imitate him. As this example illustrates, the process of checking the compatibility between the environment and the heuristic and finding the balance between "conflicting motivations and goals" requires a thorough and complicated analysis which can hardly be called an easy computation. Thus, it seems that the fast-and-frugal heuristics owe their simplicity to the cognitively costly complex mechanism behind strategy selection.

I believe that strategy selection constitutes the core of the adaptive toolbox theory. Although the researchers of the bounded rationality program also acknowledges that it is central to their theory, so far their research has primarily focused only on identifying and examining heuristics. Gigerenzer and Selten (2001b) explain that "the bundle of heuristics in the adaptive toolbox is orchestrated by some mechanism reflecting the importance of conflicting motivations and goals." (p.9) However, without offering any ideas as to how this mechanism might be working, they simply state that "This mechanism is not yet well understood" (Gigerenzer & Selten, 2001b, p.9). As the bounded rationality program rests not on the idea that people simply use heuristics, but that they use the *right* heuristic in the *right* environment, I believe that their theory will be far from complete until they address the issue of strategy selection.

Intuitions Can Be Misleading

I agree with Gigerenzer that intuitions should not be completely dismissed from the decision process and that sometimes they can offer valuable insights about the human decision making. Although the study of gut feeling can be very helpful for descriptive purposes, I believe that their use cannot be enforced, especially for decisions with high costs. Just as we all had experiences where our intuitions led us to the right choice, we also all know that sometimes our intuitions point in the wrong direction. Therefore, intuitions can only provide a hint, but not conclusive evidence for a decision.

In the bounded rationality program emotions are seen as one of the mechanisms that our intuitions manifest themselves in our decisions. The decision about getting married is an example Gigerenzer (2001a) often gives to illustrate the importance of emotions and the inadequacies of the expected utility theory in real life decision problems. He states that people do not make the decision to marry their loved ones by calculating their utilities. Also, it is impossible to check all the possible alternatives and make a comparison among them to arrive at the perfect decision, as the optimization models dictate. He claims that these models are not realistic and that in real life love acts as a stopping cue in the mate selection process leading us to make the decision to get married. Although the suggestion that love could be a decisive factor in choosing our spouse is quite reasonable, it is hard to argue that it always leads to good decisions.

Despite the fact that our emotions have an undeniable effect on our judgments, I do not think that their effects can all be called advantageous. For example, fear sometimes leads us to overestimate risk and miss important opportunities. Traumatic events distort our perception of the world and prevent us from making objective judgments. For instance, let us assume that the mother in the previous example has heard that dozens of children have died by falling from the trees, but she has personally watched a child being eaten by a crocodile. Witnessing such a traumatic event might lead her to overestimate the danger of playing in the river and let her son climb the trees instead. Similarly, prejudices can greatly distort one's judgment. For instance, the police officer at the airport could be prejudiced against the minorities and thus have a tendency to overestimate their probability of being drug couriers. Therefore, solely relying on gut feelings, especially in legal matters, can have quite dangerous implications.

Gigerenzer's misconception about intuitions being the correct guide in life comes from his incorrect belief that evolution has perfected our decision making skills. It is reasonable to argue that evolution equipped us various reasoning skills that have helped us to survive. However it is not necessarily the case that these reasoning skills have evolved to perfectly fit to the requirements of the environment or that the environment in which they have evolved in has stayed stable. For example, stress is an important mechanism that we have acquired through evolution that helped us deal with emergencies, such as the attack of a lion, by inducing fight-or-flight responses. However, the physiological and psychological effects that stress elicits in us, which were incredibly helpful in ensuring our survival in the wilderness, no longer meets the needs of the modern human. In fact, most of the time stress causes more harm than good; the anxiety, fear and aggression that stress induces in us severely impairs our judgment and decision skills. (see Sapolsky, 1998) Therefore, even if it is true that human decision making can be explained by heuristics that have evolved to fit some environments, it does not follow that these heuristics still constitute successful means of coping with the decision problems of the modern age.

The Adaptive Toolbox

As I have explained in the first chapter, the bounded rationality program asserts that we have a collection of heuristics, called the adaptive toolbox, for making decisions instead of a 'general-purpose decision-making algorithm'. I agree with the idea that we employ various strategies to make decisions in life. However I am doubtful that the adaptive toolbox provides a good model for human cognition or that it is the product of evolution.

I disagree that our brains come with a full set of strategies. The problem solving faculty of our brains could be functioning like our language faculty; we are born with the capacity to learn language, not with the rules and the lexicon of a language. Like the way we learn language, we might be born with the capacity to develop and learn strategies. In addition, it is not clear why the various strategies we adopt in life should be treated as distinct entities that have evolved separately from each other. It is possible that there exists a single decision making algorithm which has parameters allowing it to be modified to fit the needs of the decision problem at hand that gives rise to all the strategies.

Another problem with the claim that heuristics are innate is that it predicts uniform behaviour and decisions among people. For example, as I have explained before, Gigerenzer (1994, 2000) claims that the human Bayesian reasoning algorithm has evolved to work on natural frequencies. If our minds have an evolved capacity to work with natural frequencies, then how is it possible that there are people who fail to solve the questions even when they are given in the natural frequency format?

I also do not believe that the heuristics have evolved to exploit the structure of the environment. This claim is based on the misconception that the heuristics are anchored in the environment. Contrary to their claim, the use and the success of most of the heuristics that they have suggested so far are not anchored in the environment, but in the decision maker's perception of or knowledge about the environment.[14] Let us assume that two members of a species use the same heuristic in the same environment. Depending on their subjective knowledge and perception of the environment, the algorithm might yield a winning decision for one of them and an unsuccessful one for

[14] For example, the outcome of the Take-the-Best heuristic depends on the decision maker's subjective cue rankings which reflect his own knowledge and perception about the environment. The use of the recognition heuristic and its success depends on whether the decision maker has had prior exposure to the environment or not.

the other. For example, in the population guessing game both of the decision makers might be using Take-the-Best. However one might be using "presence of a university" as the most valid cue and the other one "being a capital." Similarly, they might both be using the imitate-the-most-successful heuristic, but copying different members of the group. Therefore, two members of the same species might arrive at two different decisions using the same heuristic in the same environment. As a result, performance of a heuristic cannot be uniform among the members of the same species. When the performance of a strategy is unstable, it is not possible to talk about its success in adapting to a certain environment. Moreover, for a heuristic to evolve adaption to an environment, it is necessary for it to outscore all the other heuristics in its environment.[15] However, since the outcome of a heuristic is based on the decision maker's knowledge about the environment, two different strategies do not necessarily yield two different solutions to a problem. When the performance of two heuristics in the same environment cannot be differentiated, it is not possible to argue that natural selection can favor one over the other.

Relying on evolutionary arguments to explain the success of heuristics also proves to be a problematic approach. The bounded rationality program asserts that a heuristic is successful to the degree that it is adapted to the structure of the environment. However, as Kenneth Hammond (2007, p.234) points out, this definition of success leads to a tautology; a heuristic is said to be successful if it is well adapted to the environment, but at the same time, it is assumed that it cannot be adapted unless it is

[15] For example, in evolutionary game theory, a strategy is assumed to have evolved if no other strategy can outscore it in its environment: "A strategy is *collectively stable* if no strategy can invade it" (Axelrod, 1990, p.56).

successful. In other words, success is defined by the well-adaptedness and well-adaptedness is assumed to be a consequence of success.

CHAPTER 4

CONCLUSION

Gigerenzer's bounded rationality program has filled a large gap in the area of decision making by launching a comprehensive research on natural heuristics. For example, thanks to the studies of the researchers from the bounded rationality program, one-reason decision making has become a widely used model in the decision making field. Furthermore, by initiating a re-examination of the norms, Gigerenzer has helped new insights to be gained about the nature of cognitive fallacies and made valuable contributions to the study of judgment and decision-making. However, I think that Gigerenzer was too hasty in combining the study of heuristics with that of norms and overestimated the role of heuristics in cognition.

Although this innovative approach to the study of heuristics has led to very fruitful discussions in the literature, I believe that Gigerenzer and research group have not been able to provide convincing evidence to prove that heuristics can effectively prescribe norms. They argue that their studies have shown that heuristics can be quite effective, which indicate that they are suitable for prescriptive purposes as well. I have tried to show that the heuristics whose efficiency have been demonstrated are not "boundedly rational" because their real-life execution requires comprehensive domain knowledge and complex cognitive skills. Furthermore, the success of these heuristics are contingent upon other factors, and thus their use cannot be prescribed without first addressing these issues.[16] Conversely, the heuristics which are boundedly rational, such as emotions, are in return unreliable for guaranteeing accuracy.

Additionally, even if it is true that we have heuristics that can help us solve problems, it is not clear why we should not strive to develop methods that are even more successful. There are many decision problems in the world for which the pressure for accuracy surpasses the need for simplicity and speed. Such circumstances call for algorithms that can guarantee optimal or near-optimal solutions rather than fast-and-frugal algorithms.

Gigerenzer's discussion of the cognitive fallacies has shown us that rational norms should not be applied without paying close attention to the social context to the confusion that the ambiguity in language can cause. However, when precautions are taken to avoid these confusions, it seems that the normative benchmarks of classical

[16] Such as how we rank cues for Take-the-Best, or determine that recognition is highly correlated with a criterion before using the recognition heuristic.

rationality, such as logic and probability theory, make up a solid basis for the norms. Therefore, I believe that comparing human reasoning to the classical norms still constitutes a viable method for studying human rationality.

Lastly, in my thesis I have questioned the claim that heuristics are the end result of evolution. I have argued that evolution could not have given rise to domain specific heuristics and claimed that instead we might be born with a capacity to learn and develop new strategies, rather than a fixed collection of heuristics.

Future work on the bounded rationality program should focus on investigating the specific conditions and environments that heuristics perform well in, as well as heuristics themselves, because unless they do so, their theory becomes impossible to test scientifically. The proponents of the bounded rationality program should also try to elucidate the execution of the heuristics more specifically, for example by clarifying how information from several resources can be combined to construct a single cue ranking. Finally, since their theory rests on the idea of heuristic selection, and judging the suitability of a heuristic to a certain problem is itself a very complex decision, they should focus on discovering the mechanism behind this process.

REFERENCES

Allwood, C. M., & Montgomery, H. (1987). Response selection strategies and realism of confidence judgments. *Organizational Behavior and Human Decision Processes*, *39,* 365-383.

Axelrod, R. (1990). *The evolution of co-operation.* London: Penguin Books.

Ayton, P., & Önkal, D. (1997). *Forecasting football fixtures: Confidence and judged proportion correct.* Unpublished manuscript, Department of Psychology, The City University of London.

Barnard, G. A. (1979). Discussion of the paper by Professors Lindley and Tversky and Dr. Brown. *Journal of Royal Statistical Society, Series A*, *142*, 171-172.

Barsalou, L. W., & Ross, B. H. (1986). The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 116-134.

Birnbaum, M. H. (1983). Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology, 96* ,95-94.

Borges, B., Goldstein, D. G., Ortmann, A., & Gigerenzer, G. (1999). Can ignorance beat the stock market? In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (p.59–72). New York: Oxford University Press.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics.* New York: Oxford University Press.

Dieckmann, A., & Todd, P. M. (2004). Simple ways to construct search orders. In K. Forbus, D. Gentner, & T. Regier (Eds.)*, Proceedings of the 26th Annual Conference of the Cognitive Science Society* (p. 309–314). Mahwah, NJ: Erlbaum.

Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, *115*, 199–213.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (p. 249-267). New York: Cambridge University Press.

Finetti, B. de (1989). Probabilism. *Erkenntnis*, *31*, 169-223. (Original work published 1931)

Galef, B. G. (1987). Social influences on the identification of toxic foods by Norway rats. *Animal Learning & Behavior*, *15*, 327–332.

Garcia-Retamero, R., Takezawa, M., & Gigerenzer, G. (2006). How to learn good cue orders: When social learning benefits simple heuristics. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (p. 1352–1358). Mahwah, NJ: Erlbaum.

Garcia-Retamero, R., Wallin, A., & Dieckmann, A. (2007). Does causal knowledge help us be faster and more frugal in our decisions? *Memory & Cognition*, *35*, 1399-1409.

Gibson, R. M. (1996). Female choice in sage grouse: The roles of attraction and active comparison. *Behavioral Ecology and Sociobiology*, *39*, 55–59.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is relevant for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (p. 129-161). New York: Wiley.

Gigerenzer, G. (1996a). On narrow norms and vague heuristics: a reply to Kahneman and Tversky. *Psychological Review*, *103*, 592-596.

Gigerenzer, G. (1996b). Rationality: Why social context matters. In P. B. Baltes & U. M. Staudinger (Eds.)*, Interactive minds: Life-span perspectives on the social foundation of cognition* (p. 319-346). Cambridge: Cambridge University Press.

Gigerenzer, G. (2000). *Adaptive thinking: rationality in the real world*. New York: Oxford University Press.

Gigerenzer, G. (2001a). The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (p. 37-51). Cambridge, MA: MIT Press.

Gigerenzer, G.(2001b). Content-blind norms, no norms, or good norms? A reply to Vranas. *Cognition*, *81*, 93-103.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (p. 62-88). Oxford, UK: Blackwell.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.

Gigerenzer, G. (2008). *Rationality for mortals: how people cope with uncertainty.* New York: Oxford University Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650-669.

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

Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas. *Psychological Review*, *115*, 230-237.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*, 506-528.

Gigerenzer, G., & Selten, R. (Eds.). (2001a). *Bounded rationality: The adaptive toolbox.* Cambridge, MA: The MIT Press.

Gigerenzer, G., & Selten, R. (2001b). Rethinking rationality. In G. Gigerenzer & R. Selten, (Eds.), *Bounded rationality: The adaptive toolbox* (p.1-13). Cambridge, MA: MIT Press.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (p.3-37). New York: Oxford University Press.

Gigerenzer, G., Todd, P. M., & ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice*, *45*, 219-26.

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (p. 37–58). New York: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75-90.

Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge: Cambridge University Press.

Hajek, A. (2007). The reference class problem is your problem too. *Synthese*, *156*, 563-568.

Hajek, A. (2009). Interpretations of probability. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved May 20, 2009, from http://plato.stanford.edu

Hammond, K. R. (2007). *Beyond rationality: The search for wisdom in a troubled time*. New York: Oxford University Press.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356-388.

Hintzman, D. L., & Block, R. A. (1972). Repetition and memory: Evidence for multiple trace hypothesis. *Journal of Experimental Psychology, 88,* 297-206.

Jonides, J., & Jones, C. M. (1992). Direct coding for frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 368-378.

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

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591.

Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives*. Cambridge, England: Cambridge University Press.

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

Martignon, L. (2001). Comparing fast and frugal heuristics and optimal models. In G. Gigerenzer & R. Selten, (Eds.), *Bounded rationality: The adaptive toolbox* (p.147-171). Cambridge, MA: MIT Press.

Martignon, L., & Hoffrage, U. (2002). Fast, frugal and fit: Lexicographic heuristics for paired comparison. *Theory and Decision*, *52*, 29–71.

McLeod, P., & Dienes, Z. (1996). Do fielders know where to go to catch the ball or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 531-43.

Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration*.* *Psychological Science*, *12*, 269-275.

Newell, B. R., Rakow, T., Weston, N. J., Shanks, D. R. (2004). Search strategies in decision making: The success of "success". *Journal of Behavioral Decision Making, 17*, 117-137.

Newell, B. R., Weston, N. J., & Shanks, D. R. (2003).Empirical tests of a fast and frugal heuristic: Not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes*, *91*, 82-96.

Over, D. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (p. 3-19). Oxford, UK: Blackwell.

Raiffa, H. (2002). Decision analysis: A personal account of how it got started and evolved. *Operations research, 50,* p.179-185.

Salmon, W. (1966). *The Foundations of Scientific Inference*. Pittsburgh, PA: University of Pittsburgh Press.

Sapolsky, R. (2004). *Why zebras don't get ulcers*. New York : Owl Book/Henry Holt and Co.

Selten, R. (2001). What is bounded rationality? In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (p.13-37). Cambridge, MA: MIT Press.

Simon, H.A. (1957). *Models of Man*. New York: Wiley.

Simon, H.A. (1983). *Reason in Human Affairs*. Stanford, CA: Stanford University Press.

Sternberg, R. J. (2006). *Cognitive Psychology* (4th edition). Belmont, CA: Thomson.

Teigen, K. H., Brun, W., & Frydenlund, R. (1999). Judgments of risk and probability: The role of frequentistic information. *Journal of Behavioral Decision Making, 12,* 123-139.

Tversky, A., & Kahneman, D. (1980). Causal schemata in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. 1, p. 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* (p. 153-160). New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.

Von Mises, R. (1957). *Probability, Statistics and Truth*, revised English edition, New York: Macmillan.

Warrington, E. K., & McCarthy, R. A. (1988). The fractionation of retrograde amnesia. *Brain and Cognition*, *7*, 184–200.