Gredeg
Groupe de recherche en Droit, Economie et Gestion

UNIVERSITÉ CÔTE D'AZUR

THE LICHTENSTEIN-SLOVIC-TVERSKY-KAHNEMAN NEXUS. A PREHISTORY OF BEHAVIORAL ECONOMICS (1969-1974)

Documents de travail GREDEG
GREDEG Working Papers Series

Jean-Sébastien Lenfant

GREDEG WP No. 2024-31
https://ideas.repec.org/s/gre/wpaper.html

The Lichtenstein-Slovic-Tversky-Kahneman Nexus. A Prehistory of Behavioral Economics (1969-1974)

Jean-Sébastien Lenfant
PRISM, Université Paris 1 Panthéon-Sorbonne
jean-sebastien.lenfant@univ-paris1.fr

December 1, 2024

Abstract

The purpose of this article is to provide a historical account of the contributions to judgment and decision making by four cognitive psychologists at the turn of the 1970s: Sarah Lichtenstein, Paul Slovic, Amos Tversky and Daniel Kahneman. Beyond the usual focus on Kahneman and Tversky's heuristics and biases approach, we uphold that historians of behavioral economics would gain from a broader and more balanced view of the contributions of these four psychologists to the theory of decision making. Together with the heuristics and biases approach, experiments on preference reversal and choice intransitivities represent a multifaceted criticism of standard theories of choice and decision against which the genesis of behavioral economics could be evaluated.

Keywords: Lichtenstein (Sarah), Slovic (Paul), Tversky (Amos), Kahneman (Daniel), heuristics and biases, preference reversal, intransitivity, preferences, behavioral economics, conjoint measurement, judgment, expected utility theory, mathematical psychology, cognitivism, experiments

JEL: B21, B29, D91

1 Introduction

Behavioral economics emerged in the 1980s as a set of methods and tools to provide a more realistic basis—or simply an empirically richer basis—for economic modeling and theorizing of individuals' economic behavior.[1] Over the past decades,

[1]Acknowledgments. A preliminary sketch of the present paper had been presented on the occasion of a conference honoring Richard Arena's career that was held in Nice, 19-21 May 2022. In the process of writing the article, it eventually turned out that its length made it unfit for a contribution in a collective book (another article will be published instead). I am deeply indebted to Paul Slovic for providing detailed comments on earlier drafts of this paper and a continuous support for this project of a prehistory of behavioral economics. I also benefited from the feedback of Jan Horst Keppler and Julien Gradoz whose comments have been valuable for improving the paper. My deepest thanks also to Barbara Tversky for providing precious references and support. I bear full responsibility for any remaining mistake.

behavioral methods in economics have been applied to subfields of economics (e.g. health economics, development economics, law and economics) and the approach is now officially represented in JEL classifications such as “Micro-based Behavioral Economics” (D9), “Macro-Based Behavioral Economics” (E7), “Behavioral Finance” (G4). This behavioral turn in economics highlights how specific institutional and agency settings as well as social norms can shape individuals’ behaviors and economic outcomes. Thus, behavioral economics can be described in general terms, as a collection of approaches extending the neoclassical paradigm:

> "Behavioral economics is an umbrella of approaches that seek to extend the standard economics framework to account for relevant features of human behavior that are absent in the standard economics framework." (Diamond and Vartiainen, 2012, 1)

Behavioral economics is rooted in three relatively independent streams of research about economic agency. The first focuses on the cognitive limits of human beings, the second deals with social preferences and norms and their effect on individual behavior in various environments, the third deals with broad evolutionary aspects of behavior and stresses ecological aspects of economic rationality. All these streams of research are themselves the result of various disciplinary interactions throughout the second half of the 20th century, involving economics, psychology, biology, sociology, philosophy, and behavioral sciences (Truc, 2022a,b).

However, it is generally accepted that the main impetus for the advent of behavioral economics as an institutionalized field of research stems from the contributions of cognitive psychologists to the study of judgment and decision-making in the 1970s, and the names of Tversky and Kahneman are firmly associated with this achievement (Heukelom, 2009, 2014, 2011). Indeed, during the 1970s, Kahneman and Tversky[2] laid the foundations for a new approach, the "heuristic and biases" approach (Tversky and Kahneman, 1974), and they used it to reformulate the theory of rational behavior under risk or uncertainty, namely "prospect theory" (Kahneman and Tversky, 1979). The present paper aims to enrich this narrative by widening the picture, including two other psychologists into the prehistory of behavioral economics: Sarah Lichtenstein and Paul Slovic. To date, few historians of economic thought have contributed to the early history of this huge and rich facet of contemporary and recent economics (Heukelom, 2014; Moscati, 2018; Nagatsu, 2015). Their contributions coexist with self-made histories by key protagonists (sometimes with methodologists), whose aim is to explain retrospectively the present state of their field, as they see it, and possibly to provide a broader perspective with the intellectual origins of behavioral economics (Thaler, 2016; Camerer and Loewenstein, 2004; Angner and Loewenstein, 2007).[3]

Whatever the proper balance between various accounts of rationality in behavioral economics, the fact is that a large part of economic theory is organized around a

[2] In this article, the ordering of the names usually depends on the context. Each time we are commenting a joint-paper by the two authors, we use the order as it appears in authorship. When they started collaborating, Kahneman and Tversky tossed the order of their names as it would appear on their first published article. Then, they alternated the order in subsequent articles.

[3] Camerer and Loewenstein (2004) propose to interpret behavioral economics within a much longer historical perspective stretching back to Hume, Smith and Bentham

view of market coordination in which agents are assumed to make a choice based on a set of data (alternatives, constraints) which "are not themselves the objects of rational calculation" (Simon, 1955, 100). In standard theories of riskless choice and in theories of decision, it is assumed that a number of data are fixed and well-identified (tastes, technologies, prices, quantities, alternatives, probabilities, outcomes) and that each agent knows which are under his control, and this set is assumed to be self-sufficient in order to identify an optimal behavior. These views about rational behavior would later be labeled as "substantive rationality" by Simon (1976) and the models in which a set of goals and constraints is unambiguously defined are sometimes labeled as models of "parametric rationality" in contrast with models of "strategic rationality". In two famous articles, Ward Edwards (1954c; 1961) introduced psychologists to theories and models of rational behavior used by economics, and delimited the territory on which they could contribute to improve economic modeling, drawing attention to insuperable difficulties to think about rational behavior in strategic or dynamic environments (see also Thrall et al. (1954); Simon (1955, 1959)). Subsequently, cognitive psychologists interested in issues about rational behavior in the 1960s focused on problem-solving situations and on the use of information to make judgments and decisions, and to this effect, they explicitly used as their playground parametric models of choice and decision elaborated by economists.[4]

In this respect, Tversky and Kahneman as co-authors of landmark contributions in the 1970s deserve to be put on a pedestal. Nevertheless, historical works that focus exclusively on the Tversky-Kahneman collaboration run the risk of personifying the story at the cost of ignoring something of its meaning within a broader set of contributions. A story of behavioral economics should be based on a reconstruction of the criticism and reshaping of the theory of rational behavior, and for this purpose, a broader set of protagonists and works should be integrated into the picture. This article argues that research done in psychology of judgment and decision by a group of four psychologists, Sarah Lichtenstein, Paul Slovic, Amos Tversky and Daniel Kahneman within a short time span (circa 1969-1974) provides a complete criticism of the then reigning theory of rational behavior in neoclassical economics as established by the likes of Hicks-Allen-Samuelson and von Neumann-Morgenstern-Savage.

Notably, I uphold that through a more balanced analysis of the contributions of these four cognitive psychologists, one gets a finer sense of the consistency of their approach of rational behavior. Instead, historical accounts that tend to ignore this period, taking prospect theory as their starting point, are biasing the history of behavioral economics, especially as regards the rendition of the relationships between

[4]It is beyond the scope of this paper to discuss the differences between an "old" behavioral economics associated to Simon and a "new" behavioral economics associated to Kahneman and Tversky (Sent, 2004). Be it enough to mention that Simon is definitely more interested in building new theoretical models based on a general criticism of models of substantive rationality while the cognitivist approach launched by Edwards is primarily concerned with accumulating experimental results based on controlled experiments involving information treatment and problem solving. Once this is recognized, it is quite understandable that the starting point for behavioral economics should be a critical assessment of models of "substantive" or "parametric" rational behavior and of the explicit or implicit assumptions about agents' abilities on which they are grounded.

psychology and economics that was—and maybe still is—so critical for the advent of behavioral economics. Prospect theory itself (Kahneman and Tversky, 1979) is commonly viewed as the first specific intrusion of cognitive psychology into the realm of the economic modeling of behavior under risk and uncertainty. it would would be better analyzed as the first move of cognitive psychologists towards the community of economists and their modeling practices, a move implying a selective use of the knowledge accumulated during the previous decade.[5]

Most often, what is retained from this early literature in cognitive psychology (end of the 1960s, beginning of the 1970s) is one single synthetic contribution to the heuristic and biases approach by Kahneman and Tversky "Judgment under Uncertainty: Heuristics and Biases" published in *Science* (Tversky and Kahneman, 1974) (hereafter "Heuristics and Biases"). Other contributions are simply ignored. In the following, I argue in favor of taking seriously a wider set of contributions. Notably, the present paper aims at treating Tversky's contribution to intransitivity of preferences (Tversky, 1969) and Lichtenstein and Slovic's contributions to the preference reversal phenomenon (Lichtenstein and Slovic, 1971, 1973) on an equal footing with Tversky and Kahneman's series of co-authored articles (1971; 1973; 1973) leading to their much-cited article "Heuristics and Biases".[6]

Compared with Heukelom's account of this literature (Heukelom, 2014, 2012), the present article is focused on the period 1969-1974, during which I contend that a complete deconstruction of models of choice and decision was accomplished. Hence, I deliberately make very few references to later developments in the history of behavioral economics in order to avoid a retrospective bias. It is hoped that the rational reconstruction of the research proposed here allows to question why certain contributions would attract more attention than others and why some have been ignored in the development of behavioral economics. Acclimatization of the contributions of cognitive psychology to economic theorizing and modeling started in the second half of the 1970s and is reflected in the specific outcome of prospect theory as Kahneman and Tversky conceived it. It deserves a separate historical inquiry.[7]

It is hoped that this rendition of the prehistory of behavioral economics can serve historians to highlight later orientations and choices made by economists or by psychologists in their effort to develop behavioral economics (or to oppose it). Behavioral economics being also partly the result of an interdisciplinary collaboration, it is most important to understand what economists have been likely to accept as traveling concepts and methods from one discipline to another, but also to con-

[5] It will be the subject of another article.

[6] Taking as a sample some twenty recent historical and "self-made" histories, it turns out that Tversky's contribution to intransitivity is almost ignored and that the preference reversal phenomenon receives uneven attention, as well as the series of contributions by Tversky and Kahneman in 1971-1973. Letting aside some contributions to the history of behavioral economics that do not mention any of these references, not even Tversky and Kahneman's heuristic and biases paper (Hattwick, 1989; Tomer, 2007; Berg, 2010; Levine, 2012), most often general overviews on behavioral economics contain references only to Tversky and Kahneman (1974) and to one contribution by Lichtenstein and Slovic to preference reversal (Weber and Dawes, 2005; Angner and Loewenstein, 2007; Thaler, 2016; Cartwright, 2018). Wilkinson and Klaes (2017), Angner (2020) and Nagatsu (2015) mention at least one contribution from Tversky-Kahneman co-authorship dated 1971-1973, but not Tversky (1969).

[7] See for instance Lewis (2016, chapter 9 and 10) and Heukelom (2012).

sider how cognitive psychologists themselves were deviating from usual streams of research in their own discipline. Indeed, in the 1950s and 1960s, the field of cognitive psychology and mathematical psychology was gaining a foothold in the United States, but the dominant view was that somehow—as far as a static theory of behavior is concerned—people were able to use information in a coherent way, assessing and comparing the overall value of alternatives when making choices and decisions. This state of beliefs would change by the end of the 1960s, and Lichtenstein, Slovic, Tversky and Kahneman were the most active promoters of this change.

The notion of a substantive or parametric model of rational choice can be represented as a set of of three building blocks of individual abilities :

- First, appraising the environment, what economists would call the initial constraints (like exogenous income and prices in the consumer's choice problem, or initial endowments and prices in a general equilibrium situation, or payoffs and probabilities in a risky decision problem).
- Second, valuating the various situations (initial state, possible final states), which for economists pertains to establishing a preference ordering over states or outcomes.
- Third, choosing the best final state among the various (possibly infinite) available states, which is summarized in economics as "optimizing", requiring computing abilities.

The unity behind different contributions to decision making by Lichtenstein, Slovic, Tversky and Kahneman pertains to their careful destroying of *all* these three facets of the theory of rational behavior. In an often-cited research paper, "From Shakespeare to Simon: Speculations—and some Evidence—about man's ability to process information", Paul Slovic (1972) makes a synthesis of then current research on judgment and decision making. As he points out, the state of experimental knowledge is limited to exploring simple problem-solving situations and does not cope with complex situations involving composition of beliefs and actions at the collective level. Slovic offers, I think, a similar understanding of models of rational behavior as being made of these three building blocks of individual abilities :

> "First, [the decision maker] wonders what will happen or how likely it is to happen, and his use of information to answer these questions gets him involved in processes which we call inference prediction, subjective probability, and diagnosis. He must also evaluate the worth of objects, and this often requires him to combine information from several component attributes of the object into an overall judgment. Finally, he is called upon to integrate his opinion about probabilities and values into the selection of some course of action. What is referred to as 'weighing risks against benefits' is an example of the latter combinatorial process." (Slovic, 1972, 2)

This synthetic presentation of the cognitive assumptions in standard models of rational choice lies behind the research done by Lichtenstein, Slovic, Tversky and

Kahneman. The aim of the present article is precisely to highlight the unity behind the multifaceted criticism of economic models of choice and decision by cognitive psychologists at the turn of the 1970s and to make of it a starting point for further studies in the building of behavioral economics from the mid 1970s onwards.

As will be quite evident to the reader, the protagonists of our story know each other and are well aware of their work, as is evidenced by quotations or references in their articles. However, before to embark in this long review of the advent of the psychology of judgment and decision making, it seems important to make clear that Sarah Lichtenstein, Paul Slovic, Amos Tversky and Daniel Kahneman never committed themselves to build a new theory of rational behavior, and that they did not even think of their research as a deliberate collective undertaking to destroy standard models of economic rationality. Models were taken first and foremost as providers of frameworks and relations that could be brought to test. At the university of Michigan, and even more at the Oregon Research Institute in the early 1970s, exchanges between Kahneman and Tversky on the one hand and Lichtenstein and Slovic on the other were mostly informal, and they did not intervene in any way in the design of their experiments. Slovic's paper "From Shakespeare to Simon" is the only contribution that can be read as an effort to synthesize the research done on judgment and decision making during this period, and to draw perspectives for further research, but it was not published in an academic journal. [8]

The remainder of this article is as follows. Section 2 presents the research done on decision making at the university of Michigan in the 1960s under Ward Edwards and Clyde Coombs' leadership. Notably, I point out the progressive awareness by Lichtenstein, Slovic and Tversky that the usual compensatory models of choice and decisions are unfit to describe decision making processes and individuals' abilities to make calculations and value alternatives. Section 3 presents and discusses Tversky's contribution to the understanding of intransitive behavior. Section 4 presents and discusses Lichtenstein and Slovic's work on the preference reversal phenomenon. Section 5 is devoted to the construction of the heuristic and biases approach proposed by Kahneman, Tversky and Slovic. Notably, it focuses on the building of a heuristic approach as a general level theory distinct from biases. Section 6 draws the implications of their research for a number of key dimensions of the theory of

[8] I am most thankful to Paul Slovic for clarifying the kind of exchanges and interactions that took place during this period, and especially during academic year 1971-1972, when Tversky and Kahneman were invited at Oregon Research Institute and the following years. After the series of articles on heuristics and biases, there have been occasional collaborations with Slovic. Slovic co-authored two articles with Tversky (on the Savage Axiom (Slovic and Tversky, 1974) and on contingent weighting (Tversky et al., 1988)). He also coauthored with Kahneman and Tversky for editing a book of seminal contributions on judgmental heuristics, *Judgment under Uncertainty: Heuristics and Biases* (1982) and for a synthesis on preference reversal (Tversky et al., 1990). As Paul Slovic writes (private correspondence, July 14, 2024): "Clearly, Amos and Danny were tracking the work by Sarah and me, just as we were paying very close attention to theirs. With a few notable exceptions, we were keenly aware of each other's work but we were on complementary parallel paths. We did not intervene directly with substantive advice on one another's experiments or articles. With one exception: Amos thought my 'Shakespeare to Simon' paper was good but needed more work before sending it for publication. I didn't understand his concerns and I never did publish it except as an ORI report. I do think it was an important paper and regret never publishing it."

decision making. Section 7 concludes.

2 Psychologists and decision making: the birth of a cognitive perspective (1950s-1960s)

Behavioral decision theory as a field of research in psychology emerged in the 1950s with leading figures Ward Edwards and Clyde Coombs at the university of Michigan. Other prominent figures in decision theory came from mathematics and philosophy departments (Patrick Suppes and Donald Davidson at Stanford, Duncan Luce at Columbia, then at Harvard). The specificity of Edwards and Coombs, however, lies in their mixed interest in mathematical modeling and experiments. The field of behavioral decision was thoroughly transformed in the second half of the 1960s and beginning of the 1970s by their students Sarah Lichtenstein, Paul Slovic and Amos Tversky, who pioneered new ideas about the way people select and process information in various contexts of choices and judgment, notably in risky environments. Their work on decision making covered methodological as well as conceptual aspects and articulated highly abstract inquiries in mathematical psychology and sophisticated experimental investigations. More precisely, as I hope to show below, by the end of the 1960s, new experimental facts, ideas, conjectures, methodological tenets had been accumulated that would come to full fruition in the next decade (1969-1979). The aim of this section is to document this prehistory of behavioral decision theory, setting the stage for Slovic, Lichtenstein and Tversky. In short, I shall argue that the initial tracks for psychological research on choice and decision were defined by Ward Edwards and Clyde Coombs. After presenting their work, I show how Lichtenstein, Slovic, Tversky and some others would progressively moved away from their approach, laying the groundwork for a deeper cognitive approach to decision making, focusing more systematically on information-processing aspects of choice and judgment.[9] In the course of this historical record, I shall stress some comments made by the researchers about the issue of rationality (notably transitivity) and normativity in the theory of decision.

2.1 Ward Edwards's contributions to the psychology of decision making

In the 1950s and 1960s, Ward Edwards established two research agendas linked to decision making. The first deals with an elicitation of the relevant parameters that induce individuals' judgments and decisions in risky environments, notably when these environments can be represented as gambles with probabilities of wins and losses. The second, started in the 1960s, is about individuals' learning and perception of probabilities, i.e. Bayesian decision theory.[10]

[9] Kahneman would join the team only in a later stage of this story (see section 5).

[10] Ward Edwards (1927–2005) studied psychology at Swarthmore College, then graduated at Harvard and obtained his PhD in psychology (under Fred Mosteller's supervision). Fred Mosteller introduced him to the expected utility model of von Neumann and Morgenstern. His PhD contains a first criticism of the expected utility model, objecting to the assumption that probabilities are

Edwards' research agenda in the 1950s is set out in a well-informed and influential exposition of the theory of rational choice in economics: "The theory of decision making" (1954c). In this article, Edwards provides for the first time in a psychology journal an overall presentation of the theory of riskless choices (à la Hicks-Allen) and of the theory of choice under risk and uncertainty (expected utility theory and subjective expected utility theory, hereafter EUT and SEU). He identifies research topics for experimental psychology based on these economic theories of rational (parametric) behavior.[11] He made the subject a field of inquiry for psychologists.

This is the starting point for intensive research on theories of rational choice by psychologists (Edwards, 1969). In the field of riskless choice, a commonplace observation to a psychologist is that "human beings are neither perfectly consistent nor perfectly sensitive" (Edwards, 1954c, 388). As regards EUT, psychologists were first interested in obtaining experimental utility functions for lotteries (Mosteller and Nogee, 1951). But the model contains several behavioral assumptions that need scrutinizing.

> "If [the EUT] model is to be used to predict actual choices, what could go wrong with it? It might be that the probabilities by which the utilities are multiplied should not be the objective probabilities; in other words, a decider's estimate of the subjective importance of a probability may not be the same as the numerical value of that probability. It might be that the method of combination of probabilities and values should not be simple multiplication. It might be that the method of combination of the probability-value products should not be simple addition. It might be that the process of gambling has some positive or negative utility of its own. It might be that the whole approach is wrong, that people just do not behave as if they were trying to maximize expected utility." (Edwards, 1954c, 393)

At first, Edwards' agenda was to test some assumptions contained in EUT and SEU, to identify which parameters in a gamble are more important than others to predict preferential choices or even to identify some usual preferential patterns. In 1954, Edwards had already obtained some experimental results in this direction in his PhD dissertation and follow-up articles, estimating that subjects (hereafter *Ss*)

valued linearly in the process of valuing lotteries. He identified that some values of probabilities in a two-outcome bet are preferred to others, thus distorting somehow the evaluation of lotteries. After his PhD, he held a position at Johns Hopkins University and prepared the 1954 review article on decision theory. Fired from the psychology department for not being devoted enough to his teaching duties, he was hired by the Air Force Personnel and Training Research Center headed by Arthur W. Melton, to develop a section dedicated to decision problems. In 1958, he followed Melton to the University of Michigan where Melton established the Engineering Psychology Laboratory. Melton and Paul M. Fitts Jr. later established the Human Performance Center which was to attract Tversky and Kahneman in the 1960s. In 1962, Edwards founded the Bayesian Research Conference (Fryback, 2005; Phillips and von Winterfeldt, 2006).

[11] In the course of preparing his article while at Johns Hopkins university, Edwards benefited from exchanges with Savage, Georgescu-Roegen and Allais, among others and was able to use unpublished material. At the time, Savage's theory circulated as a research paper of the Statistical Research Center at the University of Chicago ("An axiomatic theory of reasonable behavior in the face of uncertainty", dated 1952)

facing various gambles (with positive, negative or zero expected value) exhibited a specific attraction for bets with probabilities $\frac{1}{2}$, a feature that cannot be rationalized in the expected utility model (Edwards, 1953, 1954a,b). Overall, Edwards' research agenda remained driven by the idea that choices among bets are reflecting a valuation of some parameters in the bet and a way to deal with conflicting parameters when it occurs, even though this is not enough for exhibiting a complete transitive ordering of bets. The parameters studied by Edwards are not the direct parameters of a gamble (probabilities of losing and winning and amounts to be gained or lost). He focused first on structural parameters that were at the heart of various decision models in economics: expected value, variance.[12] Identifying new parameters of a gamble as determinants of choices is a first shift toward a cognitive approach to the psychology of decision making.

Instead of concentrating on the assumption that individuals follow a model of choice, like maximizing expected value when choosing among alternative bets, Edwards asserted that the logical question to be studied is "How do people actually go about making decisions in gambling situations?" (Edwards, 1953, 351) and in order to elicit what factors may influence decision other than expected value, one needs to neutralise the assumed motivation to choose according to the expected value. To this effect, Edwards devised an experiment in which *Ss* have to make binary choices among bets with the same expected value. The experiment accounts for the fact that various factors may influence the choice (e.g. playing real money or worthless chips, accounting for accumulated wealth during the experiment, choosing between bets with positive expected value or negative expected value). Behavior in the face of gambles with the same expected value (EV) (p_1, W_1, p_2, W_2) (where W_i denotes a positive monetary gain) is not linear, i.e. there is not a regular tendency to prefer long shots (low probability of winning a high amount) over short shots (high probability of winning a small amount). Instead, there is a tendency for overplaying bets with a probability of winning of $\frac{1}{2}$ and a tendency to underplay bets with a probability of winning of $\frac{3}{4}$. Given that the expected value of the bets is the same, this is interpreted typically as a separate effect, a "probability-preference", i.e. an attraction in favor of bets with a probability of winning of $\frac{1}{2}$ and a diversion from bets with a probability of winning around $\frac{3}{4}$.[13]

In the course of discussing experimental results and the difficulty in making sense of the observed behaviors in a single factorial model, Edwards makes a conjecture about a number of intransitive triads of bets, anticipating future research:

> "It is worth speculating that perhaps intransitive judgments are inevitable features of situations in which the stimuli combine several inconsistent dimensions along any one of which judgments may be made, as do the stimuli in this experiment." (Edwards, 1953, 363)

In other words, Edwards conjectures that the usual assumption that *Ss* can treat probability and outcomes separately, as assumed in EUT, is not acceptable,

[12] Irving Fisher (Fisher, 1906) had suggested that the variance is a preference parameter, as did Allais (1953).

[13] Behavior toward bets with negative expected value is symmetrical, in that *Ss* prefer bets with a low probability of losing a high amount to bets with a certain loss. This preference, however, is weak and not linear with probabilities.

unless *Ss* view probability in a different way to that which is assumed in EUT (see also Edwards (1954b, 76)). Other such phenomena must be involved in choices among bets, and it is the task of the psychologist to identify them before building a quantitative model.

In the following years, Edwards would continue working along these lines. In two follow-up articles, he aimed to specify the meaning of probability-preference and its use in modeling behavior toward risk and discussed the influence of probability-preference when choices are made between bets with differing expected values (1954a). In this case, the motivation for searching the maximum expected value becomes operational, without canceling out completely the effect of a probability-preference. Edwards (1954b) aimed to capture some new variables influencing the relative role of outcomes and probabilities within a choice among bets.

Even though Edwards's agenda was primarily linked to the models of rational choice, it moved away from the search for a utility function for money based on EUT (Mosteller and Nogee (1951), Preston and Baratta (1948)). A direct consequence of his findings is that, in order to obtain some operational and measurable concept of utility, one needs to construct it conditionally on some values of probability distributions, thus neutralizing the probability-preference effect. However, in another article covering this issue, Edwards (1955) elaborated on another idea: to provide a model of choice among bets that would mix a utility function for money outcomes and a notion of subjective probability. To him, such a model fares better than other existing models in predicting observed behaviors and it highlights the importance of subjective utility. (One can see it as an anticipation of prospect theory.) In yet another attempt, Edwards (1962b) mixed assumptions on subjective probability and variance preferences. What surfaces from all these various attempts made by Edwards is an unsolved tension between the search for a reliable descriptive model of decision under risk or uncertainty and the impossibility to account simultaneously for several significant determinants of choice.

In his update of the 1954 review article, Edwards launched the label "Behavioral decision theory" to account for the growing number of studies on rational decision in various disciplines, including discussions of SEU and of behavioral strategic decision (Edwards, 1961) In the same article, Edwards also promoted Bayesian decision theory as a promising field for overcoming the limits of static models of choice.[14] Inspired by the Bayesian approach to statistical inference promoted by Savage's *Foundations of statistics* (1954), Edwards put forward a paper co-authored with his student Harold Lindman and with Leonard Savage (Edwards et al., 1963). This was

[14]To an economist, the idea of assessing a variance preference or a probability preference sounds at odds with the very essence of preference to be a catchall of the total set of characteristics that an individual is likely to consider relevant, even though the idea of treating separately certain such characteristics as independent from one another has been suggested and studied on various occasions. Another, more telling reason for not considering a separate treatment for variance and probabilities in the study of risk behavior comes from the fact that variance—obtained from probabilities and pay-offs information—is involved in the method of measurement of utility. This is the method developed by Davidson, Suppes and Siegel (1957). Here is a first tension in the development of behavioral studies. If variance is involved in the measurement of utility, then testing variance preference would lead us to renounce this measurement, since it becomes a parameter for individual preference.

the starting point for another research thread of Edwards on Bayesian decision theory, experimenting on how *Ss* consider new information to change their views about probability judgments. During the next decade, Edwards published mainly on this area, with a view that *Ss* are poorly willing to adapt their posterior probabilities in view of new information. The main difference between these two areas of research is that the Bayesian approach, in its early developments at least, contains a normative appraisal of *Ss*' failures to make judgments in tune with Bayesian models. In the second half of the 1960s, Edwards was fully involved in Bayesian decision theory and he felt that there was no hope of finding a suitable model for choice under uncertainty. He did not want to enquire on the deviations from EUT and SEU following the path traced by Lichtenstein, Slovic and Tversky.[15]

2.2 Clyde Coombs, scaling and transitivity

Another leading figure in the field of the psychology of choice and decision in the 1950s-1960s was Clyde Coombs. It is necessary for our story to dwell a while on some of Coombs' contributions, namely his contributions on scales of measurement of psychological magnitudes.[16] Following the tradition of psychometrics of Thurstone (1927b; 1927c; 1927a; 1927d), and the reframing of the theory of measurement advanced by Stevens (1946), Coombs undertook to advance a new scale of measurement, an "ordered metric scale". Stevens (1946) had put some order in the long ranging practice of psychologists to measure psychological magnitudes along an assumed continuum of feelings or perceptions. The outcome was a distinction between (a) nominal scales, (b) ordinal scales, (c) interval (or cardinal) scales and (d) ratio scales, each allowing a specific kind of meaningful operations on values scaled.

The distinction between ordinal scales and cardinal scales overlaps the one used by economists since the 1930s and is well identified in choice theory. Consumer theory is based on a theory of choice under certainty which can be expressed completely in ordinal terms, assuming only a complete order of preferences of goods, while the theory of choice under risk or uncertainty can be described through the use of a cardinal scale of utility, which makes sense of the comparison between any two variations of utility.[17]

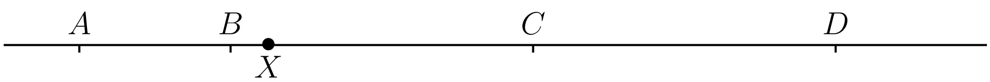
[15]Phillips and Winterfeldt (2006) recounted how Edwards progressively abandoned research on testing the descriptive validity of the subjective expected utility model, concentrating his efforts on Bayesian decision theory. Edwards was reluctant to follow the new perspectives opened by Lichtenstein and Slovic in their experiments on probability preferences, then by Tversky and Kahneman on choice by representativeness: "He was more interested in making the SEU model work in practice rather than in discovering its descriptive violations. So, he gave up on this line of research altogether." (Phillips and von Winterfeldt, 2006, 11)

[16]Coombs also contributed to other topics of the psychology of choice, notably to the literature on risk perception, always with the aim of identifying a polynomial model describing risk perception as a function of different parameters. We shall not discuss these contributions, since they are not central to our story.

[17]As is well-known, the ordinal scale implies that any monotone increasing transformation of the utility function $u(x)$ can represent equally well an individual's preferences. Therefore, the only meaningful property of numbers is that any relation $u(x) \geq u(y)$ is preserved. In the case of a cardinal scale, if $u(x) - u(y) = k \times (u(z) - u(t))$, $k \in \mathbb{R}_+^*$, then this relationship must be preserved by any acceptable utility function $v(.)$ representing the same preference ordering. The rationale for resorting to a cardinal approach to utility stems either from an assumption that statements

In “Psychological scaling without a unit of measurement”, Coombs (1950) suggested that a new scale of measurement could be conceived of, an ordered metric scale, whose property is that it makes sense of the relative distance between stimuli (whatever the choice of zero and unit values), only to the extent that distances between stimuli can be described as a partial order. In other words, the ordered metric makes sense of a statement like “the distance between stimuli A and B is greater than the distance between C and D”, but it does not claim to make sense of a statement like “the distance between A and B is twice the distance between C and D” (which is assuming cardinality). I shall present the aim of the ordered metric briefly through a hypothetical example. Then I will point out some features of the model and some of its uses and limitations. Lastly, I will focus on some applications by Coombs of the underlying notion of “psychological distance”, notably his use of criteria of transitivity as a test of the ordered metric scale. This detour will help us understand Tverky’s later views about modeling and measurement of psychological states.

The motivation for building an ordered metric scale is to know if a set of persons can share a common representation of objects (taken as stimuli) and a common view about the "similarities" and differences between these objects, from which a scale of the stimuli along a continuum could be inferred. Coombs’ overall study of this issue is called "the unfolding theory of preferential choice". Let us assume that different *Ss* can order a set of stimuli on a scale, that is, that the various stimuli contain a common latent attribute that serves to order them along a line from the one that contains the smallest quantity of the attribute on the left to the one that contains the most of it on the right. The question is whether it is possible to identify more precisely the relative distances between the various stimuli along the line and to establish this scale as a common scale for the *Ss*. Consider for instance four candidates who are seeking to be chosen to represent a certain party in the next general election. There may be a general agreement among all the members of the party that the four candidates represent various degrees of pro-market or protectionist stances (within the usual boundaries that make the identity of the party). Hence, under the assumption that all members order the candidates the same way ($ABCD$), from candidate A (representing a pro-market tendency) to candidate D (representing a protectionist tendency), the question is whether it would be possible to infer from stated preferential choices of the members an order about distances between A and B, B and C, C and D. To arrive at this supplementary "quantitative" information, Coombs assumed that each member has an ideal candidate on the scale, corresponding to the exact blend of free market and state intervention s/he would like to support. The ordered representation on a horizontal line of the set of stimuli (a candidate being $ABCD$) and of the ideal X is a J scale, a joint distribution of the stimuli and the ideal points along the line.



The unfolding theory of preferential choice consists in conducting an experiment

such as $u(x) - u(y) = k \times (u(z) - u(t))$ make sense to individual agents or, as in the case of EUT, that individuals abide by the axioms of the theory.

involving paired comparisons of stimuli. Through paired comparisons, *Ss* are asked to reveal their preferred order of candidates. It is assumed that they select the candidates according to their proximity with their own ideal on the scale. If all the *Ss* choose that way, under some conditions,[18] it is possible to analyze the order of preferences as a function of the position of the ideal on the scale. All *Ss* having an ideal near stimulus A will indicate ($ABCD$), but this order will change once the ideal is on the right of the midpoint AB (meaning the midpoint of segment [AB]). Once the ideal is on the right of AB, in between AB and AC, the preference order becomes ($BACD$). If the ideal is on the adjacent segment joining AC and BC, the preferential order becomes ($BCAD$). Hence, from a set of individual preferential orders (I scales),[19] if it is possible to identify a unique set of I scales going from ($ABCD$) to ($DCBA$), then one can recover the quantitative J scale that generated them. This J scale is quantitative, meaning that it contains additional information regarding the relative distance between stimuli.[20] If the experimenter can identify a single set of I scales generating J, he can then retrieve some additional information from the order of all the midpoints AB, AC, AD, BC, BD: "there are certain relations between the manifest data and the metric relations of the continuum" (Coombs, 1950, 150). For instance, if AD precedes BC, it means that $\overline{CD} < \overline{AB}$. This is a quantitative information on relative magnitudes on distances between stimuli, shared by all the *Ss* through their preferential choice. Hence, the total amount of information allows to build an ordered metric from the quantitative J scale, if the set of I scales that generated it is complete and unique.[21]

A condition for obtaining an ordered metric is that there is some collective representation of stimuli; hence ordered metric scales are embedded in the search for collective properties behind individual preferential orderings. They are conditional on stringent assumptions regarding the qualitative ordering of stimuli. The point, however, is that Coombs' research program is embedded in the traditional issues about measurement and the search for a common dominating latent attribute determining preferences and choices. In his contributions related to the unfolding theory, Coombs conjectured that *Ss* must be somehow able to make coherent choices based on the identification of a latent attribute common to the various alternatives and which they wish to maximize.

[18]Notably, information on relative distance depends on having a representative set of *Ss* covering a whole gamut of preferential orderings of A, B, C and D.

[19]I scales are obtained by taking all the stimuli on the left of an ideal and drawing their symmetrical value about the ideal. The half straight line originating at the ideal point is the I scale for this S.

[20]Actually, with four stimuli, there are two such sets of successive preferential orders that can generate a qualitative J scale, depending on the asymmetry of the positions of B and C on $[AD]$. Hence the quantitative content of the J scale depends critically on the possibility to identify a unique set of I scales generating J. And the possibility for alternative I sets generating a J scale are dramatically increased with five stimuli and more, all the more limiting the hopes of obtaining reliable quantitative information.

[21]To Coombs, in some situations the order of the stimuli as individuals perceive them is well known, and the psychologist devises experiments to identify quantitative J scales. In other cases, the knowledge of the latent attributes against which the various stimuli are ordered is not known and it is the aim of the experimenter to identify them or to search for the attributes that best make sense of the preferential orders. In this case, the number of alternative orders and the probability that not all *Ss* share the same order of stimuli on J leads to insuperable difficulties to analyze data.

There is one notable application of the unfolding theory for our story. In some follow-up contributions Coombs (1958; 1959; 1964) used a probabilistic interpretation of J scales. Stimuli and *Ss*' ideals are now represented as symmetrical distributions around their mean, in accordance with a probabilistic interpretation of preferences. Therefore, some overlapping between stimuli may occur, and occasions for inconsistent (nontransitive) preferential choices take place. Through a comparison of different definitions of transitivity, Coombs conceived of a test of the unfolding theory. The discussion of transitivity measures is detailed in Coombs' *magnus opus*, *A Theory of Data* (1964)

The experiment used to discuss transitivity is as follows. *Ss* are presented with series of cards with four shades of grey. They are asked to rank order the four shades of grey "from the best representative of what he meant by the label 'grey' to the poorest." (Coombs, 1964, 108). From a set of replicated choices among four shades of grey taken out of twelve different shades from the lightest to the darkest, Coombs retrieved data on pairwise probabilities $p(X, Y)$ for all pairs XY of shades of grey. In each trial,[22] since the method for obtaining pairwise comparisons based on four stimuli orderings does not impose transitivity, it is possible to use the data to check if the preferences over the range of shades of grey are transitive. To that end, Coombs provided the first comparison between various definitions of transitivity in a probabilistic framework.[23] In a probabilistic framework, A is preferred over B if more often than not A is chosen in a paired comparison of both stimuli, that is, $p(A, B) > 0.5$. Three more or less stringent definitions of transitivity can be derived from this definition of preferences:[24]

- *Strong Stochastic Transitivity* (SST)

$$p(A, B) > 0.5 \text{ and } p(B, C) > 0.5 \Rightarrow p(A, C) > max[p(A, B), p(BC)]$$

- *Moderate Stochastic Transitivity* (MST)

$$p(A, B) > 0.5 \text{ and } p(B, C) > 0.5 \Rightarrow p(A, C) > min[p(A, B), p(BC)]$$

- *Weak Stochastic Transitivity* (WST)

$$p(A, B) > 0.5 \text{ and } p(B, C) > 0.5 \Rightarrow p(A, C) > 0.5$$

From a probabilistic statement of preferences, assuming symmetric distributions of both stimuli and ideal for each subject, there is a dominant I scale for each S

[22]Actually, the data are based on a decomposition model by which the probability $p(A, B)$ is obtained from knowledge of choices among triples ABX, under the assumption that $p(A, B)$ is independent of any X "a strong form of the 'independence' from other alternatives' axiom that continually crops up in choice theory and decision making." (Coombs, 1964, 53)

[23]There are a fair number of studies aimed at testing transitivity of choices and preferences in the 1950s, both on the part of psychologists and economists. However, to our knowledge, only one such study makes use of a comparisons between more or less stringent definitions of transitivity (Chipman, 1960). On some early experiments and theories involving transitivity properties of choices or preferences, see Moscati (2007) and Lenfant (2018).

[24]A fourth definition, General Stochastic Transitivity, is provided by Chipman (1960, 77).

and there is an ordering of stimuli satisfying WST. Now, a predicted outcome of unfolding theory of preferential choice is that SST will not hold in general, while MST will hold. This is due to the fact that a shift of the ideal point (within the range of values of its distribution) will affect differently the comparison of stimuli located on the same side of the ideal (*unilateral pairs*) and the comparison of stimuli located on opposite sides (*bilateral pairs*). Hence, some laterality effect should be expressed by the data. When this idea is applied to the study of preferential choice over triples of stimuli, the theory predicts that SST should be violated more often for some kinds of triads than for others. Experimental data confirm this prediction ((Coombs, 1964, 113-114, tables 2 and 3).[25]

Also, in the case of this experiment on shades of grey, Coombs obtained that a significant proportion of triples of choice satisfy at least moderate stochastic transitivity, and the same qualitative results were obtained on a study involving gambles (Coombs and Pruitt, 1961).

Coombs' contributions are characteristics of works in mathematical psychology of choice and preferences in this period. The research agenda was centered on building sophisticated scales of measurements and devising experiments aimed at eliciting such scales and testing their robustness in a probabilistic environment. In this case, definitions of transitivity are simply operational tools for testing a measurement scale. Individuals are seen as consistent calculative machines within the boundaries of what a stochastic definition of consistency allows. This view is quite representative of the inquiries on the consistency of choices in the 1950s. It was to change by the end of the 1960s.

When Sarah Lichtenstein, Paul Slovic and Amos Tversky came to prepare their PhDs at the university of Michigan, they were involved in working along the lines of research agenda promoted by Edwards and Coombs. Their contributions between 1965 and 1970 bear the mark of a progressive emancipation from this agenda.[26] A sketch of the first contributions of Slovic, Lichtenstein and Tversky is in order at this point, to identify their first doubts regarding the approach promoted by Coombs and Edwards and the usual approach to the modeling of choice and risky decisions.

2.3 Amos Tversky: from scaling to information processing

Tversky was born in 1937 in Haifa, Israel. He obtained a Bachelor degree from the psychology department at Hebrew University, Jerusalem, then flew to the US in the fall of 1961 to graduate at the university of Michigan after reading Edwards (1954c). There, he became more attracted by Coombs' personality and research. In 1964,

[25]Triples of stimuli can be classified according to the distribution of the three stimuli around the ideal. There are three cases : 1) all stimuli are on the same side (*unilateral triples*); 2) the stimulus which is single on one side can be either the nearest or the farthest from I (*bilateral adjacent triples*), or 3) it can be in between the two opposite stimuli from I (*bilateral split triples*). SST should be violated more often than MST in the case of bilateral adjacent triples than for unilateral triples, and violations of STT for bilateral split triples "should almost never occur." (Coombs, 1964, 113-114)

[26]Slovic, Lichtenstein and Tversky co-authored a few articles with Edwards (Edwards and Slovic (1964); Tversky and Edwards (1966); Slovic et al. (1965), and Tversky co-authored a handbook on mathematical psychology with Coombs (Coombs et al., 1970) and co-edited *Decision Making*, book of selected readings (Edwards and Tversky, 1967).

he defended his PhD entitled "Additive choice structures" under his supervision (see Lewis (2016, 100 *sqq*)).

Tversky's early work was devoted to the theory of measurement and its applications to psychology. Grounding his analysis and thoughts on experimental results and on a deep understanding of measurement theory, Tversky came to the conclusion that, more often than not, choices and decisions cannot be described through the use of compensatory models and measurement scales.

In the 1950s and 1960s, there was a wide consensus in the field of mathematical psychology and psychology of decision that, when making decisions, individuals maximize some criterion of worth and that worth can be decomposed into its basic elements. The role of psychologists was to find out how these elements are compounded and can compensate for one another. Edwards and Coombs' work shared this view. From the beginning, Tversky set out to discuss the theoretical foundations and the validity of tools used to substantiate these ideas.

> "The decomposition of a complex phenomenon into a set of independent factors with a specifiable rule of combination may be regarded as one of the goals of scientific investigation. Performance, for instance, may be decomposed into learning and drive factors, or conflict may be analyzed into approach and avoidance components. The choice of addition as the basic rule of combination of the various components is basic to most models that have been constructed to analyze and explain psychological data." (Tversky, 1964, 1)

Actually, Tversky's dissertation on additive structures was a contribution to (simultaneous) conjoint measurement, a kind of measurement structure introduced in the 1960s and heavily discussed during the decade. Tversky's PhD (1964) deals with measurement issues in mathematical psychology and with applications of some measurement device to the study of decision making (riskless and risky choices).[27] After Stevens' reformulations of the categories of measurement for the social sciences (Stevens, 1946), mathematical psychologists focused on the study of axiomatic foundations for measurement systems adapted to various constraints on psychological attributes, notably the absence of any natural scale and quantitative knowledge about these attributes. Standards of scientificity were felt to depend on the possibility to establish some properties connecting various psychological states. To that end, ordinal scales are not informative enough and it was expected that psychological states and behaviors could be approached at least through interval scales. Adopting a conventional 0 and unit would then lead to ratio scales, putting psychological measurement on a par with measurement in physics. Research was thus oriented

[27] A detailed exposition of measurement theory and ideas about measurement in psychology and economics is beyond the scope of the present article. This complex subject has been discussed by Michell (1986; 1999; 2007; 2021) for the case of psychology and Moscati (2018) for utility measurement. For a broader, interdisciplinary view, see Boumans (2012, 2015). The story started in the 1920s with physicist Norman Campbell's rejection of the possibility of "fundamental measurement" in the social sciences (Michell, 1986, 2007), i.e. the possibility of a rule of concatenation of the elements to be measured. In the field of psychology, Thurstone contributed to establishing a list of conditions for useful measurements of individuals' attitudes.

toward obtaining forms of measurement for the various independent attributes "in terms of axioms that lead to scales of the highest repute: interval or ratio scales" (Luce and Tukey, 1964, 4).[28]

The axiomatic theory offering a solution to this challenge is simultaneous conjoint measurement.[29] The aim of conjoint measurement theory is to provide a set of axioms and operational recipes in order to link a set of independent psychological attribute with another, dependent, psychological attribute. For instance, the feeling of loudness could be modeled as a compound of two independent feelings: tone and intensity; or the maximum sum of money to be spent on a basket could depend on the mixture of goods it contains; or else, the feeling of physical discomfort could be the compound outcome of temperature and humidity. For this purpose, *Ss* are assumed to be able to order in a coherent way the dependent data representing the compound effect of two independent attributes (taken as stimuli). Theories of simultaneous conjoint measurement provide sets of axioms about the properties of the ordering of the compound effect of two (assumed) independent attributes that are sufficient to assess the possibility to express the joint effect of the two attributes as a function of the effect of each independent attribute. Consequently, measurement scales would be obtained simultaneously for the compound attribute and for the independent attributes.

Apart from its axiomatic structure, the whole theory relies on several assumptions regarding the existence of independent attributes and some continuum of values that each attribute can take on. It is beyond the scope of the paper to discuss this theory. In the following, our goal is to present the first axiomatic treatment of the theory of simultaneous conjoint measurement and to assess Tversky's amendments and criticisms of it.

Over the period 1960-1970, the theory of conjoint measurement has been a hot issue. It was investigated notably by Adams, Davidson, Debreu, Fagot, Fishburn, Krantz, Luce, Scott, Suppes, Tukey, and Tversky (see Krantz et al. (1971, chaps 6 and 7)). The most widely studied variety of conjoint measurement structures is additive conjoint measurement, a set of axioms about the comparability of data and properties of their orderings such that the joint effect of two independent attributes is represented as the sum of their scaled value. An additive structure offers the possibility to compare different theories of decision making with a simple two-dimensional model. The first axiomatic treatment of additive conjoint measurement is due to Luce and Tukey (1964), who give a general presentation of the additive conjoint measurement model in the case of an infinite number of stimuli. It is usually regarded as a seminal contribution (Trendler, 2019).[30]

Let $\mathcal{A}$ be a set with elements $A, B, C, \ldots, F, G, H$ and $\mathcal{P}$ a set of elements

[28]Besides, ordered metrics and other information on relative proximities of independent variables (stimuli) may be a starting point for elaborating interval scales (cardinal scales) (Abelson and Tukey, 1959).

[29]This theory has been adapted in marketing studies to elicit consumers' preferences (see Carroll and Green (1995)

[30]Debreu (1959) is usually regarded as a less general presentation than Luce and Tukey (1964) because it relies on a topological approach while Luce and Tukey's approach is algebraic. Debreu (1959) contains some references to previous works in mathematical psychology and economics dealing with similar issues.

$P, Q, R, \ldots, X, Y, Z$. The data set of typical elements $D(A, P)$ of all pairs $\mathcal{A} \times \mathcal{P}$ is the set of subjective effects associated with the couple of stimuli (A, P). Let $\geq$ be a binary relation on elements of $\mathcal{D}$. The question is: What set of axioms on the structure $\langle D, \geq \rangle$ is equivalent to representing values in $\mathcal{D}$ as the sum of real valued functions on elements of $\mathcal{A}$ and $\mathcal{B}$, $f(A)$ and $g(P)$: $\phi(A, P) = f(A) + g(P)$?

Luce and Tukey's theorem of the existence of a conjoint measurement structure (1964) relies on four axioms.

- Axiom 1 (Weak ordering): $\geq$ is a weak ordering over $\mathcal{D} \times \mathcal{D}$ (reflexive, transitive, connected)
- Axiom 2 (Solvability): $\geq$ fulfills a solvability condition. For any $A \in \mathcal{A}$ and any $(P, Q) \in \mathcal{B}^2$ there exists $F \in \mathcal{A}$ such that $(F, P) = (A, Q)$ *and* for any $(A, B) \in \mathcal{A}^2$ and any $P \in \mathcal{B}$ there exists $X \in \mathcal{B}$ such that $(A, Y) = (B, P)$
- Axiom 3 (Cancellation): For $A, B, C \in \mathcal{A}$ and $P, Q, R \in \mathcal{P}$, $(A, Q) \geq (B, R)$ and $(B, P) \geq (C, Q) \Rightarrow (A, P) \geq (C, R)$
- Axiom 4 (Archimedean): Define infinite sequences of (A, P) in $\mathcal{A} \times \mathcal{P}$ denoted by (A_i, P_i), $i = 0, \pm 1, \pm 2, \ldots$. For any (B, Q) in $\mathcal{A} \times \mathcal{P}$ there exist integers n and m such that
$$(A_n, P_n) \geq (B, Q) \geq (A_m, P_m)$$

The central axiom for conjoint measurement is the cancellation axiom. It states that "if a change from A to F overbalances a change from X to Q, and if a change from F to B overbalances one from P to X, then the over-all combined change from A to F to B overbalances the combined change from P to X to Q." (Luce and Tukey, 1964, 4)[31] The literature on conjoint measurement explores various extensions and trade-offs between empirical operationality of the theory (number of observations), realism (natural boundaries on human beings as sensorial systems), and adaptation to the standards of fundamental measurement as physicists conceive it. Tversky's PhD (Tversky, 1964) contains three important contributions to the literature on additive conjoint measurement.

First, it reformulates the theory (peculiarly the cancellation axiom) for a finite set of basic elements in order to avoid the solvability and Archimedean axioms. The solvability axiom is not acceptable in psychology. "In spite of its interesting consequences, the solvability axiom restricts the applicability of the system by requiring the factors to be both dense and unbounded." (Tversky, 1964, 29). Unboundedness ignores some limits inherent to the situation under study. As an illustration taken from utility theory, consider a gamble (Win \$100 with probability 0.9, Win nothing with probability 0.1). If utility and subjective probability are independent, the matrix of data representing the effect (\$Win, p(Win)) is additive, and the solvability axiom implies that one can associate with any value, even small, of the amount to win, say \$1, a given probability q such that (\$100, 0.9) is indifferent to (\$1, q). "It seems impossible to find such a value of q because probabilities are bounded and

[31]Thus, the cancellation revolves on the possibility to compare transitions and to order them, a property that is known to be fundamental in utility theory to obtain a cardinal mesurement (Alt, 1971).

any reasonable subject will prefer the former gamble even when q = 1."(Tversky, 1964, 30). Density (also assumed by the Archimedean axiom) asserts the possibility to generate an arbitrary number of intermediate values by a suitable choice of independent variables, contradicting the knowledge that "many psychological variables are naturally discrete or finite, and although they may be conceived of as embeddable in some continuum, the form and the nature of such an embedding is neither immediate nor clear." (Tversky, 1964, 30).

Second, Tversky (see also Tversky (1967a)) discusses conditions for obtaining (or choosing) a unique additive decomposition of the scale representing the mixed effect ϕ into the scales representing each independent stimulus, such that $\phi(A_i, B_j) = f(A_i)+g(B_j)$.[32] In general, there is no unique set of functions, and the measurement structure gives only an ordered metric *à la* Coombs. Additional axioms are necessary to obtain an interval scale (Davidson and Suppes, 1956), that is, a scale unique up to a linear transformation. The choice of additional conditions to obtain an interval scale depends on the kind of problems one wants to address. Tversky discusses different methods (algorithms) to obtain the best additive representations in various circumstances.[33]

Third, Tversky applies the additive choice structure to the study of riskless choice and risky decision. As far as riskless choice is concerned, he regards the assumption of additivity as a building block of the history of choice under certainty in economics, even though they are known examples of interdependent goods.[34] Empirical testing by psychologists of the additive model dates back to Thurstone (1931) and it met with mixed success.[35] As regards decision under uncertainty, four types of expectation model are to be compared and tested (SEU, EU, EV and SEV) (models differ in whether probability or value are regarded as objective or subjective), and they all assume that both the amount won and the probability contribute independently to the worth of a gamble (be it through subjective transformation). "More specifically,

[32]Note that this representation is adapted to the representation of simple multiplicative relationships between the independent variables, since it can be transformed into an additive relationship through rescaling with a logarithmic transformation.

[33]For instance, when a set of numerical values is given from the start or attributed by the theoretician, a unique solution can be the one that minimizes the distance from the data. When some information about difference between numerical values is regarded as relevant, one may choose functions f and g so as to maximize the number of pairwise differences orders that are preserved (a *faithful* solution). Another possibility is to reduce the interaction term of an analysis of variance, hence identifying the unique mean additive solution.

[34]Actually, economic theory in the wake of Pareto and the ordinal evolution has distanced itself from the additive model, even though the basic two-good model and the utility representation by a Cobb-Douglas utility function still assumes additivity. The fact is that since Slutsky (1915) and Hicks and Allen (1934a; 1934b), we know that the only way to do justice to subjective interdependence between goods entails the consideration of at least three goods in the model. Consequently, Tversky's simple example of right-shoe and left-shoe as interdependent goods is irrelevant (Lenfant, 2006).

[35]The main attempts to deal with the measurement of marginal utility or utility of money are Gulliksen (1956); Thurstone and Jones (1957); Adams and Fagot (1959). Notably, in Adams and Fagot (1959), the alternatives of choice were job applicants with two attributes, intelligence and ability to handle people, and this led 18 *Ss* out of 24 to reject the additive model, exhibiting instead intransitive behavior. Tversky lucidly concludes: "No attempt has been made to specify the psychological conditions under which additivity holds or fails to hold." (Tversky, 1964, 34)

the models claim that utility and subjective probability are compensatory but not interacting" (Tversky, 1964, 38). Tversky improves on the experimental testing of the SEU model by providing a test of a necessary and sufficient condition of SEU for a specific class of gambles (A, p) (and probability $1 - p$ to win 0). Based on experiments, Tversky concludes that the additivity assumption is well supported for the SEU model. However, he identifies that either the assumption of independence of utility from risk or the assumption that subjective probabilities sum to 1 is not supported.[36]

As I have already pointed out, it is beyond the scope of the present study to enter into the history of conjoint measurement and its connections with the history of utility measurement in economics. To get an intuitive account of the stakes of the debates between the researchers involved in this history, one can read it as the search for a balance between an instrumental-descriptive view of measurement and a cognitively meaningful modeling of valuation processes. According to the latter view, entities involved in the model should play an independent role in the formation of an overall value, and this in turns entails discussion of *Ss*' abilities to identify various stimuli as similar and to treat them unambiguously as items of the same independent variable. On this account, Tversky, from his first works, was concerned about the possibility to identify various items as similar and to make categories of them.

Tversky (1967b) develops a more complex model to deal with the possibility that some factors are additive and others multiplicative (see also Krantz and Tversky (1971)). In the wake of Coombs, he reflects on the notion of "psychological distance" or, equivalently, on the feeling of "dissimilarity" between multidimensional stimuli. Fundamentally, the additivity assumption expresses that an ordering of dissimilarity varies monotonically with the Euclidean distance. If any two stimuli are represented as vectors whose components are the values on each dimension $x = (x_1, \cdots x_n)$ and $y = (y_1, \cdots y_n)$, then the Euclidean distance is $d(x, y) = [\sum_{i=1}^{n}(x_i - y_i)^2]^{1/2}$ and the model assumes that an impression of dissimilarity varies monotonically with an additive combination of contributions from the differences along each dimension.

An article co-authored with Krantz (Tversky and Krantz, 1969) provides a test for an additive representation of the impression of dissimilarity among simple three-dimensional structures of schematic faces. Faces are represented through three attributes (face contour / eyes / mouth), each of which can take on two values (rounded vs flat / white or black / smiling or neutral). Paired comparisons allow to obtain an ordinal scale of dissimilarity that satisfies the additivity assumption. However, the validity of this result is limited to a simple three-dimensional case and a sample of two values for each continuous attribute. "Hence, the conclusions are subject to the reservation that the fuller space has not been thoroughly explored, and the introduction of additional levels and/or dimensions may reveal some interactive patterns" (Tversky and Krantz, 1969, 127) Notably, Tversky and Krantz recognize that judgments of similarity usually involve some configural or "interactive" treatment of stimuli, that is, some Gestalt effect. However, they believed that some specific iden-

[36]Note that the model assumes u(0)=0, hence the interpretation is that measures of the variation of utility are derived from the amount won, as compared with a *status quo*. Hence, the total wealth is not included in the valuation.

tification of non-interacting features would allow recourse to an additive-difference model (a generalization of the Euclidean distance model) to represent the ordering of stimuli. A major characteristic of these models is a joint-factor independence law: If some combination of differences on $n-1$ dimensions exceeds that produced by some other combination on the same dimensions, the same ordering will prevail even if the value of the difference on the nth dimension is varied. The article thus conveys the view that individuals' choices may be analyzed as the result of a mental combination of "independent subimpressions" (ibid, 127) (see also Krantz (1972)).

Tversky was involved in this literature on conjoint measurement up to the beginning of the 1970s, with the publishing of the first volume of the landmark *Foundations of measurement* (1971)(written with Krantz, Luce and Suppes). In these contributions, Tversky sometimes challenged the usual assumptions about measurability and the mere possibility of analyzing judgments about the similarity between various stimuli (Beals et al. (1968), Krantz and Tversky (1971)). Within the community of mathematical psychologists, Tversky seemed particularly mindful of the need to develop sophisticated models of measurement that would be cognitively meaningful, that is, models accounting for all the relevant elements involved in the evaluation of the alternatives and of their combination. "The usefulness of utility theory for the psychology of choice ... depends not only on the accuracy of its predictions but also on its potential value as a general framework for the study of individual choice behavior." (Tversky, 1967a, 199)

This called into question the legitimacy of the use of compensatory models. In an article co-authored with Russo (Tversky and Russo, 1969), the authors wondered about the cognitive assumptions associated with the use of compensatory models. Choices are usually interpreted as the result of a confrontation between the values of two alternatives. Endowing *Ss* with abilities to attribute values and to choose according to values is a scalability assumption. If a set of binary choices (choice of x over y) are represented through a probability function, Tversky and Russo established the equivalence of the following properties:

- *Simple scalability.* Over the choice set, the probability function can be represented as a function of the difference between the values of x and y :

  $$p(x,y) = F[u(x), u(y)],$$

  u increasing in x and y, F increasing in $u(x)$ and decreasing in $u(y)$[37]

- *Strong Stochastic Transitivity*

  $$p(x,y) > \frac{1}{2} \text{ and } p(y,z) > \frac{1}{2} \Rightarrow p(x,z) > max[p(x,y), p(y,z)]$$

- *Substitutability* For all x, y and z in the choice set, $p(x,z) \geq p(y,z) \Rightarrow p(x,y) \geq \frac{1}{2}$[38]

[37] $p(.)$ has the property that $p(x,x) = \frac{1}{2}$. Thurstone (1927) and Luce (1959) choice model are a special case, with $p(x,y) = F[u(x) - u(y)]$.

[38] A result obtained by Block and Marschak (1960), theorem 4.1

- *Independence*: $p(x,z) \geq p(y,z) \Rightarrow p(x,w) \geq p(y,w)$

All four properties, therefore, capture the same principle that pairwise choice probabilities can be expressed as a monotone function of some underlying scale values. These properties were challenged on many occasions (Krantz, 1964; Coombs, 1958), thus weakening the general assumption of simple scalability:

> "Although this principle has dominated much of theoretical work in the field, research exists indicating that choice probabilities are affected by comparability factors which cannot be accounted for by any model based on simple scalability. ... These results indicate that choice probabilities reflect not only the scale values of the alternatives, but also the degree of difficulty of the comparison. Consequently, substitutability is violated as alternatives may be substitutable in some contexts but not in others." (Tversky and Russo, 1969, 4).

Thus, failures to satisfy simple scalability prove that subject sometimes face real difficulties in comparing alternatives in paired-choice comparisons, in turn jeopardizing the current representations of *Ss* has perceiving complex stimuli within a psychological continuum. To test if these axioms are satisfied equally well, Tversky and Russo carried out an experiment. *Ss* were asked to compare different shapes of rectangles and lenses with two standard shapes of the same size (two extreme standard lenses and two extreme standard rectangles). The independence principle turned out to be violated, leading to a final conclusion:

> "If, as available data indicate, simple scalability is violated in many contexts, then both theoretical and applied research on choice behavior should be fundamentally reevaluated." (Tversky and Russo, 1969, 11)

Moreover, the experimental result opens to the possibility to influence choices:

> "Finally, to the extent that the similarity hypothesis is applicable to the decisions of consumers or voters, it suggests the intriguing possibility of influencing choice probabilities between products or candidates by manipulating the similarity between them." (Tversky and Russo, 1969, 12)

By 1969, Tversky had set up an alternative research agenda. Over the same period, Paul Slovic and Sarah Lichtenstein followed a similar evolution, questioning the dominant modeling of information processing and decision making.

2.4 Slovic and Lichtenstein on probability judgments

Around 1962-1964, Sarah Lichtenstein and Paul Slovic were preparing their PhDs at the University of Michigan. Both were involved in the study of risk-taking behavior.[39]

[39] Sarah Lichtenstein (1933-2017) was born in 1933 in Seattle (Washington). She graduated from Swarthmore college (Pennsylvania) in 1956, then went to Michigan University, Ann Arbor, where

In her PhD dissertation, Lichtenstein pioneered the study of three-outcome bets (Lichtenstein (1965) is based on it). Her contributions during the Michigan years are related to decision under risk. Several of her studies deal with understanding the effect of contextual elements of an experiment on the *Ss*' responses (effects of verbal phrasing, quantitative perceptions of loose verbal phrases like "very likely", "probable", "better than even") in order to devise well-controlled experiments on behavior under risk. Notably, she identified a systematic asymmetry in quantitative responses to mirror-image phrases like "quite likely" vs "quite unlikely" (Lichtenstein and Newman, 1967).

From his first work, Slovic (1962) studied the perception of risk and its modeling. In the same period, Slovic's research agenda was focused on the ability to provide descriptive models for various psychological continua associated with risky decisions. Already in his first works, he expressed skepticism about the possibility to provide any measurement of an attitude toward risk: "willingness to take risks may not be a general trait at all but rather one which varies from situation to situation within the same individual" (Slovic, 1962, 68).[40]

Questioning the use of information in decision making, Slovic comes to the problem of multiple "cues" and their combination. The structure of the research involved issues about similarity, coherence and information processing. Facing a problem of valuation of a situation, cues convey information about the properties of a system (or of a situation): How do *Ss* "combine information from multiple cues, each with a probabilistic relationship to a criterion, into a unitary judgment about that criterion?" (Slovic, 1966a, 427) Slovic went beyond a simple search for a descriptive additive model. The heart of the issue is how judges confront the various cues at their disposal, appraise their coherence and draw conclusions about the properties of the system. If cues are judged to be consistent, they can be used additively to make a judgment. Instead, if cues are judged inconsistent, the judge can resolve this inconsistency by focusing on a consistent subset of cues or by looking to previously ignored cues.

To establish how *Ss* proceed in the case of conflicting cues, Slovic proposed to test a linear model to describe judgments of intelligence made by *Ss* based on cards describing people through nine attributes with varying values of the attributes. The conclusion was that when two dominant cues are coherent (high grades in school and high values for English Effectiveness) then *Ss*' judgments can be described through

---

she earned a Master's degree and a PhD in Psychology under the supervision of Ward Edwards in 1962. Paul Slovic was born in 1938 in Chicago. He earned an undergraduate degree at Stanford in 1959, then went to University of Michigan, Ann Arbor where he earned a MA in psychology in 1962. As a graduate student at Michigan, Slovic became a research assistant in Clyde Coombs' laboratory in 1959, then worked in Edwards' laboratory in 1960 while Coombs was on sabbatical. There, he met Sarah Lichtenstein and Amos Tversky. He started collaborating with Lichtenstein and Edwards, and he earned a PhD in psychology in 1964 under Edwards' supervision. He then joined the Oregon Research Institute (Eugene, Oregon) as a researcher, after Lewis Goldberg's invitation (his former mentor at Stanford). He was joined by Lichtenstein in 1966. From this time on, Slovic and Lichtenstein started working on their own topics. In 1976, together with Baruch Fischoff, they founded Decision Research, a research institute dedicated to research on risk and decision making.

[40]Through an experiment on risk-taking in children, he found that risk-taking as a differentiated boy's trait appears around the age of 10 years and is thus a result of social pressure (Slovic, 1966b).

the additive model. Instead, when the values of cues have contradictory implications, one of them is excluded from consideration and other cues are then taken into account. Thus, the selection of information to be used to make a judgment is part and parcel of the judgment process. In short "These data call for a modification of the commonly accepted 'linear model' of information combination." (Slovic, 1966a, 434)

Lichtenstein's PhD dissertation led to the publication of a seminal article (Lichtenstein 1965). In this experimental study of risk behavior, she undertook to elicit the relative importance of various indirect variables in gambles (variance (V), expected value (EV), skewness and kurtosis), possibly obtained from the knowledge of simple information: probability to lose $p(L)$ and to win $p(W)$, amount to lose ($\$L$) and to win ($\$W$). The aim was to compare various normative theories of risky choices: Expected value (EV), Expected Utility Theory (EUT) and Subjectively Expected Utility (SEU). Lichtenstein's innovation was a system of three-outcome bets that makes it possible to monitor the value of skewness and kurtosis. Assuming that one model fares better than another entails taking into account different layers of cognitive assumptions. Notably, the SEU model *per se* is too tolerant given that it does not put enough constraints on the subjective probability function. The experimental design consists in asking *Ss* how much they would bid for various bets. Bets are built in such a way that some bets have the same EV, V, and skewness. Bets that vary only along one variable are compared and statistical tests based on *Ss* bids along different objective data are carried out. Lichtenstein's results are that both EV and V play a significant role in influencing choice, other things equal. However, when EV or V are equal, other items of information become relevant. However, Lichtenstein points out that models based on structural indicators (EV or V) are not satisfactory, because they cannot account for the specific importance of other combinations of single variables, such as the value of the least likely outcome.[41] Indeed, experiments show that the largest least likely outcome explains the choice when two bets have the same variance but not when variance differs between two bets. In this case, *Ss* choose the bet with the least variance. This suggests some discontinuity in information processing:

> "These findings suggest that *Ss* approach bets with a lexicographic ordering of relevant variables: EV is of primary importance. When deciding between bets of equal EV, *Ss* prefer low V. When both EV and V are constant, *Ss* attend to the least likely amount, avoiding a small chance of large loss." (Lichtenstein, 1965, 168)

In one of their first co-authored papers at the Oregon Research Institute, "Relative importance of probabilities and payoffs in risk taking", Slovic and Lichtenstein

[41]She rejected the probability-preference identified by Edwards (1953; 1954b; 1954a) or Coombs and Pruitt (1960) as an artifact due to the structure of two-outcome bets (knowing one probability implies knowing the probability distribution). Also, Lichtenstein et al. (1969) show that explicit information about EV does not influence choice. A bias in favor of the least likely outcome "indicates that *Ss* overestimate small probabilities, so that the subjective probability curve is not linear with objective probability" (Lichtenstein, 1965, 168), anticipating Kahneman and Tversky's twisted function for probabilities (Kahneman and Tversky, 1979).

(Slovic and Lichtenstein, 1968) reported an experiment aimed at identifying the independent effect of $p(W)$, $p(L)$, \$W and \$L on the perception of risk. To this purpose, they used a new type of gamble, the "duplex gamble", an innovation that enables the probability of winning to be varied independently from the probability of losing and the amounts to lose and gain independently from each other. For example, the probability that "perceived risk" is correlated with the probability of losing, while "attractiveness" is highly correlated with the probability of winning. Moreover, the results were consistent with previous studies showing limitations in information processing and its influences on risk-taking. This led to a first reformulation of SEU and a criticism of various attempts at interpreting risk behavior in the literature, highlighting the interpretative bias due to the impossibility, in experiments, to discriminate between effects of varying probabilities and effects of varying amounts to be gained or lost. In the end, no simple weighted linear model can account for observed behavior and more complex models have to be designed, in which decisions are the result of several steps of information processing.

In the same paper, Slovic and Lichtenstein also innovated by allowing differentiating *Ss*' answers obtained through different response modes. A response mode describes the kind of task performed by *Ss*. It aims to verify whether the kind of response expected from *Ss* when performing a given task influences the way they use the information at their disposal. In their study, the important discriminatory variable is whether the task of valuing gambles entails making a monetary valuation of the gambles or not. One group of *Ss* is asked to rate different duplex gambles, hence to indicate on a scale how attractive they are. Another group is asked to bid on these gambles, i.e. to announce the lowest price they would accept to be paid for selling the gamble.[42] In the end, to account for some unexplained puzzles, having eliminated variance-preference and probability-preference explanations and various stretching of subjective probability, a satisfactory model is one that allows a step-by-step process of decision. *Ss* first evaluate the attractiveness of a gamble, then they accomplish a certain task. During the first stage of judging whether a bet is attractive (1) or not (0), $p(W)$ is the most important determiner. During the second stage, that of bidding or rating, *Ss* focus on the same risk dimensions: for about 90% of the *Ss* in the bidding group, "when [he] found a bet attractive, his judgment of the relative attractiveness was principally determined by the size of \$W; and when he disliked a bet, \$L was the primary determinant of the degree of dislike." (Slovic and Lichtenstein, 1968, 12). In sharp contrast, for *Ss* in the rating group, "the difference between the method used to estimate degree of attractiveness and that used to determine the degree of unattractiveness was considerably less marked." (ibid., 12). Highlighting the influence of simple information in a gamble and the step-by-step information processing showed that a new path of studies was opening, deprived of some normative principles. As Slovic and Lichtenstein saw it, theories of risk taking had been, explicitly or not, influenced by "normative models" and "prescribed principles" that "were incorporated into descriptive models against which performance could be compared." (Slovic and Lichtenstein, 1968, 14). The normative aspects have drawbacks: "preoccupation with normative aspects seems,

[42]If the gamble is unattractive (repulsive), the maximum price they would be ready to pay the experimenter to be relieved from the gamble.

in this case, to have led to neglect of certain psychological aspects of the risk-taking process." (ibid., 14)[43]

The psychological description provided by Slovic and Lichtenstein was provisional, but it already conjectured a two-stage process (encoding and valuing) and more, it suggested for the first time an anchoring and adjustment phenomenon. *Ss* would start with prior beliefs about the relative importance of probabilities and payoffs to evaluate a gamble, then they would try to make evaluations based on these beliefs. This effort to make a satisfactory evaluation would turn out to be difficult, because *Ss* struggle "to integrate several sources of information into a single choice or judgment" (ibid., 15) Clearly, the overall representation is one of mixing beliefs and strategies in a two-stage reasoning.

> "The decision maker is guided by certain beliefs (e.g., probabilities are more important than payoffs), which he combines with strategies designed to make his task less complex. One such strategy is to subdivide the evaluation of a gamble into two stages. In the first stage, gambles are classified as either attractive or unattractive. In stage 2, the degree of attractiveness is quantified. When bidding, this quantification seems to proceed as a crude adjustment of $W if the bet is attractive or $L if the bet is unattractive" (ibid, 15)[44]

To sum up, in just a couple of years or so, Lichtenstein and Slovic discarded the research agenda built by Edwards and Coombs which was structured in accordance with normative economic models of choice and decision, adopting ideas of maximization of a psychological continuum based on some predefined treatment of information. Slovic (1969a,b) synthesized the results obtained and acknowledged the birth of a new research agenda. The structural characteristics of a problem-solving situation (patterns of importance among cues, information display and response mode) bear on the complexity of the information-processing operations, causing systematic changes in cue utilization to reduce cognitive strain. Also, an item of information (a cue) makes sense according to whether it is associated with some other items or not: information is used configurally (using patterns mixing several dimensions) and not always incrementally and independently. To Slovic, the consequences of these cognitive features had to be acknowledged. Expert judgment is crucial in all domains of activity, and research must find out the "hidden cognitive processes of the judge. Hopefully, by understanding these processes it will be learned why some

[43]The two-stage modeling proposed here also accommodates some puzzling findings, notably regarding the ability of *Ss* to perform coherent choices and to exhibit transitive preferences or choices. When the experimental design controls for motivation, it is shown that choices performed with a high degree of motivation are more often intransitive or inconsistent than when motivation is low (Slovic et al., 1965; Pollatsek, 1966). The rationale for this is that unmotivated *Ss* tend to "reduce the cognitive strain" by adopting a simple rule to make choices (e.g. always select the highest probability of winning or always choose the lowest monetary loss). The more risk dimensions (elementary or compounded) are conflicting the more experimental results show intransitiveness.

[44]Also, choices performed with a high degree of motivation are more often intransitive or inconsistent than when motivation is low (Slovic et al., 1965). Unmotivated *Ss* may tend to "reduce the cognitive strain" by adopting a simple rule to make choices.

judges are more accurate than others, and this knowledge will, in turn, help us to train persons to make better judgements." (Slovic, 1969a, 256)[45]

The next sections set out three directions of research offered by Lichtenstein, Slovic, Tversky, and Kahneman as a newcomer in the team. The first set of contributions came from Tversky alone, the second set was due to Lichtenstein and Slovic, and the last set of contributions, that was to be by far the most influential on the building of behavioral economics, was produced by Kahneman, Tversky, and Slovic.

3 Intransitivity of Preferences

Whether the assumption of transitivity is a red line that no economist should ever cross has been a running issue at the margins of the most sophisticated treatments of utility and choice since Georgescu-Roegen (1936).[46] The issue of transitivity as a tenet of rational choice theories became more urgent in the 1950s in the hands of psychologists and some economists (Lenfant, 2018). It is beyond the scope of the present study to summarize all the complex stakes that lie behind the issue. Be it sufficient to mention the most evident one: an intransitive agent cannot be endowed with a utility function representing her preferences, hence both issues of predicting observable choices or elaborating welfare statements become meaningless.

The main methodological shift in studies on intransitive behavior in the 1950s concerned the probabilistic restatement of preferences and the transitivity assumption. Given that *Ss* seldom replicate exactly the same choices when compared with the same decision context and the same set of alternatives, the notion of preferences was now expressed as a probabilistic notion, and transitivity could be expressed as a stochastic notion too. As we have seen already with Coombs' works, the most commonly used definition of transitivity was——and still is—Weak Stochastic Transitivity, which relies on a probabilistic definition of preferences. This probabilistic definition was commonly accepted in various studies in the 1950s (Mosteller and Nogee (1951); Papandreou (1953); Papandreou et al. (1957); Davidson et al. (1957); Marschak (1955); Simon (1959); see also Lenfant (2018) and next subsection). To account for fluctuations in the evaluative process, a probabilistic version of preferences is as follows. Given two alternatives x and x', let $p(x, x')$ be the probability that x is chosen over *x'* in a repeated comparison. The preference relation $x \succsim x'$ is now expressed by $p(x, x') \geq 1/2$.[47] Consider three alternatives x_1, x_2 and x_3, the

[45]Slovic carried out an experiment with two brokers having to evaluate 128 hypothetical companies through qualitative information on 11 different cues (yield, price earning ratio, sales volume trend, etc.). Based on former studies, he concluded that a significant part of the overall judgment relies on configural judgment, that is, on a combined analysis of the values of two or three different cues, for instance resistance trend and sales volume trend.

[46]Georgescu-Roegen questioned the meaning of transitivity of preferences in terms of agents' abilities to perform the task of comparing the alternatives between which they have to choose. He would return to the subject in the 1950s, proposing a new axiomatics of preferences allowing for some intransitivity (Georgescu-Roegen, 1958)

[47]Georgescu-Roegen had already proposed a probabilistic definition in 1936, but it was not adopted at the time.

Weak Stochastic Transitivity property is defined as:

$$p(x_1, x_2) \geq \frac{1}{2} \,\text{and}\, p(x_2, x_3) \geq \frac{1}{2} \Rightarrow p(x_1, x_3) \geq \frac{1}{2}$$

In an update of his 1954 review article, establishing in passing the label "Behavioural decision theory" (Edwards, 1961), Edwards calls for a comprehensive experimental setting that would establish the conditions under which intransitive behavior is likely to happen and to be monitored. [48]

> "No experiment yet reported has created conditions deliberately designed to be unfavorable to transitivity, strong or weak, and ended up accepting even weak stochastic transitivity. In short, as a basis for psychological theorizing, algebraic transitivity is dead, and stochastic transitivity, strong or weak, has yet to be exposed to the adverse climate of hostile experiments. It seems likely that conditions can be designed in which subjects choose intransitively most of the time (unpublished research so indicates), it is even possible that the direction of intransitive cycles can be controlled by experimental manipulation. If so, the question for experimenters to answer is not whether any form of transitivity holds, but rather under what circumstances do various assumptions about transitivity hold, and under what circumstances do they not." (Edwards, 1961, 483)

Setting aside the mysterious "unpublished research" mentioned by Edwards, the most sophisticated study fulfilling his wish is Amos Tversky's "Intransitivity of preferences" (1969) (hereafter "Intransitivity"). As we have seen in the previous section, at the time of elaborating the experiments reported in "Intransitivity", Tversky had already accumulated several clues regarding the connections between transitivity, measurement issues and the way human beings process information and value outcomes. The study is a landmark in that it combines four aspects relative to intransitivity of preferences or choices: *i)* a probabilistic definition of preference, *ii)* a modeling of actual decision processes and a discussion of alternative models, *iii)* a use of subjective utilities to compare alternatives, *iv)* general comments on the meaning and scope of intransitivity in relation to rationality. Moreover, though "Intransitivity" was published in a psychology journal, Tversky established several links with economics and mentioned the main experiments carried out by economists on transitivity so far.[49]

[48] Among other things, the issue of intransitive behavior and the way of identifying genuine intransitivity is also plagued with methodological difficulties, to the extent that *Ss* tend to exhibit more consistency in the case of imaginary experiments while they tend to behave in a more complex way (multi-peaked behavior) in the case of real gambles (Slovic et al., 1965)

[49] All the examples in the article were taken from economic decisions such as buying a car, choosing a job, playing a bet, selecting a job applicant or a candidate. Also, Tversky did not hesitate to use the language of economists, writing about "preferences" as the underlying pattern of ordering of alternatives by *Ss*.

3.1 Provoking Intransitive Behavior

The starting point for "Intransitivity" is the usual statement that the assumption of transitivity is "of central importance to both psychology and economics" (Tversky, 1969, 31). It is central for economists as a necessary condition for building a (ordinal) utility function. It is also instrumental for measurement models of "sensation and value" in psychology. That said, Tversky asserted that psychologists need to devise models of choice connected with *actual* decision processes. In many situations, decision processes might not follow compensatory computations that are commonly used by economists and psychologists. This situation might generate intransitive cycles of binary choices incompatible with well-behaved preferences. Based on a stochastic definition of preferences, WST is "the most general probabilistic version of transitivity" (Tversky, 1969, 31). The advantage of a probabilistic definition of preferences and transitivity is that if there are genuine intransitivities in *Ss*' choices, they can be differentiated from other false intransitivities that may appear in the case of a deterministic definition of preferences. Tversky also insisted that probabilistic versions of transitivity had already been studied experimentally by psychologists and economists (Savage, 1951; May, 1954; Quandt, 1956; Morrison Jr, 1962; Edwards, 1953; Papandreou et al., 1957; Davis, 1958; Davidson and Marschak, 1959; Chipman, 1958; Griswold and Luce, 1962). None of them found a violation of WST, except Shepard (1964) in a study on acoustic pitch. Given this stock of past studies, Tversky's goal was to take a step forward in elaborating choice situations such that WST is violated and then to identify the elements in the choice situations that lead to this violation and to provide a model adapted to it. Building a model and building an experiment are indeed two sides of the same coin. As he puts it, previous studies were unable to detect violations of WST because they did not rely on a model that guides the construction of the alternatives and, consequently, offer at best fragile statistical evidence to accept WST (Tversky, 1969, 44).

Tversky then introduced a simple example of a multidimensional choice situation and a decision process that lead to a violation of WST. The decision rule and choice situation have the following four properties :

1. Alternatives have at least two dimensions;

2. *Ss* have to choose between two alternatives (binary choices);

3. When choosing, *Ss* give priority to the first dimension over the second dimension but this priority is subjected to a judgmental threshold. If the difference between two magnitudes along the first dimension is big enough, the alternative with the highest value is chosen and the second dimension is ignored. If the difference between two magnitudes along the first dimension is too weak (below the threshold), then *Ss* will make a choice by comparing magnitudes along the second dimension;

4. There is some objective or subjective uncertainty regarding the true value along each dimension.

These four properties being assumed, an experimental device involving series of judgments on pairs of alternatives is easily set out so that *Ss*' responses exhibit some cyclicity in terms of choices, that is an "intransitive chain of preferences" (ibid., 32).

The model that is likely to integrate this set of features is the lexicographic semiorder model (LS model), borrowed from Luce (1956 and 1955). The lexicographic property of this model implies that the process of choice does not necessarily rely on the whole set of information: it is a *non-compensatory* model of choice. If the differential value between two alternatives on the most important dimension is above his subjective threshold, an S will make his decision based on this dimension only, selecting the alternative with the highest value on this dimension. If the differential value is below or equal to the subjective threshold, the S will regard the two alternatives as equivalent on this dimension and switch to the second most important dimension, and so on.

The gist of Tversky's experimental design is to compare the principle of transitivity of a preference-or-indifference relation when alternatives are multidimensional, i.e. where alternatives vary along different dimensions all relevant to choice. The condition for applying an LS model on objects of choice is that they contain at least two dimensions. Tversky presented and discussed two experiments in which *Ss* face a choice context that is likely to induce them to follow a sequential decision process similar to the lexicographic semiorder, and consequently to exhibit cycles of intransitive choices.

3.2 Experimental design and results

The second part of the article deals with the presentation of two experiments. Experiment I is a study of preferences among gambles, experiment II deals with the ranking of applicants to a college based on information along three relevant dimensions (intellectual ability, emotional stability, social facility). Each experiment is designed and carefully prepared so that *Ss* will be induced i) to adopt a rule of decision conforming to the lexicographic semiorder model and ii) to make use of of this decision rule in the experiment.[50]

Experiment I consists in binary choices among gambles of the form $(x, p, 0)$, that is, gambles with probability p to win x and (1-p) to win nothing. The chance events are determined by the spinning of a spinner on a disc divided into a black and white area. The *Ss* have a precise knowledge of payoffs x but can only approximate the probabilities of winning p by the relative size of the black area on the disc. The series of gambles to be paired satisfy two properties: p decreases when x increases, and the expected value of the gambles increases with p (and decreases with payoffs). The gambles are also designed to form a chain of "adjacent" gambles differing from one another only by a small variation in payoff x and a small change in p. Thus, starting from a gamble with the lowest p two adjacent gambles (p, x) and (p', x') are such that $p' > p$, $x' < x$ and $p'x' > px$. The main idea behind this experimental design is that when presented with adjacent gambles, *Ss* would tend to choose based

[50] Preparatory tests are devised to filter *Ss* that would not abide by the LS model and to account for the fact that different *Ss* will have potentially different preference thresholds regarding the various attributes.

on payoffs whereas, when presented with distant gambles for which probabilities are sensibly different, *Ss* would tend to choose based on the expected value. Since payoffs and expected value are negatively correlated, *Ss* would violate transitivity if they were presented with conveniently chosen series of paired comparisons. As an application of the scheme of the LS, it is thought that *Ss* will consider each gamble as a compound of two attributes, the first being an *approximation* of the expected value, the second being the payoff. They choose according to the payoff as long as some threshold of probability difference has not been reached. Otherwise, they choose according to expected value.

The running of the experiment consists in presenting the *Ss* with adjacent pairs and with a single pair of extreme gambles, with all due precautions in selecting *Ss* and avoiding some undesirable effects (randomization, control for coherence and pattern of preference). Each *S* went through five test sessions. In each session, he was presented with the same pairs of gambles four times. For each subject, a statistical treatment over the twenty choices for each paired comparison gives a probability value of choice. The number of violations of WST and of LS is then computed. Statistical treatment of the results leads to the conclusion that five *Ss* out of eight reject WST and that one rejects the LS (Tversky, 1969, 35-36, tables 2 and 3). Postexperimental interviews show that *Ss* who consciously apply a LS criterion do not think that it can lead to intransitivities.

In Experiment II, *Ss* have to select applicants to a college. They are presented with pairs of candidates and have to select the one they would preferably accept based on cards presenting the profile of each applicant as a set of scores on different dimensions, namely intellectual ability (I), emotional stability (E) and social facility (S). Before choosing, they are told that the values have been obtained by different techniques and that the scores may not be perfectly reliable. They are also told that "naturally, intellectual ability would be the most important factor in your decision, but other factors are of some value too." (ibid., 37) The profiles are such that there is a negative correlation between the scores on the most important dimension I and the scores on dimensions E and S. A choice which is made according to the higher score on dimension I is said to be "compatible with Dimension I". Since small differences along this dimension are not easy to figure out from the cards, it was supposed that some *Ss* would adopt an LS decision rule, switching to dimension E and S when scores along dimension I are almost equal. A preliminary round of the experiment was run to identify *Ss* adopting this kind of rule and to gather some information about each subject's preference threshold e, that would be used in the design of new profiles for the test session. *Ss* were then grouped according to their preference threshold and new sets of only five profiles were prepared for each group such that the score difference on dimension I between adjacent profiles was smaller than e and such that the score difference between extreme profiles was above e. Statistical treatment of the experimental results, using a maximum likelihood ratio test, leads to rejection of WST.[51]

[51] In both experiments, several sessions of paired comparisons took place, with usual precautions to avoid memory effects (presenting *Ss* with irrelevant pairs along with relevant pairs, randomizing the order of presentation and the display of pairs). The test of the use of the LS model is important, since it is used to assess that observed intransitivities are not the result of random decisions. All

3.3 Theoretical discussion

After exposing the experimental design and the results, Tversky engages in a theoretical discussion in two steps. The overall motivation for this is to provide arguments in favor of the LS model as a good candidate for *explaining Ss*' decision processes and the resulting intransitivities. Tversky's discussion aims at establishing that intransitivity, in the sense of rejection of WST, is a phenomenon that can be shown to exist, that can be stimulated by adequate experimental devices and that can be predicted. In "Intransitivity", this phenomenon concerns binary choices among multi-dimensional alternatives. On the basis of this summary, a few themes are discussed.

The first theme revolves around the choice of a model to represent choices. Tversky is clear that the LS model is not the sole candidate. He also stresses that "[LS] is based on a noncompensatory principle that is likely to be too restrictive in many contexts" (Tversky, 1969, 40). Two formal models of choice are discussed. Tversky ponders the benefits and drawbacks of models regarding both descriptive accuracy and explanatory power. The first, the "simple additive model"—the one widely used in economics—supposes that each alternative in a choice situation can be evaluated independently from each other and given a utility or scale value. If each dimension in an alternative is valued, then the value of the alternative is simply the addition of dimensional values. The second model, the "additive difference" model, assumes that choice between two alternatives results from a comparison between the values on each dimension ("attribute" or "component") (Tversky, 1969, 41). Considering x_i^1 the i-th dimension (or component) of the alternative x^1 and $u_i(x_i^1)$ the separate valuation of this dimension, each S is assumed to calculate the difference value $\delta_i = u_i(x_i^1) - u_i(x_i^2)$ for each dimension and to apply to it a weight function ϕ_i expressing the contribution of dimension i to the overall evaluation of the alternatives ($\sum_i \phi_i(\delta_i)$). The alternative that outweighs the other is chosen. Both models suggest "different processing strategies" (Tversky, 1969, 42) though mathematically, the simple additive model is just a special case of the additive difference model when the weight function is linear (see below). Here is a first theme of enquiry that Tversky posited in the debate on decision making and its modeling, that of the relationship between a model of choice and an actual decision process.

"More specifically, the simple additive model is more likely to be used when the alternatives are displayed sequentially (i.e., one at a time), while the additive difference model is more likely to be used when the dimensions are displayed sequentially." (Tversky, 1969, 42) Tversky then argued in favor of the additive difference model over the simple additive model, putting to the fore "several general considerations" (ibid.).

test sessions were run individually. In the case of experiment I, intransitivities, "all violations were in the expected direction, and almost all of them were in the predicted locations... This result is extremely unlikely under the hypothesis that the intransitivities are due to random choices." (Tversky, 1969, 34). In experiment II, the observed proportion π of triples of pairs violating WST is compared with the expected proportions based on the maximum likelihood estimate under WST and under LS. The observed values of π exceed the maximum likelihood estimate of π for all but one S and the LS model predicts the observed proportion better than does WST for 11 out of 15 Ss (Tversky, 1969, 39-40).

Model generality arguments: The additive difference model can accommodate a wider variety of preference structures. Notably, the LS model is a limiting case of the additive difference model, when the difference function approaches a step function where $\phi(\delta) = 0$ whenever $\delta \leq e$.

Information processing arguments: Intradimensional ("component-wise") comparisons are "simpler and more natural" (Tversky, 1969, 43) because they are expressed in the same units, and the choice process is simpler, especially when one alternative gives a better result on most of the components.

The second theme is about the relationships between each model of choice and intransitivity. The simple additive model satisfies transitivity (simple scalability). The question is thus to understand under what conditions the additive difference model can lead to intransitivities. Tversky provides a theorem on the necessary and sufficient conditions for transitivity when the process of choice is described by an additive difference model. In the general case of more than two dimensions, transitivity holds (in the case with three or more dimensions) if and only if all difference functions are linear, that is, there is a set of $0 < t_i < 1$ such that $\phi_i(\delta_i) = t_i\delta_i$ and $\sum_i t_i = 1$, thus imposing "extremely strong constraints on the form of the difference function" (Tversky, 1969, 43).[52]

To sum up, the structure of the argument runs as follows. The additive difference model and the simple additive model can describe the same choice behavior, but the choice process is more satisfactorily described by the additive difference model. Except under strong constraints imposed on the weighting function, the additive difference model will imply intransitivity "somewhere in the system" (Tversky, 1969, 43). Therefore, if one considers that the modeling of choice in terms of utility has something to do with the representation of the choice process, then the additive difference model has much to be recommended over the simple additive model and it reveals how stringent usual properties of intransitivity attached to the simple additive model are. However, one important question is whether the model is accurate in accounting for the tested transitivities and if it is likely to induce more intransitivities than usually observed. The LS model of choice, for instance, is an extreme form of the additive difference model, a pedagogic device calling for a "more plausible" (Tversky, 1969, 43) rendition of the intransitivity phenomenon. Notably, the theorem shows that intransitivity *does not* depend upon deliberately ignoring some dimension in the alternatives: "intransitivities can occur in the additive difference model in a fully compensatory system where all the information is utilized in the evaluation process". (Tversky, 1969, 44)

Eventually, Tversky discusses the normative strength of transitivity. Post-experiment interviews confirm that most *Ss* used a LS decision process and at the same time that they were not conscious of behaving intransitively, being confident that their choices are transitive: "People are not likely to admit the existence of consistent intransitivities" (Tversky, 1969, 46). Concerning the hot issue of rationality, Tversky left the question open, concluding that intransitivity is not necessarily irrational:

[52] In the case with two dimensions, the necessary and sufficient condition for WST is that there is some real t such that $\phi_1(\delta) = \phi_2(t\delta)$

> "Is this behavior necessarily irrational? It seems impossible to reach any definite conclusion concerning human rationality in the absence of a detailed analysis of the sensitivity of the criterion and the cost involved in evaluating the alternatives." (Tversky, 1969, 45-46)

Thus, the main outcome of the article is that violations of transitivity from people that are not to be regarded as irrational "reveal something about the choice mechanism and the approximation method that govern preference between multidimensional alternatives" (Tversky, 1969, 46). In some complex environments or when the alternatives of choice have many components, there are good reasons to privilege a process of decision that makes the choice easier even though we may be aware and accept that it is likely to imply intransitivity. Thus, Tversky reached conclusions similar to those of Slovic and Lichtenstein about the necessity to study how people make use of a rich set of information, somehow anticipating the importance of heuristics:

> "When faced with complex multidimensional alternatives, such as job offers, gambles, or candidates, it is extremely difficult to utilize properly all the available information. Instead, it is contended that people employ various approximation methods that enable them to process the relevant information in making a decision." (Tversky, 1969, 46)

If one considers the overall balance of arguments and references in "Intransitivity", Tversky must be credited for providing probably the first synthesis between the methods of economics and psychology in the field of decision theory, opening the possibility of a complete reevaluation of the assumption of transitivity in the theory of economic rationality. Tversky changed the way transitivity was considered from both a cognitive and a normative perspective. On the cognitive side, he opened up to the possibility of modeling intransitivities and unveiling their causes. On the normative side, he pointed to the fact that transitivity as a requisite of rational choice is contingent on some external conditions of the choice situation.

In two related papers, Tversky provided a descriptive model of choice named "Elimination by aspects" (Tversky, 1972b,a) which he presented as an alternative to other random-type models due to Thurstone (1927b) and Luce (1959). The main idea of elimination by aspects is that in some situations, when *Ss* have to choose among complex multidimensional alternatives, they might use a step-by-step lexicographic process of elimination of alternatives that do not contain some desired feature. The order in which these desired features are considered is important for the final choice. Instead of being determinate, it is given by some probability law reflecting the importance of each feature among other features. The main outcome is to obtain a model of choice that is lexicographic—reflecting a covert process for making a decision—and still representable as a compensatory model of choice. Such a model, however, is likely to be descriptively accurate only if each feature can reasonably be taken as independent from the others in the individual's mind, thus excluding any configural or Gestalt effect. The elimination-by-aspects process makes it possible to conceive that any final choice will be possible from any initial available alternative: "no matter how inadequate it might be, one can devise a sequence of selected aspects or, equivalently, describe a particular state of mind that leads to the choice of that alternative." (Tversky, 1972b, 298) In this respect, Tversky took into account the fact that in a complex choice environment, people are limited in their computational abilities (Slovic and Lichtenstein, 1971) and will therefore make a decision based on their current state of mind. For instance, they may be influenced by an advertisement inducing them to order the relevant aspects of alternatives in a specific way. Of course, the final choice does not claim to be superior to other alternatives and has, thus, no normative value. It merely reflects, once more, that a person's choice process can be influenced by carefully monitoring this person's state of mind, especially when the choice alternatives are difficult to compare.

4 Preference reversal

The preference reversal phenomenon was first identified and labeled in a study by Sarah Lichtenstein and Paul Slovic published in the *Research Bulletin* of the *Oregon Research Institute* in October 1970, entitled "Reversals of Preference Between Bids and Choices in Gambling Decisions" (Lichtenstein and Slovic, 1970). It was published shortly after in the *Journal of experimental psychology* (Lichtenstein and Slovic, 1971). Lichtenstein and Slovic wre very clear in interpreting their results as a shift from the current trends of research on decision making. The way *Ss* perform various tasks of rating, bidding or choosing among gambles shows that "such decisions do not rely solely on expected utilities" (Lichtenstein and Slovic, 1971, 46); experimental results on gambling suggest instead that "information processing considerations [influence] the method by which a gamble is judged." (ibid., 47) and that some anchoring phenomenon biases the elicitation of preferences. Additional experiments with gamblers at the Four Queens Casino, Las Vegas, supported the hypothesis (Lichtenstein and Slovic, 1973)

I will first present the method of analysis and the experimental results. I will then discuss some complementary reflections on the interpretation of preference reversal in connection with intransitivity and the meaning of preferences in this context.

4.1 Preference reversal as a robust phenomenon of individuals' judgment

The motivation for the Lichtenstein-Slovic experiment is that the economic theory of rational choice assumes that individuals' choices mirror an underlying scale of valuation of alternatives. Thus, the theory of rational behavior assumes that agents will order various alternatives exactly the same way whatever the method used to motivate them to use their subjective valuations. A *S* would reveal the same preference ordering if asked to perform a rating of attractiveness of gambles, or asked to bid to buy or sell gambles, or else asked to choose among gambles. Opposing this view, Lichtenstein and Slovic asserted that the revealed or stated preference ordering will depend on the kind of task *Ss* have to perform: tasks are not neutral to the result. Notably, as the study showed, there is a systematic and significant difference between, on the one hand, the tasks of choosing or rating gambles and, on the other hand, the tasks of bidding to buy or sell gambles.

Lichtenstein and Slovic carried out three experiments in order to compare *Ss*' stated preferences over pairs of bets with *Ss*'s preferences revealed through "pricing" of the same bets. The idea of the experiment originated in previous experimental results that did not fit well with neither the theoretical postulates of rational choice in economics, or with the main ideas about *Ss*' abilities to calculate. As we shall see, it complements Tversky's ideas about the difficulty to compute harmoniously several dimensions in choice, introducing the idea of anchoring and (mal)adjustment. At the same time, it adds to the reflection about the modeling of information processing. Last, Lichtenstein and Slovic (1971) offer one of the first explicit discussions of decisions as *biased* decisions, an orientation clearly reaffirmed in a follow-up article "Response-induced reversals of preference in gambling: An extended replication in Las Vegas" (Lichtenstein and Slovic, 1973) (hereafter "Las Vegas").

The starting point of the experimental device is that choice among gambles (as implemented in expected utility theory) is not the only way to discover a person's "opinions about the utilities of various gambles" (Lichtenstein and Slovic, 1971, 46). For instance, for any set of bets, the person could rate the bets on an arbitrary scale according to their degree of attractiveness to him; or he could announce bids to buy (maximum buying prices) or bids to sell (minimum selling prices) for these bets. According to utility theory, these various techniques should be taken as equivalent ways to know about an individual's preferences, since the responses are governed by "the same underlying values" (ibid, 46). The first aim of Lichtenstein and Slovic was to discard this assumption of procedural invariance, i.e. the assumption that the kind of task performed to reveal value or attractiveness is neutral to the result. Here, as in Tversky's analysis of intransitivity, the source of incoherent responses stems from the fact that a gamble "is a multi-dimensional stimulus whose various attributes have differential effects on individual decision-making behavior." (ibid, 46). The second aim was to identify patterns of information processing according to the task to be performed. In a previous study, Slovic and Lichtenstein had identified that choices and attractiveness are correlated with gambles' probabilities, whereas selling bids (of attractive gambles) are positively correlated with payoffs Slovic and Lichtenstein (1968). They conjectured that the task to be performed influences the relative importance of the various dimensions as the individual perceives them and processes them. Thus, they engaged in a new experimental design to give more substance to the idea that different tasks involve different ways to process information and decide on it (choosing or bidding). The specific assumption here is that when the task is to evaluate a gamble through a single dimension, namely money, *Ss* tend to focus primarily—and subsequently weight more—on the subset of characteristics of the gamble that has the same dimension, hence on the monetary payoff. The idea that different information processes are used for judging and comparing led to the conception of an experimental design that, in a sense, proves the assumption by exploiting it to trigger specific information processing on the part of *Ss*, a strategy similar to that of Tversky (1969). The experiments were devised so as to provoke a (revealed) preference reversal by comparing *Ss*' preferential choices and bids for pairs of gambles.

Consider two broad categories of gambles, labeled P-bets (probability bets) and \$-bets (dollar bets). P-bets are gambles with high probabilities (ranging from .99

to .80) of small gains (e.g., .99 to win \$4, .01 to lose \$1) while \$-bets are gambles with balanced probabilities (ranging from .33 to .10) of higher gains (e.g., .33 to win \$16; .67 to lose \$2). According to the hypothesis about information processing, "one might expect that *Ss* would choose Bet P over Bet \$, but bid more for Bet \$." (ibid., 47)

Lichtenstein and Slovic (1973) carried out three experiments with bets of positive expected value (selecting only those that were judged to be "attractive" by all *Ss*). Each pair of bets presented to *Ss* (bet A and bet B) contained bets with the same or approximately the same expected value. For each pair, each *S* indicated first his preference between A and B, then had to bid for each, one at a time. Experiment I settled the main hypothesis by comparing choices and selling bids. Experiment II compared choices with buying bids. The third experiment aimed to control *Ss*' motivation. I present the results in detail.

Experiment I consisted in comparing choices and selling bids for 13 bets. *Ss* were 173 male undergraduates who were paid to participate in the experiment but did not play gambles for real. They had to choose from pairs containing a \$-bet and a P-bet with expected value ranging from \$1.40 to \$4.45. *Ss* were asked to indicate their preferred bet and how strongly they preferred it (from "slight" to "very strong").[53] Then, the second part of the experiment consisted in bidding for each of the 13 bets. *Ss* had to imagine that they owned a ticket for the bet and "to name a minimum selling price ... such that [they] would be indifferent to playing the bet or receiving the selling price." (ibid, 47) For each pair and each *S*, four possibilities are:

i) choosing P-bet, bidding more on \$-bet;

ii) choosing P-bet, bidding more on P-bet;

iii) choosing \$-bet, bidding more on \$-bet;

iv) choosing \$-bet, bidding more on P-bet;

Only case i) strictly conforms to the hypothesis that preference reversal is linked with specific information processing. Case iv) should be explained in another way. The proportion of preference reversal is obtained from cases i) and ii), and is called "proportion of conditional predicted reversals." An *S* can exhibit either systematic preference reversal (if s/he satisfies i) in all instances) or occasional preference reversal. Of the 173 *Ss*, 127 (73 %) showed systematic preference reversal (ibid, 48). Case iv) is a case of conditional unpredicted reversal behavior that goes strictly against the hypothesis and, to the authors, it "is hard to rationalize under any theory of decision making" (ibid, 48). Actually, 83% of the *Ss* never showed unpredicted preference reversal.[54]

Experiment I set the stage for investigating the effect of other tasks, possibly involving differential information processing. Experiment II explores the generality of preference reversal by asking *Ss* to make buying bids instead of selling bids

[53]Actually, these labels were suggested to be "moderate" and "strong", and answers were coded with values 1-4.

[54]Also, *Ss* who reversed their preference according to the cognitive hypothesis, choosing the P-bet and bidding more on the \$-bet, were those who expressed on average the strongest preference for the P-bet they had chosen in each pair (mean = 3.10).

and by using a greater variety of P-bets and $-bets. *A priori* considerations are that performing a Buying-bid instead of a Selling-bid means imagining having to pay to acquire the ticket, and therefore—anticipating some sort of endowment effect—preference reversals should occur less often. The experiment was carried out with 74 students along similar lines, only $-bets and B-bets are presented differently (probabilities being expressed as ratios) and the range of expected values is greater. Results are consistent with expectations. Buying-bids were lower on average than Selling-bids for the same expected value.[55]

Other specific results with the buying bid task are worth mentioning. A higher proportion of unpredicted conditional reversal is observed (though still less than predicted reversals). When choosing, *Ss* were sensitive (negatively) to the absolute value of the loss in both kinds of bets. This sensitivity to loss did not exist when bidding. Conversely, statistical analysis showed that the amount to win (the ratio of winning amounts of $-bets over P-bets) influenced bidding but not choice. The comparison of preferential choice with buying-bids (experiment II) reinforces the idea that "different modes of information processing are used in bidding and choosing." (ibid, 50). This first behavioral hypothesis (differential sensitivity to loss and win amounts) leads to the conclusion that the preference reversal phenomenon manifests itself acutely when the pair contains a large $-bet loss relative to the P-bet loss and a large $-bet win relative to the P-bet win. This is in tune with experimental results; indeed, 60 % of the *Ss* showed predicted preference reversal with the following pair:

P-bet: 9/12 to win $1.10 and 3/12 to lose $.10

$-bet: 3/12 to win $9.20 and 9/12 to lose $2.00

Hence, this result favors the idea of a different treatment of information (in a determinate direction) when making preferential choices and when making a monetary valuation (be it through a selling bid or a buying bid).

Experiment III aims to control *Ss*' motivation for making the best choices and valuations, and feeling concerned about it. It was carried out with 14 male students each of whom received "lengthy and careful instructions" (ibid, 51). Notably, gambles were played for real, using a roulette wheel, and a conversion procedure warranted that all *Ss* would gain; each pair of bets was presented three times. On the third round the experimenter warned *Ss* of their past choices before making their final choice. Likewise, the Selling-bids were organized at best to obtain reliable information about the true value.[56] The results obtained from experiment III with well-motivated *Ss* confirm preference reversal: 6 *Ss* out of 14 always made conditional predicted reversals, 5 *Ss* sometimes made them. A final touch to the establishment of the preference reversal hypothesis is to compare the observed responses with a pure random choice model based on *Ss* failure to make choices and

[55] In Experiment I, S-bids were $3.56 above expected value while B-bids in Experiment II were 4 cents below average expected value.

[56] Notably, the Becker et al. (1964) method was used in order to motivate *Ss* to announce their true willingness to pay. This method consists in making a randomly generated counteroffer to the Selling-bid. If the counteroffer is lower than the Selling-bid, the bet is played, if it is higher, it is paid.

bids in accordance with their true preferences. Assuming probabilities for such mistakes and combining them, Lichtenstein and Slovic obtained a base random mistake model with a distribution of cases that can be compared with actual observed proportion. The inferred values of probability parameters cannot account for observed values. Post experimental interviews with *Ss* were conducted to persuade them to change their decisions in case of conditional predicted reversal. Two *Ss* refused because their choices reflected their true feelings. Interviews confirmed that *Ss* did not focus on the same information when choosing and bidding.

In 1972, Lichtenstein and Slovic carried out another experiment, with the aim of testing their hypotheses in situations where *Ss* were highly motivated and stakes were higher.[57] The experiment took place at the Four Queens Hotel and Casino, Las Vegas (Nevada) and was part of a project directed by Ward Edwards. Tversky and Kahneman's comments were acknowledged by the authors. The new hypothesis about unattractive bets was that when pricing an unattractive bet, *Ss* would "use the amount to lose as a starting point and make an adjustment upward in an attempt to account for the other aspects of the bet." (Lichtenstein and Slovic, 1973, 16). If this adjustment is insufficient—especially when the \$-bet loss is large, this will bias the buying- and selling-bids. In contrast, no such bias is expected in the choice task. Series of P-bets and \$-bets with the same absolute expected value are built. The experimental design ensures that players are motivated: they are clients at a casino, volunteering to participate in the experiment, they play real money, and they can choose the stakes by selecting the unit value of chips.[58] During the stage of bidding, *Ss* indicated how many chips they would like to pay to avoid playing an unattractive bet or how much they would want to be paid for not playing an attractive bet, and they knew that a randomized counteroffer would decide whether the bet would be played or paid. "With this set of counteroffers, the player who states the EV of the bet as his price for each bet will maximize his expected winning." (Lichtenstein and Slovic, 1973, 18)

Over 10 weeks, 53 games were completed (involving 44 *Ss*, of whom 7 were casino dealers). The statistical treatment of responses (after removing situations where a pair of bets contains an attractive and an unattractive bet) confirmed former results about preference reversal. Lichtenstein and Slovic obtained 81 % of predicted preference reversal.[59] In the case of unattractive bets, the rate of predicted reversal was 76%.[60]. In sum, the results obtained with motivated players at a casino are similar to those of previous experiments with college students, and the preference reversal phenomenon applies as well to unfavorable (unattractive) bets. Moreover, additional biases that could be at work in the case of attractive bets (a

[57]Another aim also was to introduce unattractive bets in the experiment and to address the issue of preference reversal with these bets. Harold Lindman, another student of Ward Edwards, had obtained results similar to Lichtenstein and Slovic with unattractive bets (though *Ss* did not play real money) (Lindman, 1971).

[58]Actually, though they could choose unit values of \$1 or \$5, all *Ss* preferred to play with 5¢, 10¢ or 25¢chips. Players had to buy 250 chips to start with.

[59]Out of 229 pairs for which *Ss* had chosen the P-bet over the \$-bet, there were 185 instances of *S* pricing the \$-bet higher than the P-bet

[60]Out of the 112 pairs of unattractive bets for which \$-bets were chosen over P-bets, *Ss.* gave higher prices to P-bets on 85 instances.

pure pleasure for playing, thus overbidding) did not operate with unattractive bets, and however predicted reversal were similar. This gave even more support to the cognitive approach to biases;

> "The widespread belief that decision makers can behave optimally when it is worthwhile for them to do so gains no support from this study. The source of the observed information-processing bias appears to be cognitive, not motivational." (ibid, 20)

4.2 Analyzing preference reversals: between rationality and heuristics

The conclusions drawn from the experiments are threefold, and open to comments on the complementarity between preference reversal phenomenon and the case of intransitive choices.

First the reversal of preferences phenomenon is incompatible with the subjective expected utility model, since this model assumes the same fixed system of functions for probabilities and utilities whatever the mode of response used to evaluate lotteries. Actually, since the cause for reversal originates in the response mode adopted, the preference reversal should also apply to situations of choice under certainty. To the authors, "these reversals clearly constitute inconsistent behavior and violate every existing theory of decision making." (ibid, 54) It is beyond the scope of the present article to discuss this complex subject. Be it enough to mention that the first reactions of economists to this result were hostile. In any event, it shows that utility theory was a target for Lichtenstein and Slovic, in accordance with several other studies by the same authors, notably in connection with behavior on financial markets.

Second, regarding the specific mechanisms in choosing tasks and in monetary valuation tasks, Lichtenstein and Slovic noted that in the case of choice among alternatives, the whole set of information common to both alternatives can be used as a basis for valuations and combinations. In contrast, "bidding techniques provide an obvious starting point: the amount to win." (ibid, 54) This amount serves as an anchor to set a first valuation, and it is then adjusted in order to account for probabilities and amount to lose, which together form a complex multidimensional set of information which is not used properly, and the adjustment that would have been necessary with $-bets is insufficient. In other words, the preference reversal phenomenon is presented here as a consequence of some heuristics of anchoring and adjustment, in which anchoring is determined by the response mode: attributing a monetary value in the bidding task for an attractive bet leads to take (mentally) as a starting bid the amount to win and to account grossly for other variables in a second step.

Third, the hypothesis of differential information processing when different modes of responses are used challenges the meaning of rationality, which needs to be built upon some standard of preferences. On this, Lichtenstein and Slovic sided with Tversky's conclusions in "Intransitivity of Preferences":

> "[Tversky] noted that it is impossible to reach any definite conclusion

concerning human rationality in the absence of a detailed analysis of the cost of the errors induced by the strategies S follows as compared with the cost to S of evaluating alternative strategies. The approximations Ss follow in order to simplify the difficult task of bidding might prove to be rather efficient, in the sense that they reduce cognitive effort and lead to outcomes not too different from the results of optimal strategies. *In using such approximations, the decision maker assumes that the world, unlike the present experiments, is not designed to take advantage of his approximation methods.*" (Lichtenstein and Slovic, 1971, 55, our emphasis)

The last part of this quotation is important to our story. The default in using all relevant information poses itself in cases of monetary valuations of alternatives, i.e. when the number of dimensions that characterize the item are greater than the number of dimensions that support the valuation (typically a single dimension in the case of a monetary evaluation). Hence, it is in this case—not in the case of a preferential choice—that Ss can be said to use information only incompletely. If ever a bias in judgment was to be attributed to Ss, one can speculate that this bias would manifest itself in the case of a monetary valuation, not in the case of a preferential choice, or only to a lesser degree and in an indeterminate direction.

Commenting on the Las Vegas experimental results, Lichtenstein and Slovic push the idea a bit further, hypothesizing that such decision biases would be induced by manipulating the commensurability between a response mode and the various dimensions of the cues at disposal of the decision maker. In general, they consider that the choice of a uni-dimensional response mode (e.g. money) leads to a bias as compared with response modes that are less specific.

> "The strain of amalgamating different types of information into an overall decision may often force an individual to resort to judgmental strategies that do an injustice to his underlying system of values. The systematic bias in the pricing responses of the present study is one demonstration of this. Most individuals are unaware of the biases to which their judgments are susceptible."(Lichtenstein and Slovic, 1973, 20)

In this respect, Lichtenstein and Slovic's contribution complements Tversky's. Starting from the same theoretical apparatus (information processing, multidimensionnality, weights of arguments), Tversky identified conditions in favor or against the existence of a coherent preference ordering for a fixed response mode, whereas Lichtenstein and Slovic identify conditions under which a more reliable pattern of preferences, i.e. more in tune with individual's overall values, is likely to be elicited through some experimental setting. At this stage, only two revelation principles were studied (monetary valuation and preferential choice), and the case of pure exchange (item-item) was not considered.[61] Lichtenstein and Slovic's experiments tend to show that rationality cannot be thought of independently from the conditions under which some system of preferences is obtained, thus leading towards a

[61] As we know, the case of pure exchange triggered the identification of the endowment effect.

much more complex representation of rational behavior. On this account, we must postpone a broader discussion to the concluding comments.

Also, the issue of rationality relates to issues that are deemed irrelevant in economics. First, individuals are assumed to compare psychological costs and monetary costs. Second, individuals are assumed to appraise the characteristics of the environment; notably, it is assumed that they take stance on their feelings about the possibility to be subjected to some manipulation. This is excluded in economics in a pure competitive environment (parametric behavior), and is addressed only in strategic environments. Instead, to a psychologist, rational behavior is defined in relation to a person's beliefs about his environment, hence the boundaries for a pure parametric behavior are blurred.

In another respect, there is a notable difference between Tversky's study on intransitivity and Lichtenstein and Slovic's study on preference reversal, even though both studies deal with information processing. Tversky's experiment deals only with judgments of preference while Lichtenstein and Slovic deal with both judgments of preference and monetary valuations and decisions. While Tversky's interpretation is focused on the possibility for a person to deal with multidimensionality of desirable attributes, Lichtenstein and Slovic focus on the influence of the choice task on the use of information. In the case of Tversky, the question is whether a person could use and weight all the information at his disposal and however exhibit an intransitive pattern of choice. In the case of Lichtenstein and Slovic, the question is whether a specific stimulus (the response mode) can influence the relative importance of the various pieces of information used by individuals. Since the response mode is imposed on individuals by the experimenter, it means that responses can be influenced by the environment. In this sense, a monetary valuation twists the use of information, inducing *Ss* to focus on some information—some "natural starting point"—and to ignore or downplay some other information. In this respect, Slovic and Lichtenstein seem to be Simonian, in the sense that they do accept the principle that *Ss* will always use shortcuts to make judgments and always fail to be rational in the sense of utility theory. In this respect, there is *no* neutral response mode, *no* specific task inducing the revelation of true preferences. It is a task for the psychologist to determine whether some specific tasks induce an unbalanced or incomplete use of information: "The reversal phenomenon is of interest, not because it is irrational, but because of the insights it reveals about the nature of human judgment and decision making." (ibid, 55) For all that, one cannot assert that some elicitation mechanism delivers the true preferences. [62] By assuming that the observed reversal

[62] In economics, the assumption that the way preferences are revealed (procedural invariance) does not change preferences had been heralded in the axiomatization process of the 1930s and 1940s, notably with Samuelson's revealed preference approach (Samuelson, 1938, 1965, 1950). According to the revealed preference approach, agents and theoreticians are put on equal terms: the agent reveals to herself what her preferences are and reveals them likewise to the theoretician-experimenter simultaneously. The agent is not assumed to know better (neither is the theoretician). Older approaches (e.g. Pareto) considered that the individual agent does know something about her preferences, obtained through trial, error and adjustment, knowing potentially more than what the theoretician might discover about her through various methods of investigation. The endless debate regarding the proper way to elicit preferences (through market behaviors or through questionnaires) was not unraveled before Samuelson's kind of behaviorist formulation of it (Lenfant, 2012). It stands as a heroic statement, assuming that preferences are elaborated without consid-

of preference ordering is biased in the task of bidding, they somehow admit that the amount of information leading to the choice is reduced and that therefore not all relevant information has been accounted for.

5 Heuristics, Cognitive Biases and Agents' Decisions

During exactly the same period, between 1966 and 1971, an extensive body of literature documented the fact that human beings are biased in almost all kinds of judgments involving probability calculations. They can even be extremely bad in some supposedly simple inference tasks. At the beginning of the 1970s, several kinds of biases had been identified and sometimes replicated. These biases were studied first and foremost with physicians, especially in order to test the reliability of their diagnoses (Goldberg, 1968). Fewer studies endeavored to adopt the same pattern of experiments in the field of business and finance, especially with a view to understanding finance analysts' performances. Slovic contributed to this literature. In this section, our goal is to focus on the most salient of these studies and to identify the kind of theoretical representation that emerges from this literature. Notably, the importance of the distinction between heuristics and biases is at stake, and Slovic, Tversky and Kahneman's contributions are central for understanding this distinction.

The fact that individuals are poor statisticians was well documented in the 1960s. Studies on human judgments and behaviors involving probability calculations are numerous and can be grouped under two headings. In the first place, several studies focused on the way experts in various medical fields perform clinical judgments. In one such study, Goldberg (1968) made a useful distinction between three types of reliability. Consistency refers to the stability of a judge across time when faced with the same set of data. Consensus refers to the agreement across judges using the same data. Convergence refers to the stability of a clinical judgment by the same judge based on different sets of data. Goldberg's conclusions are clear: often, clinical judgment is not convergent, suggesting changing ways of information processing.

The literature on clinical judgments has been analyzed through the lens of an opposition between the objective or scientific use of data in order to perform probabilistic calculations against the intuitive or subjective use of data to infer a diagnosis. Taking clinical judgment as a reference point assumes that the value of things to be judged is known. For instance, the experimenter can assess whether a diagnosis of some disease by a radiologist is a good or bad judgment, and it becomes possible to enquire about the data that are likely to make the clinical judge underperform. These studies tend to enquire about the right way to combine clinical judgment and statistical methods to improve diagnosis.

A first bias concerns Bayesian learning and the tendency to be "conservative", that is, to give too little importance to new information to update our views about the statistical properties of a given environment. As we have seen, the study of individual Bayesian judgment was launched as a research program by Edwards.

---

ering the moral values and conceptions of life of the *Ss*.

Individuals are conservative in the sense that their revisions of probability estimates based on new information go in the direction prescribed by Bayes but are too small (Edwards, 1968).

Between 1971 and 1974, Amos Tversky and Daniel Kahneman published a series of articles that, taken together, established a new theoretical approach to the study of human judgment and decision, the heuristics and bias approach, breaking with Edwards' ideas. Before presenting this approach, a word is in order here about Kahneman and his collaboration with Tversky.

Born in 1934, Kahneman spent his early youth in France, hiding during the war, then leaving for Palestine in 1946. He graduated in psychology at the Hebrew University of Jerusalem (1954), then was enrolled in the Army for his military service as a psychologist, conceiving of tests for the orientation of officers. There, he designed a process based on interviews to select and allocate combat-unit recruits to various combat jobs. On this occasion, he got aware of a persistent cognitive illusion. Young recruits were selected and oriented toward a specialty in the face of mere observation of their personality traits in performing a collective task, and the process of selection was not adapted despite continuous statistical feedback by military authorities regarding the evaluation of each recruit (Kahneman, 2002). In 1958, he went to the university of California (Berkeley) to prepare a PhD in psychophysiology (1961). In 1961, he became a lecturer at the Hebrew University of Jerusalem. During the period 1961-1970, he made regular research residencies abroad (Michigan, Harvard, Cambridge (UK)). In 1965, he went to the University of Michigan for postdoctoral study with Gerald Blum. His main research topics in the 1960s were related to visual perception and its relationship with attention, learning and performance. However, at the time of engaging in a collaboration with Tversky, he had already experimented informally with some judgmental biases (through teaching or observation of clinician's diagnoses.) As Maya Bar-Hillel puts it (2024, 2), "Danny brought his fresh observations and original insights into an area that had been governed by a different approach—formal and normative, which was Amos's forte."

Kahneman and Tversky started collaborating in 1969. Kahneman had invited Tversky to present his research in his seminar at the Hebrew University. Strangely enough, Tversky chose to present some work done on Bayesian decision at Edward's lab. The story goes that Kahneman welcomed Tversky's speech by saying: "Brilliant talk, but I don't believe a word of it." Kahneman's opinion was that real people are not statisticians at all, that they miss too much of a statistical way of thinking to ever come close to having their beliefs influenced by statistical thinking. This was the starting point for their collaboration on the psychology of judgment and decision. As we have seen, Tversky in 1969 had already started challenging certain views about people's abilities to compare multidimensional alternatives, but he still thought that more simple situations could be handled through the holistic measurement paradigm reigning in mathematical psychology. During the academic year 1970-71, Tversky was a visiting researcher at Stanford. Tversky and Kahneman accumulated a lot of experimental and questionnaire material during this year. Then started the most fruitful period of their collaboration, hosted by the Oregon Research Institute (Eugene), during which they would write several papers on the heuristic and bias approach. Back in Jerusalem, they prepared a synthesis of their research, to be published in *Science* "Judgment under uncertainty: Heuristic and biases" (Tversky and Kahneman, 1974).

The heuristics and biases approach was presented as a general approach that can be applied to any problem-solving situation. The claim made by Kahneman and Tversky is that in any situation of choice, when the goals can be precisely defined, individuals usually use mental shortcuts to appraise a situation, make a decision or value the alternatives available to them. The generality of this theoretical hypothesis is noticeable. For all that, it is intended to be limited to a theory of judgment and decision once a problem situation has been identified; it does not claim to explain the reasons why a person interprets a situation the way she does. The heuristics and biases approach aims at establishing some regular patterns of thought, common ways to process information and frequent mistakes that ensue from it. It is based on specific experiments pertaining to the use of certain information directly available or retrieved. it is based also on some experiments involving individual calculations.

In this section, we present the main results of the literature as they were organized by Kahneman, Tversky, Lichtenstein and Slovic from 1970 to 1974. We have seen already that some pointers about decision processes had led Lichtenstein and Slovic to advance a heuristic of anchoring and adjustment to explain *Ss*' bidding behavior. Two other heuristics were put forward by Tversky and Kahneman: the heuristic of representativeness and the heuristic of availability. Historically, the heuristic of representativeness helped to structure the whole program later called "heuristics and biases". The following describes the historical progress towards constructing the heuristic approach and making a theory out of it.

5.1 The heuristic of representativeness

The heuristic of representativeness was an important step in the shift away from the program of research on individuals' calculation abilities and in the launching of a new approach. The starting point was Tversky and Kahneman's study of judgments about the reliability of statistical tests in academic research, especially in psychology. In a famous article entitled "Belief in the Law of Small Numbers" (Tversky and Kahneman, 1971), their first co-authored article, they identified very common wrong intuitions or expectations among *Ss* about chance, statistical calculations and properties of samples drawn from a population. They proposed a first interpretation of these mistakes. Most *Ss* believe that samples must be representative of the population from which they are drawn. This first intuition about representativeness was developed in follow-up articles to establish the pervasiveness of a mode of construction of beliefs and judgments based on the way people think that a thing or event A is representative of a set of elements or events B.

5.1.1 Building the hypothesis of representativeness

"Belief in the law of small numbers" is important in two respects. Firstly, it addressed the specific and fundamental issue of the use of statistical results in scientific practice. Secondly, it set the authors on the path to an intuition about the use of information in everyday life, each time people elaborate beliefs about the properties of things or causal relationships based on fragmented information about these things or based on information about the parent population from which they have been generated.

The starting point for this intuition about human beliefs was a deep investigation of the beliefs and practices of researchers in psychology when making a statistical test of significance. Tversky and Kahneman showed that the standard practice of tests and checks in this community was functioning as a self-sustaining system of wrong beliefs. Initial beliefs, though fragile, were reinforced through a series of practices about statistical tests with new samples, because searchers did not take into account enough the inherent variability of sample properties. Identifying the functioning of this system led to a first presentation of the cognitive pattern by which *Ss* tend to attribute properties to things. Even though the article dealt with scientific practices in psychology—but it could apply as well to meteorology or economics—the authors grounded these practices on common biases associated with the representation of probabilities, statistical values and randomness.

The first general belief about randomness is a belief of "similarity". People believe that random samples drawn from a population, even small samples, must be "representative" of—or "similar" to—the properties of the entire population. For instance, when *Ss* are asked to imagine a random sequence of tosses of a fair coin, they tend to generate a sequence with the same number of heads and tails, scattered more or less evenly in the sample. The average value for heads and tails stays closer to 1/2 than what the laws of chance imply, as if any sample, even short, should bear the mark of the “fairness” of the coin. This is what Tversky and Kahneman ironically labeled the “law” of small numbers, that is, the belief that a small sample has similar properties to large samples (which obey the law of large numbers). Another common intuition about the properties of random series and samples was that any apparent deviation from randomness at some point in a series must be corrected by a symmetric deviation soon after. This is the "gambler’s fallacy". A famous example of this bias is as follows.

- **Description**. “The mean IQ of the population of eighth graders in a city is known to be 100. You have selected a random sample of 50 children for a study of educational achievements. The first child tested has an IQ of 150. What do you expect the mean IQ to be for the whole sample?”

- **Responses**. a large number of *Ss* believe the right answer is still 100.

- **Correct answer**. The answer given by *Ss* assumes a self-correcting process, that the rest of the sample will correct the deviation from the average caused by the initial value of 150. “The laws of chance do not work [as *Ss* believe it]: deviations are not canceled as sampling proceeds, they are merely diluted.” (Tversky and Kahneman, 1971, 106) The correct answer is simply to expect that the rest of the sample of 49 children has an average IQ of 100, thus the expected value for the whole would be 101.

Tversky and Kahneman called the idea that *Ss* held these beliefs about randomness and random samples a “representation hypothesis”—the words "representativeness" and "heuristic" were not used in 1971—i.e. a theoretical hypothesis that people's responses in experiments reflect their beliefs or intuitions about chance.[63] Actually, there were two representation hypotheses: i) a representation hypothesis of similarity (between large and small samples), i.e. "the law of small numbers" and ii) a representation hypothesis about self-correction, the gambler's fallacy, which contains some implicit belief of causality between independent events, as if chance was animated by some will ("fairness").[64]

The gist of the paper consists in analyzing the way this two-part representation hypothesis (gambler's fallacy, law of small numbers) operates as a vicious circle of self-confirmation of wrong beliefs, with no correcting feedback. Tversky and Kahneman took as their field of study the behavior of psychological scientists, but their reasoning was designed to apply to other social sciences.[65]

It is beyond our scope to present the various results obtained through questionnaires. The dynamics of self-confirmation is the following. A scientist believing in the law of small numbers "gambles his research hypotheses on small samples without realizing that the odds against him are unreasonably high." (ibid, 109) He conceives some confidence from sizeable deviations from the null hypothesis. This confidence then works as an anchor to evaluate replications and search for "explanations" for future discrepancies. "Thus, [the scientist] has little opportunity to recognize sampling variation in action. His belief in the law of small numbers, therefore, will forever remain intact." (ibid, 109) The only hope for controlling the effects of the law of small numbers would be through statistical education and good practices. The authors are skeptic, though, since "related biases, such as the gambler's fallacy, survive considerable contradictory evidence." (ibid, 109)

> "The true believer in the law of small numbers commits his multitude of sins against the logic of statistical inference in good faith. The representation hypothesis describes a cognitive or perceptual bias, which operates regardless of motivational factors." (ibid, 110)

Two comments are in order. First, the article launched several important ideas

[63]Kahneman and Tversky (1972) gave more details about *Ss*' expectations regarding samples' properties. What counts as representative is the way a sample seems to have been generated through a random process, the way it "reflects randomness as people see it; that is, all its parts are representative and none is too regular." (Kahneman and Tversky, 1972, 436). Hence, people expect that even small sequences will be representative of a parent population, and at the same time they expect these sequences not to have too regular a pattern. For instance, a majority of *Ss* wrongly think that a sequence of eight coin tosses such as HTHTHTHT or TTHHTTHH does not reflect randomness, while HHTHTTTH does. Also, *Ss* asked to simulate random sequences of coin tosses systematically overrepresent locally representative sequences, avoiding long runs of H or T.

[64]This representation hypothesis is sustained by experiments and answers to questionnaires in the community of psychologists. It is also sustained by former experiments reviewed by Tune (1964) on the ability of *Ss* to generate random series of numbers or heads and tails.

[65]Regarding the use of statistics in science, notably in psychology, Tversky and Kahneman elaborated on the fact that researchers in psychology tend to be overconfident about their initial experimental results even though they have been obtained with a small sample of *Ss* responses, i.e. even though another sample of the same size might have led to different results invalidating their hypothesis against the null hypothesis. This belief, they continued, derives from a more general belief about similarity between samples, i.e, the law of small numbers. Contradicting results between two samples led to search for a causal "explanation"of it which is actually an "exercise in explaining noise."(Tversky and Kahneman, 1971, 108)

that would be studied more systematically in further papers, leading to the principle of a representativeness heuristic. At this stage, the distinction between the heuristic as a general pattern of thought and its associated biases was still blurred. More experiments on other kinds of beliefs relating to other kinds of problems and situations needed to be accumulated to elaborate a heuristic of representativeness. However, it is notable that the distinction between a pattern of thought and a bias was already assumed through the distinction between beliefs on the one hand and practices on the other hand (hence judgments of validity of a scientific result). By the end of the article, the gambler's fallacy is presented as a *bias* deriving from the belief in the "fairness" of chance; and this belief appears as a very strong intuition preventing corrective experience from shaking our belief that small samples represent large samples (the law of small numbers). However, both the gambler's fallacy and the law of small numbers come to mind as "intuitions" or "beliefs".

Second, the statement of a self-sustaining system of beliefs leading to a judgmental bias was interpreted by Tversky and Kahneman as a shift from the usual program associated with probabilistic learning and Bayesianism. People are not conservative in the sense that they are reluctant to change their views about a probabilistic distribution. Instead, they are prone to project onto statistical results *more* information and properties than they actually contain. This view is definitely a change of perspective on judgment.[66]. It would become the central theme of their second co-authored article "Subjective probability: A judgment of representativeness". (Kahneman and Tversky, 1972)

Kahneman and Tversky (1972) widened the experimental situations studied and elaborated a general principle of thought based on the idea that human beings tend to perceive objects by associating them with existing categories or classifications of known objects. The "representation hypothesis" now became a "judgment of representativeness" or a "representativeness heuristic". This identification process is obtained by focusing on some salient features of an object that are judged to be common with salient features of the class to which this object can belong. This cognitive process assumes that human beings somehow shrink the area of characteristics contained in objects (be it in the reference object or in the object to be judged or in both) in order to reach this assignment.

Two domains of application of the representativeness heuristic were addressed in the article. The first was a continuation of the study of *Ss*' beliefs about randomness and sampling. The second opens onto issues about Bayesian judgment of probabilities. I shall present the main results in each domain in order to pinpoint the arguments establishing the heuristic of representativeness.

"The law of small numbers" had left us with the idea that people expect too many properties from small samples because they simply convey these properties from large samples. At the same time, people expect some irregularity to be prevalent in the way properties of large samples are transmitted to small ones. How far

[66]It is mentioned in a footnote: "Edwards (1968) has argued that people fail to extract sufficient information or certainty from probabilistic data; he called this failure conservatism. Our respondents can hardly be described as conservative. Rather, in accord with the representation hypothesis, they tend to extract more certainty from the data than the data, in fact, contain." (ibid, 109 fn4)

does this expectation of irregularity go? In order to establish firmly that *Ss* derive the same expectations of variability for two different samples that are perceived as equally representative, Kahneman and Tversky devised a new experiment about sampling distribution. According to the hypothesis of representativeness, *Ss* will be indifferent to size effects, they will expect the same levels of variability (variance of an outcome) in small and large samples. Hence, one must devise an experiment to reveal expectations, based on a known objective knowledge about probabilities in the parent population. The well-known illustration of this bias associated with population sampling is the large vs small hospital maternity unit example (although several other experiments with different scenarios replicate the results).

- **Description**. In a town with two hospital maternity units, the larger one records about 45 births every day, while the smaller 15. The percentage of boys varies from day to day (more or less 50%). For one year, each hospital records the number of days when boys represent 60% or more of the daily births. Which hospital do you think recorded more such days?
  - The larger hospital
  - The smaller hospital
  - About the same (i.e. 5% of each other)
- **Responses**. About 20% of *Ss* indicate the correct answer (the small hospital) 24% choose the large hospital and 56% choose "about the same number of days (± 5%)"

The question calls for an ordinal answer. The modal answer ("about the same number of days") shows that to a majority of *Ss*, the size of the sample is irrelevant to consideration of the variance of the proportion boys/girls. Since they regard a sample of 15 and of 45 as equally representative of the whole population, they expect from these samples the same properties, notably the same deviations from the mean (the same variance). The construction of beliefs regarding sampling is thus enriched: "The notion that sampling variance decreases in proportion to sample size is apparently no part of man's repertoire of intuitions." (ibid, 444).[67] Kahneman and Tversky developed their intuitions about the role of representativeness, building on existing experimental results and adding their own set of experimental results in favor of a heuristic approach to statistical inference.

Several important aspects of the heuristics and biases approach are put to the fore in dealing with statistical inference. First, as for any new problem-situation, a new task being performed in a new informational context requires one to define which information is likely to become prominent in terms of representativeness and which is likely to be ignored although it should not be ignored (from an objective

[67] This disconnection is reflected in the fact that people trust information displayed in percentage, deprived of any information about the sample size, and at the same time can distrust information generated from large samples on the ground that it is just a sample and thus it is not representative of the whole population. This ignorance could be partly corrected by explaining people the influence of sample size on variances; but only partly, since the tendency to overestimate the results obtained with small samples is firmly established.

measurement standard). Second, once this theoretical framework has been defined, a set of experiments can be carried out to corroborate the theory and draw conclusions in favor of a heuristic approach and against the dominant view that behaviors are influenced by "appropriate variables and in appropriate directions." (Peterson and Beach, 1967, 43) I summarize the contribution of Kahneman and Tversky on these two points.

First, the statistical inference approach (or Bayesian approach) can be used for two kinds of tasks. The first deals with the revision of opinion based on new information (assessing probability after a first draw, a second draw, etc.): this is a dynamic-sequential problem. The second deals with the evaluation of evidence. For instance, *Ss* are presented with the whole sample drawn from one population and make a single estimate of posterior probability. While the revision of opinion attracted attention and was instrumental for the standard view of man as a conservative Bayesian statistician (Edwards (1962a); Edwards (1968); Edwards (1971)), Kahneman and Tversky deliberately opted for experiments on single estimates. Consider for instance the following description:

- **Description**: Bag A contains 80% of red poker chips and 20% of blue poker chips; bag B contains the reverse proportions. Bag A and B contain the same number of chips. One bag is selected by chance and a random sample is drawn from it. From observing the number of red and blue chips in the sample, you estimate the probability (or the odds) that the sample has been drawn from bag A.

Typically, in this case, there are two well defined populations with some known characteristics and the task is to infer from which population a given sample has been drawn. The correct answer consists in calculating the probability of obtaining the observed sample in each population (each bag), then computing the ratio of probabilities (likelihood ratio).

If a sample (not too small) has been drawn, it contains enough information to be compared with the two populations from which it may have been drawn: if people infer probability from representativeness, they will tend to compare the proportions in the sample with the known proportions of the originating populations. Moreover, as with subjective sample distributions, it is expected that *Ss* will tend to ignore the size of the sample (the absolute values of red and blue chips) and will concentrate on proportions of red and blue chips to infer probability. Hence Kahneman and Tversky's conjecture—in tune with previous experiments and theoretical building blocks—"that subjective posterior probability should depend on the same salient characteristics of the sample which determine subjective sampling distributions." (ibid, 446). In the case of the blue/red chips example (with symmetric values for probabilities in bags A and B) "the objective posterior probability depends only on the difference between numbers of red and blue chips observed in the sample." (ibid, 446-447).[68]

[68] If b is the number of blue chips and r the number of red chips in the sample, p being the probability of red chips in A (and the probability of blue chips in B), the likelihood ratio is $(\frac{p}{1-p})^{(r-b)}$. Hence what counts for the objective result is the absolute value $r-b$, not the proportion r/b.

Second, Kahneman and Tversky provided numerous experimental results showing that there is a marked and systematic discrepancy between objective posterior probability and subjective posterior probability. They tested their theory through ten experiments involving decks of cards with various proportions of X and O in them and a 12-card sample drawn at random. The main results are that first, the values of p (the probability for red chips or X-cards) play no role in subjective posterior probabilities, and second, that subjective posterior probabilities depend on sample ratios (X/O) and not—as they should—on the sample difference ($X - O$). The conclusions are unambiguous:

> "*Ss*' estimates violate the Bayesian rule qualitatively as well as quantitatively. Not only are the estimates systematically conservative ..., but they are actually determined by the wrong variable. Consequently, the subjective and objective [posterior probability] estimates are not even monotonically related." (ibid, 448)

This last quotation shows that what was at stake in the whole paper was to show that the representativeness heuristic approach accounts for a richer set of behavioral observations than the conservatism approach championed by Ward Edwards. It can even account for subjective posterior probabilities that violate the correct ordering of likelihoods (Kahneman and Tversky, 1972, 449). This is not to uphold that sample difference plays no role at all in evaluations of likelihoods (both in the sequential and evidence tasks); however, if sample difference is relevant, it will be of secondary importance in the final judgment. Therefore, the experiments conducted on the basis of a hypothesis of judgment of representativeness supported Slovic and Lichtenstein's (1971) earlier criticisms of this normative Bayesian view and encouraged the search for a broader heuristic approach to judgment and decision making.

Spurred on by this victory, Kahneman and Tversky advanced the heuristic approach a step further in two directions. First, they conjectured that representativeness is also relevant when considering situations in which no objective parent population exists. Second, they conjectured that another type of heuristic operated in some other circumstances, each time that the construction of a parent population needed to make a judgment is effortful: availability. The following sections address these two issues.

5.1.2 Generalizing the heuristic of representativeness

"On the psychology of prediction" (Kahneman and Tversky, 1973) (hereafter "Prediction") was the last seminal contribution promoting the heuristic of representativeness in this period. It develops arguments for understanding the generality of the heuristic approach and it posits some reflections regarding other heuristics and how they can coexist with each other in explaining judgments and choices. In this section, we focus on the representiveness heuristic, which was developed along the lines of the conjecture advanced in 1972.

> "We conjecture that the same [representativeness] heuristic plays an important role in the evaluation of uncertainty in essentially unique situations where no 'correct' answer is available. The likelihood that a

particular 12-year old boy will become a scientist, for example, may be evaluated by the degree to which the role of a scientist is representative of our image of the boy. Similarly, in thinking about the chances that a company will go out of business, or that a politician will be elected for office, we have in mind a model of the company, or of the political situation, and we evaluate as most likely those outcomes which best represent the essential features of the corresponding corresponding model." (Kahneman and Tversky, 1972, 451)[69]

"Prediction" widened the scope of the argument to new classes of tasks and informational contexts. It deals both with "judgments" and "predictions" when some stable objective data in the context allow to compare individuals' responses with expected responses along the lines of "the normative principles of statistical prediction." (Kahneman and Tversky, 1973, 237). In contrast with previous experiments, these objective data are not given by the experimenter or by some natural element in the description of the task. They are rather homogenous intersubjective representations of a prior distribution that each *S* shares to a reasonable degree with other *Ss*, a distribution that is expected to work as a stable prior distribution in the construction of expectations on the part of each *S*.

The aim of Tversky and Kahneman was to understand how representativeness might influence the way *Ss* make predictions. Two different classes of predictions were studied, which differ by the mode of response. In the case of a *category* prediction, the mode of response is nominal, it consisted mainly in ordering categories of possible outcomes according to their likelihood (e.g. being a doctor, an engineer, a lawyer). In the case of *numerical* predictions, *Ss* are asked to provide a numerical response (e.g. an estimation of the future value of a student grade). In both cases, the goal of the experiments is to test whether *Ss* account for their prior objective distribution of categories or if they rely predominantly on poorly specific information about the individual instance to be categorized. For categorical predictions, what is at stake is the tendency to ignore objective base-rates. In the case of numerical predictions, the subject of study is the influence of uncertainty in the description of the variability of the predictive assignments. In both cases, the heuristic of representativeness provides an explanatory scheme to account for systematic deviations from normative principles. In the following, I concentrate on a subset of experiments in Kahneman and Tversky (1973) and I discuss their attempt to draw general conclusions regarding individuals' abilities to integrate regression calculus in everyday life.

The behavioral economics literature has largely publicized the experimental results that *Ss* tend to ignore base-rate information when making a judgment, giv-

[69]There is an ambiguous statement in this quotation. In the case of the boy and the politician, Kahneman and Tversky call for a model of the parent population: the population of scientists and the characteristics attributed to them *qua* scientist; the set of characteristics of the political situation (opinions in the population of voters on various subjects, economic indicators, etc). In the case of the company, they call for a model of the company and not, as expected, of the parent population that consists of the set of companies operating in the same business and all known financial and non financial information on these companies. We can postpone the discussion of this case and insist that, in general, representativeness calls for a representation of the parent population *and* of the unique element with which it is compared

ing disproportionate importance to poorly specific information and stereotypes. Consider the following experimental setup:

- A group of *Ss* (**the *base-rate* group**) is first asked to indicate their best guess about the distribution of all first-year students in the US among nine disciplines (Medicine, Engineering, Humanities and Education, Law, . . .). This establishes an objective base-rate for statistical expectations.

- A second group (**the *similarity* group**) is then asked to read a personality description about a young man named Tom, his strong and weak points, his hobbies, etc.[70] *Ss* in this second group are then asked to rank the nine disciplinary categories according to how Tom's description fits well with their views about the "typical graduate student" to be found in each category. The average score for each category is then computed (giving 1 to the lowest ranking and 9 to the highest ranking).

- A third group (**the *prediction* group**) is given the personality sketch of Tom and warned that it was written by a psychologist when he was in high school; Tom is now a graduate student. *Ss* in this third group are then asked to rank the nine fields according to the likelihood that Tom is graduating in each of these fields. The average score for each category is then computed (again attributing 1 to the lowest ranking and 9 to the highest ranking). It is thus possible to compare the three sets of values: the objective mean base-rate (expressed in % of the total population of students), the similarity judgment and the likelihood judgment (Table 1).

Statistical analysis of the correlation between values in columns indicates a strong correlation between similarity and likelihood judgments, while instead there is no correlation with base-rate. For instance, in their experiment, the prior odds for humanities against computer sciences is 3:1 (based on subjective distribution of students among disciplines) whereas 95% of *Ss* in the prediction group judged that Tom is more likely to study computer science than humanities.

> "Judgments of likelihood essentially coincide with judgments of similarity and are quite unlike the estimates of base rates. The result provides a direct confirmation that people predict by representativeness, or similarity. . . . [They] drastically violate the normative rules of prediction." (Kahneman and Tversky, 1973, 239).

The assessment that such a behavior manifestly contradicts normative rules of statistical thinking needs further arguments. "Prediction"compares, within an experimental setting, a current state of beliefs regarding a whole population (the objective prior) with a set of descriptive information about a single individual case in the

[70]The description is as follows. "Tom W. is of high intelligence, although lacking in true creativity. He has a need for order and clarity, and for neat and tidy systems in which every detail finds its appropriate place. His writing is rather dull and mechanical, occasionally enlivened by somewhat corny puns and by flashes of imagination of the sci-fi type. He has a strong drive for competence. He seems to have little feel and little sympathy for other people and does not enjoy interacting with others. Self-centered, he nonetheless has a deep moral sense." (ibid, 238)

Graduate specialization area	Mean judged base rate (%)	Mean similarity rank	Mean likelihood rank
Business Administration	15	3.9	4.3
Computer Science	7	2.1	2.5
Engineering	9	2.9	2.6
Humanities and Education	20	7.2	7.6
Law	9	5.9	5.2
Library Science	3	4.2	4.7
Medicine	8	5.9	5.8
Physical and Life Sciences	12	4.5	4.3
Social Science and Social Work	17	8.2	8.0

Table 1: Estimated base rates for nine areas of graduate specialization and summary of similarity and prediction data for Tom W. (Kahneman and Tversky, 1973, 238)

population. So, the issue is: To what extent can the latter information overshadow the former or, at least, be taken as relevant against the later? Bayes rule expresses the idea that singular information about individual instances can outweigh prior odds if they are judged to be "both accurate and diagnostic" (ibid, 239). Typically, this was not the case, as indicated by post-experimental interviews with *Ss*, who did not think that the information about Tom's future was established on a reliable scientific basis.

From a normative standpoint, Bayesian inference states that statistical prediction relies on three kinds of information and beliefs: i) prior information or base-rates, ii) specific information on the individual case, iii) degree of belief in the relevance of specific information, its effectiveness, or in Kahneman and Tversky's terminology "the expected accuracy of prediction" (ibid, 239). According to Bayes' rule, iii) is fundamental to figuring out the importance of ii) compared with i).

> "A fundamental rule of statistical prediction is that expected accuracy controls the relative weights assigned to specific evidence and to prior information. When expected accuracy decreases, predictions should become more regressive, that is, closer to the expectations based on a priori information." (ibid, 239).

In the present case, because *Ss* judged and made prediction by representativeness, they gave prominent weight to specific evidence against prior probabilities, as they did in judging similarity. The two rankings are quite similar, revealing that *Ss* ignored several elements that should have been in favor of giving more weight to prior base-rate: Tom's description is several years old and probably invalid (hence reducing the accuracy of the information and increasing the variability of the true characteristics), and "even if the description is still valid, there are probably more people who fit that description among students of humanities and education than among students of computer science, simply because there are so many more students in the former than in the latter field." (ibid, 240) In other words, this experiment allows to identify a bias of judgment due to the use of a representativeness heuristic when specific information is inaccurate (uncertainty).[71]

Contrary to that, when the description of a would-be first-year student is not specific at all, i.e. contains no information about his personality, hobbies, etc., then *Ss*' ranking of the likelihoods of being a student in a discipline is correlated with their base-rate estimate. In light of previous articles, everything happens as if the lack of usable information prevents the activation of a representativeness heuristic. It is an indication that people tend to become overconfident about the possibility to classify based on little information. There is a leap from using only base-rate to using mainly individual description.

> "In the statistical theory, one is allowed to ignore the base rate only when one expects to be infallible. In all other cases, an appropriate compromise must be found between the ordering suggested by the description and the ordering of the base rates." (ibid, 241)

As a litmus test for the compelling action of representativeness against base rate information, Kahneman and Tversky presented the now well-known lawyers-engineers experiment, in which the prior distribution is perfectly controlled.

- **Description**. A description of an individual displayed on a card is given to two groups of *Ss*. Each *S* receives the same description.[72] Each *S* is asked to give his best guess about the profession of the person described : lawyer or engineer.
- The **first group** is told that descriptions are drawn at random from a set of cards containing 70 lawyers and 30 engineers.
- The **second group** is told that there are 70 engineers and 30 lawyers.
- The task is replicated with five such descriptions.
- **Responses**. The average probability estimate is 50 % for the first group and 55 % for the second group.
- **Conclusions**. The base-rate influences poorly the responses: "Explicit manipulation of the prior distribution had a minimal effect on subjective probability". This confirms that representativeness is a disruptive cognitive process, i.e. that it does not shift our judgments by degree but in a discrete way.

[71] Kahneman and Tversky used an additional experiment to control for the effect of a variation in accuracy. One group of *Ss* were told that from past experiment, it is known that *Ss* like them are good predictors. Another group was told that from past experiment, it is known that they are poor predictors. This information had no effect whatsoever on the ordering of likelihoods, even though it affected *Ss* views about the accuracy of their own personal judgments. "This pattern of judgments violates the normative theory of prediction, according to which any decrease in expected accuracy should be accompanied by a shift of predictions toward the base rate." (ibid, 240)

[72] An example of description is as follows: "Jack is a 45-year old man. He is married and has four children. He is generally conservative, careful and ambitious. He shows no interest in political and social issues and spends most of his time on his many hobbies which include home carpentry, sailing, and mathematical puzzles."

> "When no specific evidence is given, the prior probabilities are properly utilized; when worthless specific evidence is given, prior probabilities are ignored. . . .[This is] perhaps one of the most significant departure of intuition from normative theory of prediction." (ibid, 242-243)

Experiments on numerical predictions aim to replicate the results obtained with category predictions. More precisely, in this case, Kahneman and Tversky's goal was to control for the variation in uncertainty that a subject should feel in performing two different tasks. The first task is to judge the value of an information relatively to some outcome or performance. The second task is to make a prediction of outcome or performance based on the same information.

- **Two groups of *Ss*** are given descriptions of first-year college students (for instance, a list of adjectives referring to intellectual qualities and character).
- For each description, *Ss* in the **evaluation group** are asked to answer the question: "Estimate the percentage of students in the entire class whose descriptions indicate a higher academic ability."
- For each description, *Ss* in the **prediction group** are asked to answer the question: "What is your estimate of the grade point average that this student will obtain?"

Since one should take into account the greater uncertainty in the predictive task, the normative model says that the prediction should be more regressive than the evaluation, that is, it should be closer to 50%. Instead, according to the heuristic of representativeness, *Ss* would be poorly sensitive to the difference between the two tasks. Once information is accurate enough to trigger a judgment by representativeness, it is accurate for any kind of task. Experiments constructed with descriptions of students confirmed that *Ss* who evaluate and *Ss* who predict produce equally extreme judgments.[73]

Now, given the apparent pervasiveness of representativeness heuristic, the next question is about understanding how an *S* makes sense of an array of various inputs and derives from them a sense of representativeness. In the previous experiments, cues or stereotypes were controlled by the experimenter and were deliberately coherent, not contradictory or difficult to match together. This issue of incoherence and conflict within a multiattribute input was shown to be central in "Intransitivity of preference" (Tversky, 1969). At the end of "Prediction", Kahneman and Tversky advanced some thoughts about the way we may tend to filter or neglect information at our disposal in order to enhance the confidence in our judgments. They connected this feature with Slovic's (1966) findings that internal inconsistency in the input decreases the confidence in predictions. For instance, one feels more confident in predicting a B as a next grade if previous grades were (B,B,B,B) rather than

[73]Kahneman and Tversky also compare the results of their experiments with results obtained in Bayesian learning experiments, where psychologists study the ability to change probability estimations as new information is displayed as feedback. In this case of learning, *Ss* response sequences that are also representative of the associations between inputs and outcomes.

(A,C,A,C). Definitely, this result is another reason for rejecting the usual multivariate (linear) models of prediction in which it is usually assumed that "expected predictive accuracy is independent of within-profile variability" (Kahneman and Tversky, 1973, 65).

One intuitive way of thinking about this essential matter is to consider that two pieces of highly correlated information (or inputs) together provide a more accurate information to infer a prediction than would two pieces of information with a lower correlation. Contrary to this intuition, if each information is equally good to infer outcome, then a higher predictive accuracy will be obtained by using the two uncorrelated pieces of information than by using the two correlated ones. As expected, experiments show that people feel more confident when they elaborate predictions from two correlated inputs. When coping with a wealth of inputs, this fact may well be an incentive for people to make predictions for which they feel a high degree of confidence and thus to sort out information and retain as relevant inputs those that are highly correlated because they belong to a stereotype. Consistency within inputs increases confidence in predictions, although it should not since it is obtained at the cost of ignoring some piece of relevant information. The main conclusions of "Prediction" stand an an attempt to establish the generality of the heuristics of representativeness. It was soon complemented by another heuristic : availability.

5.2 The heuristic of availability

The heuristic of availability was proposed in the wake of the heuristic of representativeness, in order to account for a different cognitive process of judgment and decision that *Ss* would tend to use in some specific contexts, when representativeness is not an option or in coordination with representativeness.

What are the features of such a context? As we have seen, issues about representativeness revolve around the use and selection of relevant information, hence it applies in situations where information is in excess and cannot be easily handled in totality. In the opposite case, when the initial quantity of relevant information at the disposal of *Ss* is limited or must be remembered or built in some way, representativeness is irrelevant, it cannot operate. In this case, availability appears as the relevant heuristic to deal with these situations. Also, when there is no objective knowledge about categories shared by *Ss* that can stand up against the task to be performed, availability assumes that *Ss* provide for missing information, build scenarios, in order to arrive at a judgment. Thus, the main hypothesis is that availability is predominantly at work in situations that are perceived as information demanding. In these cases, *Ss* have to build this information, and the ease with which this task is performed should determine the probability attributed to an event. As they did for representativeness, Tversky and Kahneman posited the availability heuristic from simple experiments, in which the difficulty to retrieve relevant information can be controlled. In a second stage, they investigated the possibility to extend the analysis of availability to real life situations.

5.2.1 Basics of Availability

The seminal contribution to availability is Tversky and Kahneman's article "Availability: A Heuristic for Judging Frequency and Probability" (Tversky and Kahneman, 1973) (hereafter "Availability", which appeared first as a technical report of the *Oregon Research Institute* in 1972. It is also announced in "Representativeness" (Kahneman and Tversky, 1972, 451)[74]

Kahneman and Tversky indicated from the start that the availability heuristic was an alternative to the representativeness heuristic, opposing two processes of judgment making. It deals with situations where instances of a class of events are recalled more or less easily.

> "Life-long experience has taught us that instances of large classes are recalled better and faster than instances of less frequent classes, that likely occurrences are easier to imagine than unlikely ones, and that associative connections are strengthened when two events frequently co-occur. Thus, a person could estimate the numerosity of a class, the likelihood of an event, or the frequency of co-occurrences by assessing the ease with which the relevant mental operation of retrieval, construction, or association can be carried out. ... A person is said to employ the availability heuristic whenever he estimates frequency or probability by the ease with which instances or associations could be brought to mind." (Kahneman and Tversky, 1973, 208)

"Availability" identifies different "classes" of availability heuristics, depending on the kind of task and information presented to *Ss*. The main classes studied in the article are: availability for construction under a given rule (section III), availability for retrieval (section IV), and availability for construction of scenarios (section V). Within each class, complementary distinctions are considered according to the kind of task performed (identifying frequency of repetitions or frequency of co-occurrences). The heuristic of availability contains also a dimension of psychological cost necessary to build the set of information that will be used to make a judgment. Hence, two features of availability have to be taken into account in the final outcome: one has to do with logical inference; the other one has to do with economizing cost of computation.

As regards logical inference, availability implies some kind of internal process in the mind, a computation of instances and then a comparison of two outcomes in order to assess probability or frequency. It goes in reverse order to the usual view that we tend to strengthen the association between two terms when we record a high frequency of co-occurrences. Indeed, when we judge by availability, the fact that association between two terms is easier to retrieve is taken as an indication of the frequency of co-occurrence of these terms.

---

[74]Like "Representativeness", "Availability" was published in *Cognitive Psychology*, a journal launched in 1970. Its Chief Editor was Walter Reitman, and among the editorial board members, one can mention Noam Chomsky, Donald Davidson and Herbert A. Simon. Most of the articles published in the early 1970s dealt with issues of memory, retrieval, and learning.

As regards computation, when judge by availability, we can quickly perform some partial operations that, we think, will be enough to make a satisfactory judgment or estimation of some outcome by extrapolation.

The classic experiment illustrating the logical inference aspect of availability is the experiment on judgments of word frequency.

- **Description**: In substance, *Ss* are given the following instruction: "A typical text in the English language has been selected and the relative frequency with which a letter appeared in the first and third positions in words was recorded (excluding words with less than three letters). You are asked to judge for several letters whether they appear more often in the first or in the third position, and to estimate the ratio of the frequency with which they appear in this position."
- Consider letter R, answer the questions.
  - I consider that R is more likely to appear in the first/the third position? (circle your choice)
  - My estimate for the ratio of the first position/third position is ( )
- **Results**: the great majority of *Ss* (about 70%) judge the first position to be more likely and the median estimate is that there are twice as much R-first words as R-third words. Actually, there are more R-third words. (The experiment gives the same results with usual letters K, L, N, V.)

The computation problem is first illustrated by the result of fast computation of the numerical product $1 \times 2 \times 3 \ldots \times 9$ or $9 \times 8 \ldots \times 2 \times 1$. Later, this experiment will be associated with the heuristic of anchoring and adjustment (see below). However, the main idea behind it is that in the case when the task cannot be performed using the raw information at the disposal of the *Ss*, they will make an intermediate computation, hence they will produce in a certain way the available data rather than just retrieving it through immediate memory. These data in turn can be used to make a judgment.

Of course, availability can lead to biases, as in the two previous examples, according to the structure of the problem to be solved. Also, the computational aspects and the logical inference aspects may be involved simultaneously, as in the next problem (let us call it the “Bus-Stations” problem).

- **Description**: “In the drawing below, there are ten stations along a route between START and FINISH. Consider a bus that travels, stopping at exactly r stations along this route. START _ _ _ _ _ _ _ _ _ _FINISH
- What is the number of different patterns of r stops that the bus can make? Answer this question for $n = 2, \ldots 8$.”
- **Results**: The exact answer is given by the binomial coefficient $\binom{10}{r}$.[75] Its true values start at 45, go to a maximum of 252 when r =5 then decrease up to 45.

[75]The number of arrangements of r terms among 10 possibilities without repetition—not two times the same station on the way ($\frac{10!}{(10-r)!}$). This value has to be divided by the number of different orderings of r terms ($r!$), which makes $\binom{10}{r}$.

> The average responses start at about 50 (slight overestimation) and decrease continuously up to about 18, far below the right answer.

As in other similar problems, “there is marked underestimation of all correct values, with a single exception in the most available case where $r = 2$” (Tversky and Kahneman, 1973, 215). In this case, it is reasonable to imagine that both the difficulty to compute and the difficulty to represent (or enumerate) various cases (especially with intermediate values of r) explain for the gap between stated values and true values.

Overall, for each experiment, Kahneman and Tversky engage with a possible reconstruction of *Ss*’ way of thinking in order to explain the gap between estimated values and exact values. When people extrapolate from partial computations or when they extrapolate from a feeling that it is easier to build instances of two-stops itineraries than eight-stop itineraries in the bus-stations problem, the same kind of easiness for obtaining instances serves as the criterion of judgment.

Hence, what is suggested is that easiness is the criterion for judging, and that this criterion applies to various kinds of information processing (calculations, retrieval). A series of experiments carried out by Tversky and Kahneman aim precisely at confirming this view. These experiments focus on the ability to control for the easiness to memorize certain occurrences or associations of occurrences. When *Ss* have to solve a problem based on a given list of items that has been shown or read to them, they will judge about the frequency of some classes of items or some association of items according to the easiness with which they can retrieve the items of the list, and this easiness can be linked with various controllable aspects of the items (e.g. the name of a celebrity *vs* the name of a less well-known person, a long word *vs* a short word, a natural association of words (eggs-bacon) *vs* an unusual one (lion-eggs). In all cases, availability as easiness-to-retrieve-instances will serve as a proxy for judging frequency of occurrence. For instance, eggs-bacon associations (and similar) will be estimated as more frequent than lion-eggs (and similar) even though they were equally frequent in the list.

5.2.2 Generalizing availability: illusionary correlation and scenarios

The important step in the theory of availability concerns the possibility to extend it to experimental cases that go beyond pure mnemonic exercises, and possibly to extend it to ordinary situations of judgment. The transition from pure recall or pure construction problems to expert judgment and real-life situations involving availability is organized around the notion of illusory correlation.

The concept of illusionary correlation had been introduced and discussed by Loren J. Chapman and Jean P. Chapman (Chapman (1967); Chapman and Chapman (1967) and (1969)). Illusory correlation is defined as the fact for an observer to report a correlation between two classes of events which in reality are not correlated (or are correlated in the opposite direction from the one reported).[76] Chapman (1967)

[76]Illusory correlation applies to various classes of stimuli, from magical beliefs to prejudices. The main contribution of Chapman is to suggest that "some systematic principles are operating to distort the observational report ... regardless of the differences in the subject matter being observed" (Chapman, 1967, 152)

identified that the overestimation of the number of occurrences of a pair of words takes place when words are related and then more easily recallable, as is the case with "lion-tiger" or "bacon-eggs" but not with "lion-egg" or "bacon-tiger". This overestimation is persistent even though *Ss* are warned before reading the list of paired words that they will have to indicate how many times each word is paired with each other. This illusory correlation between related words was explored further in experiments with naïve clinical judges who had been given data of test material (like draw-a-person and Rorschach tests) and diagnoses. Naïve judges tend to overestimate significantly the occurrence of stereotypical associations of symptoms and diagnoses. The illusory correlation effect appeared resistant to contradiction and prevented judges from detecting real correlations.

To Kahneman and Tversky, "availability provides a natural explanation for illusory correlation" (223). They describe a sort of self-reinforcing mechanism of strengthening of an actual association of two items and of perception of frequency of this association. This occurs either when some structural property of the problem facilitates some "natural" associations, or when the association reflects some prior association in people's minds. For instance, when pairs of words consist of long-short or long-long words, long-long pairs will "naturally" be remembered more easily . And we can assume that a "lion-tiger" pair is remembered more easily because the two animals have some attributes in common, such as wilderness or being a predator. So, the phenomenon of illusory correlation opens to the possibility of extending the availability heuristic to cases beyond closed systems of data, i.e. to cases involving references to imagination, personal memory, or social stereotypes. The stakes are important enough that Kahneman and Tversky undertook to replicate Chapman's original study.

- **Description**: In a pair of words, the first is called stimulus, the second is the response word. Consider five pairs of highly semantically-related words (like knife-fork or lion-tiger) and five pairs of highly phonetically-related words (like cake-fake, blade-blame), and ten unrelated pairs obtained after replacing all the ten stimuli by another list of (unrelated) words. Each response word is thus paired with one related word and one unrelated word. A message containing repetitions of the pairs in a randomized order is played with two groups of *Ss* (pairs with the same response have the same number of occurrences). The experimenter then measures the probability of recall (group 1) and the judged frequency (group 2).

- **Results**: *Ss* recall significantly more highly-related pairs (group 1) and judge them to be more frequent (group 2). This replicates Chapman's findings and corroborates the identification of an illusory correlation.

The next step is to provide an explanatory element for illusory correlation. If, as is hypothesized, "the illusory correlation effect is due to differences among item pairs in the strength of the associative bond between their members ... [then] the same effect should also occur in a noncorrelational design, where each response is paired with a single stimulus, and vice versa." (ibid, 226). In other words, Tversky and Kahneman conjecture that if a new experiment is designed such that *Ss* cannot rely on the comparison between pairs having the same stimulus, then the structure of the information is changed. With this new design, the origin of the associative bonds between words becomes an inner property of the single pair (not a comparative property between pairs) and it can be expected that judgment of frequency will mirror "the degree to which the response word is made available by the stimulus word." (ibid, 226). The new experimental design is as follows.

- **Description**. Consider eight highly-related pairs and eight unrelated pairs of personality traits (kind-honest, selfish-greedy, .. *vs* nervous-gentle, eager-careful, ...). A list of all pairs, with various levels of repetition, is recorded and all *Ss* listened to the record. They were then dispatched in three groups ($62 < N < 73$). The first group had to recall the response associated with each stimulus. The second group had to assess their probability of recalling the response from stimulus. The third group had to indicate how often each pair appeared.
- **Results**. Experimental results corroborate an illusory correlation effect. More precisely, the "recall" and "assess recall" tasks confirm that highly-related pairs are significantly more remembered than unrelated pairs and that *Ss* are aware of their ability. As regards the perceived number of occurrences, "although judgments of frequency were generally accurate, a slight but highly systematic bias favoring related pairs was present" (ibid, 227)[77]

After this specific experiment devised to foster the idea that illusory correlation can be generated *endogenously* by a single instance of an associative bond, Kahneman and Tversky aim to extend the availability heuristic to real life cases, that is to cases when there is no objective correct answer because not all instances can be enumerated. This attempt at extending the prevalence of availability is part of the heuristics and biases approach and echoes similar considerations for the representativeness heuristic.

> "Each occurrence of an economic recession, a successful medical operation, or a divorce, is essentially unique, and its probability cannot be evaluated by a simple tally of instances. Nevertheless, the availability heuristic may be applied to evaluate the likelihood of such events" (ibid, 228)

Kahneman and Tversky's examples of real-life cases where the availability heuristic is at work are somewhat disordered and pertain to different kinds of rationalization and generalizations. Broadly, they advance two mental processes. When we are to judge about the probability or likelihood that a certain situation or instance will be associated with a certain outcome, our minds are prone to provide us with a set of similar instances with which a specific event will be compared (first process), and this process may be somehow biased (path dependency). If this set of similar cases cannot be obtained, then people tend to rely on stories (called "scenarios")

[77] Note that the design of the experiment does not seem to allow for overestimation of frequency. Thus, the reinforcing effect between association and repetition may have been tempered.

somehow plausibly connected with the unique instance they want to assess (second process).

Let us first examine situations where someone has to recall similar cases. Consider a person who tries to judge the likelihood that a particular couple will divorce (by comparing with similar couples that come to this person's mind), or a clinician who tries to assess the likelihood that one of his patients commit suicide (who tries to recall similar patients). In both cases, judges are assumed to be able to build a set of similar cases against which they will compare the current instance. The question then is "How are relevant instances [of past couples, of past patients, of past recessions] selected?" (ibid, 228). The way a clinician will construct the set of similar cases to judge the probability that a depressed patient may commit suicide can be distorted by the fact that mostly salient cases of suicidal patients come to her mind more easily than cases of depressed patients. In this case, the sample of cases against which the present instance is compared is not relevant. There may be a tendency to confuse Pr(attempting suicide | being depressed) and Pr(being depressed | attempting suicide) or to focus on depression/suicide co-occurrences to the detriment of depression/no-suicide occurrences.

The immediate problem with the examples discussed by Kahneman and Tversky is that they mix aspects of representativeness (judgments of similarity between the instance to be judged and past instances)[78] and aspects of availability, i.e. identification of the principles that have been operating to build the list of similar or relevant cases with which the current case will be compared. It is thus clear at once that availability in this case is not the only heuristic at play in making a judgment, and that more often than not, it is sort of a preliminary or intermediate heuristic that will provide the representativeness heuristic with the material it needs to operate. What can be the consequences on the quality of the representativeness judgment? Answering this question becomes a crucial point whose discussion must be postponed to the last section of the present study.

Consider now the extreme case where a situation and a likely event or outcome that can proceed from this situation are judged to be so unique that no natural set of comparable elements to make a judgment comes to mind.

> "In thinking of such events we often construct scenarios, i.e., stories that lead from the present situation to the target event. The plausibility of the scenarios that come to mind, or the difficulty of producing them, then serve as a clue to the likelihood of the event. If no reasonable scenario comes to mind, the event is deemed impossible or highly unlikely. If many scenarios come to mind, or if the one scenario that is constructed is particularly compelling, the event in question appears probable." (Tversky and Kahneman, 1973, 229)

In building scenarios, the mind reduces the complexities of interdependencies or uses *ceteris paribus* reasoning and attributes higher value to intuitive outcomes. Availability of scenarios may depend on context. In situations of conflict, one cannot

[78] "When several instances come to mind, they are probably weighted by the degree to which they are similar, in essential features, to the problem at hand." (ibid, 228)

easily consider the vantage point of the opponent and one tends to downplay the evolution of his thinking. Scenarios may be dependent also on the most immediate experience about a class of phenomena, like “a temporary rise in the subjective probability of an accident after seeing a car overturned by the side of the road." (ibid, 230)

At the same time Kahneman and Tversky uphold that compelling scenarios, once adopted, tend to organize the rest of our views and are not easily modified or abandoned. This echoes views already put forward (Tversky and Kahneman, 1971) regarding the difficulty in discarding a scientific hypothesis once a statistical test has suggested some significant correlation,[79] and it pertains to ideas about a third heuristic: anchoring.

The last comments on the effects of past experience on judgment and decisions in real life revolve around the idea that past experience about a phenomenon sets a bound on the representations we are ready to accept when we have to judge about the likelihood of similar phenomena. Someone who has been told about the risk of huge floods or droughts in the area where he lives may only think of the floods or droughts he has experienced in the past. However, Kahneman and Tversky advance that we are also likely to be influenced by recent fictional stories or news about a type of extreme phenomenon in appraising its likelihood. This brings us to the heuristic of anchoring.

5.3 Anchoring and adjustment

The heuristic of anchoring and adjustment has a special place in the trio of heuristics proposed in the early 1970s. As we have seen, the idea that *Ss* may be led to overvalue certain information when processing information had been conjectured by Slovic and Tversky in the late 1960s and it was central to Lichtenstein and Slovic’s argument about preference reversal. However, it became strictly speaking a heuristic in an research paper by Slovic (1972) "From Shakespeare to Simon", in which Slovic offered the first attempt to figure out the theoretical coherence behind the contributions of Tversky, Kahneman, Lichtenstein and himself.[80] Anchoring was

[79] Tversky and Kahneman refer to Bruner and Potter (1964, 425), who had identified that when people are exposed to a blurred image which comes into focus slowly tend to build hypotheses about the identity of the object and take more time to recognize the object as compared with *Ss* exposed directly to a clear picture. “hypotheses about the identity of the picture are made despite the blur. The ambiguity of the stimulus is such that no obvious contradiction appears for a time, and the initial interpretation is maintained, even when the subject is doubtful of its correctness.”

[80] The title of this article was inspired by the opening opposition between the bounded rationality assumption upheld by Simon and Prince Hamlet’s views (*The Tragedy of Hamlet, Prince of Denmark*, Act II, scene 2) of the high faculties of human beings: "What a piece of work is man, how noble in reason, how infinite in faculties, in form and moving, how express and admirable; in action how like an angel, in apprehension, how like a god: the beauty of the world, the paragon of animals." (from the 1623 edition, *The Complete Works of William Shakespeare*). Note that at the time of writing "From Shakespeare to Simon" (published in the *Oregon Research Institute Research Bulletin*, vol 12 n°2), Slovic knew about all the contributions of Kahneman and Tversky on availability and representativeness that were to be published in 1972—1974. As regards the heuristic of anchoring, a paper by co-authored by Tversky and Kahneman is cited, titled "Anchoring and calibration in the assessment of uncertain quantities", and was published in the *ORI Research Bulletin* vol 12, n°5 (1972); it was later published as part of "Judgment under uncertainty: heuristics and

then discussed in greater detail as a heuristic proper in Tversky and Kahneman's "Judgment under uncertainty: heuristics and biases" (Tversky and Kahneman, 1974) and also in Tversky (1974), the published version of a lecture Tversky delivered at the Royal Statistical Society (February 13th 1974). Anchoring was also mentioned, incidentally, in a study on insurance for natural hazards by Slovic, Kunreuther and White (1974), where they noted that one possible effect of anchoring is to make our expectations about the frequency of rare events too small.

In the following, I shall concentrate on the presentation of anchoring in Tversky and Kahneman (1974), and I shall make occasional remarks based on the other contributions.

The driving idea about anchoring is that on many occasions, the way we process information to reach a judgment or a decision leads us to focus on some specific parts of the information at our disposal in order to make a first evaluation of the final outcome. Once computed or identified, this first evaluation is bound to exert too much weight on our final judgment.

> "In many situations, people make estimates by starting from an initial value that is adjusted to yield the final answer. The initial value, or starting point, may be suggested by the formulation of the problem, or it may be the result of a partial computation. In either case, adjustments are typically insufficient. ... We call this phenomenon anchoring." (Tversky and Kahneman, 1974, 1128)

As a heuristic, anchoring indicates that a certain framework of information combined with a certain task (implying a certain response mode) will generate a sequential processing of information. When compared with another task, or another response mode, or another initial framework of information, it may be possible to identify that some anchoring has generated a systematic bias as compared with another anchoring. The ability to identify a systematic effect of one setup against another is essential to identify a bias linked to anchoring. On this point, it is notable that Kahneman and Tversky connect this heuristic to Slovic and Lichtenstein's (1971) analysis of preference reversal. A pure effect of anchoring attributed to a biased arbitrary starting point is presented by the following experiment.

- **Description**. Groups of *Ss* have to estimate the percentage of African countries represented at the United Nations. Before starting to think about this value, a number between 0 and 100 is drawn at random (one for each group).
- *Ss* have to answer the question "Is that number higher or lower than the exact value of the percentage of African countries at the UN?"
- Then, the second question is "What is your estimate of the percentage of African countries at the UN?"
- **Results**. The starting values drawn at random have a "marked effect on estimates". The median estimate for the group that had drawn number 10 was 25, while the median estimate associated with number 65 was 45.

biases".

Another anchoring effect, associated with a prior partial computation of values, is illustrated by the fast (5 seconds) numerical estimation of two equivalent algebraic products.

- **Description**. Two groups of *Ss* have to perform a mental calculation.
- **Group 1** of *Ss* is asked to estimate $8 \times 7 \ldots 2 \times 1$
- **Group 2** is asked to estimate $1 \times 2 \times 3 \ldots \times 8 \times 9$
- **Results**. The median answer for the first group is 2,250 and the median answer for the second group is 512. The true value is 40,320. The cost of calculating implies that *Ss* calculate only the first steps and approximate based on the speed of increase. There is wide underestimation in both cases, even though one would be tempted to assess that one bias is weaker than the other.

Interestingly enough, this example of a partial computation that works as an anchor for further calculations was first presented as an instance of the heuristic of availability in Tversky and Kahneman (1973, 215-216), showing that the anchoring heuristic had to gain ground to be analysed as a specific heuristic apart from availability. After a few sets of computations, *Ss* extrapolate to give an estimate, but their adjustments "are typically insufficient" (ibid, 185). Indeed, while the result of the partial calculation is an anchor since it determines the value obtained through extrapolation, one may also consider that the first steps of computation provide some relevant intuitive indicator to *Ss*, and as such are made available through calculation based on other information.

Next, Tversky and Kahneman present various biases associated with anchoring and adjustment. They start with a bias in the evaluation of conjunctive and disjunctive events borrowed from experiments carried out by Bar-Hillel (1973).[81]

- **Description**. Consider three kinds of events.
  - (i) **Simple event**. Drawing a red marble from a bag containing 50% red and 50% white marbles (p=0.5),
  - (ii) **Conjunctive event**; Drawing a red marble seven times in succession, with replacement, from a bag containing 90% red marbles and 10% white marbles ($p \simeq 0.48$);
  - (iii) **Disjunctive event**. Drawing a red marble at least once in seven successive tries, with replacement, from a bag containing 10% red marbles and 90% white marbles ($p \simeq 0.52$).[82]

[81]Maya Bar-Hillel's experiments (1973) were fostered by the program set out by Slovic, Lichtenstein, Tversky and Kahneman on "systematic biases in intuitive judgments of probability" (396). They consist in a systematic analysis of *Ss*' preferences when facing a choice between a simple gamble (one single draw in an urn) and a compound gamble (drawing n times in an urn with a known proportion of colored marbles.)

[82]The results presented in this description are purposefully selected as the most noteworthy instance of a bias obtained by Bar-Hillel in the whole set of experiments with different values for opportunities to draw out of an urn and proportions of colored marbles. Indeed, in this case of seven draws out of an urn with 90% of colored marbles, 12 out of 15 *Ss* choose the conjunctive event gamble. And in the complementary case with 10% of colored marbles, 14 out of 20 *Ss* prefer the disjunctive event gamble (see Bar-Hillel (1973) table 1, row c and table 3, row 1).

- **Task**.*Ss* are presented with pairs of such gambles ((i) *vs* (ii) and (i) *vs* (iii)) and asked to choose the one they would prefer to play (assumed to be the one that offers the highest subjective probability of winning).

- **Results**. *Ss* significantly bet on the conjunctive event against the simple event and on the simple event against the disjunctive event. Hence, their ordering of subjective probabilities is (ii) > (i) > (iii), the opposite of the real values. This illustrates a general pattern of experimental results on choice among gambles, a systematic bias: people tend to overestimate the probability of conjunctive events and to underestimate the probability of disjunctive events (Cohen et al. (1972), Bar-Hillel (1973)).

To Tversky and Kahneman, the heuristic of anchoring provides for an explanation for observed biases.

> "The stated probability of the elementary event (success at any one stage) provides a natural starting point for the estimation of the probabilities of both conjunctive and disjunctive events. Since adjustment is typically insufficient, the final estimates remain too close to the probabilities of the elementary events in both cases." (Tversky and Kahneman, 1974, 1129)

In other words, in order to establish that an anchoring and adjustment heuristic is operating, one needs a clear-cut identification of a natural starting point (in this case the probability of a single event) and of the process of adjustment. "The direction of the anchoring bias can sometimes be inferred from the structure of the event." (ibid, 1129)

The anchoring and adjustment heuristic in the case of conjunctive events can be used to think of the tendency to be overoptimistic about the success of an undertaking that must however pass several independent stages, each of which is necessary for success. Conversely, when several events are disjointed, each being an independent risk of failure for the working of a complex system, people tend to underestimate the overall risk of failure.

The next step in establishing the heuristic of anchoring and adjustment seeks to show how two different experimental designs to solve the same problem might lead to two different processes of anchoring and adjustment.

Consider the situations where an expert establishes a subjective distribution of values, a set of values X_i such that i percent of the real values observed fall below X_i. Experiments on subjective probability for such distributions (Winkler (1967), Alpert and Raiffa (1969), Staël von Holstein (1971)) indicate large systematic departures from normal or plausible distributions, underestimating the probability that real values are smaller than X_i for small values of i and high values of i (typically, for X_{99} and X_{01} or X_{90} and X_{10}). The rationale in terms of anchoring is that people start from an expectation about the average value, then adjust towards the extreme values, taking too narrow confidence intervals. To Kahneman and Tversky, the systematic bias that people underestimate the values of the density distribution for small and high values of X_i "is attributable, in part at least, to anchoring" (ibid, 1129). Starting from the average estimation, adjustment to consider extreme values of the distribution are insufficient.

Now, there is another way to ask *Ss* to arrive at an estimation of the probability distribution of a random variable (e.g. the value of the Dow Jones tomorrow), by asking *Ss* to assess the probabilities that true values exceed some specified value. This procedure suggests a different anchor than the best estimate.: "the subject may be anchored on the value stated in the question ...[or on a] 50-50 chances" (ibid, 1129). If this were the case, then this procedure should yield less extreme odds.

This assumption that the procedure suggested can involve different processes of anchoring and adjustment with systematic differences in outcomes (estimation of a distribution) is tested through the replication of an experimental device on 24 different quantities to be evaluated.

- **Description**. Consider *Ss* divided into two groups.
- **Group 1** of *Ss* is asked to assess the value X_{90} and X_{10} for, e.g., the distance between New Delhi and Peking.
- **Group 2** of *Ss* is given the median of Group 1 responses for X_{90} and X_{10} (anchors) and for each value, they have to assess the odds that it exceeds the true value.
- **Responses**. Median odds stated by group 2 are 3:1, significantly different from the 9:1 odds that were estimated by group 1.

None of the procedures used allowed to obtain a subjective probability near the true value. While group 1 underestimated the extreme values, group 2 overestimate probabilities for extreme values. In other words, each procedure triggers a specific way to process information which, given the pattern of information and the task to be performed, leads to systematic biases in a given direction. There is probably no procedure that would be neutral with certainty. However, if one were to provide favorable conditions for a proper calibration of a subjective probability distribution, then, as Tversky (1974, 156) puts it, there is "the intriguing possibility that an appropriate combination of the two methods could yield properly calibrated probability distributions." In this experiment, the anchoring effect is not independent from a response mode effect, i.e. from the kind of response asked, conditional on performing an equivalent task, and none of them can be taken as more reliable in reducing biases.

In sum, anchoring and adjustment identifies a variety of elements in a problem situation that can influence the way an *S* makes a judgment or reaches a decision. These elements can be purely exogenous to the problem (e.g. the pure anchoring effect) or they can pertain to the kind of task or response mode expected from *Ss* or to the way information is displayed. The last case opens onto the idea of a pure framing effect. As it is, anchoring appears to be lurking in every problem-solving or choice environment and to complement the analysis of representativeness and availability.

6 Perspectives opened on the notions of probability, biases and rationality

Based on the material of the previous sections, this section explores the main subjects of comparison between formal theories of decision and choice in economics in the 1970s (EUT and SEUT) and the representation of decision making as it stands out from the set of experiments and thoughts accumulated by Lichtenstein, Slovic, Tversky and Kahneman circa 1965-1974. Even though almost all experiments bear on the analysis of choice under risk or uncertainty, several comments apply *mutatis mutandis* to the analysis of choice under certainty. Our goal is to provide an overview of the opportunities and obstacles to the inception and development of behavioral economics at this juncture, a rational reconstruction of the ways cognitive psychologists could promote a dialogue with economists on the notion of rational decision and the search for normative or prescriptive models of decision making.

The main difficulty in accomplishing this task stems from the fact that the whole set of new knowledge accumulated in the field of cognitive psychology in the 1960s and early 1970s led to to the establishment of complex relationships between the concepts of probability, preferences and rationality, thus challenging the usual axiomatic-normative categories and representations associated with rational choice theory in economics, and beyond, the very notion of optimality. As we shall see, the claim of our protagonists is a rather subtle view on the conditions under which it is possible to assess individuals' departure from a rational choice based on well-identified preferences. In order to grasp the variety of attitudes towards rational choice theory and the possibility to design systems to improve decision making, I shall first delve into the notion of subjective probability. Then, I shall clarify the distinction between heuristics and biases. Last, I shall focus on the reshaping of the concepts of rationality and optimality.

6.1 Uncertainty and subjective probabilities

As is well-known, the meaning of probability has been a hot issue in economics at least since Knight, Keynes and Ramsey's treatments of it, and the development of the standard modeling of choice under risk and uncertainty is firmly grounded on De Finetti's and Savage's views. Quarrels about interpretations of the descriptive or normative contents of these theories are still lively (Brady and Arthmar, 2016; Feduzi et al., 2017). It may be interesting to examine whether promoters of the heuristics and biases approach adopt a specific view about subjective probability. In this subsection, I pinpoint some statements about probabilities scattered through the early literature on heuristics and biases.

A common characteristic of the research done by Lichtenstein, Slovic, Tversky and Kahneman during this period is worth mentioning here. They definitely favor a non-Bayesian, static approach to the study of judgment and decision. This is, in my view, a methodological tenet that has to do with the very possibility of highlighting the disruptive consequences of cognitive psychology on the modeling of choice (Slovic and Lichtenstein, 1968) For instance, Lichtenstein's study with three-outcome bets allows to reject Edwards' theory of a probability preference. And she does this through the use of a model in which objective probabilities are given as initial data, and showing that these probabilities are then twisted by *Ss*. It does explain observed choices better that the probability preference account reported by Edwards, and better than any model purely based on moments (expected value, variance, skewness, kurtosis) which cannot account for the fact that other things being equal, the value of the least likely outcome is preferred by *Ss*. Lichtenstein interprets the tendency to favor of the least likely outcome as a bias indicating "that *Ss* overestimate small probabilities, so that the subjective probability curve is not linear with objective probability" (Lichtenstein, 1965, 168). Hence, replacing the set of objective probabilities with a set of subjective probabilities obtained by twisting objective probabilities is a modelling trick that allows to write a model that describes well-observed choices.One may remark that the assumption of objective probabilities in Lichtenstein's experiment makes it unfit to criticize Savage's axiomatic. Actually, the fact of assuming subjective probabilities obtained from some transformation of objective probabilities permits such a criticism. The main problem with Savage's subjective probability function is that it is too liberal. A great number of observed choices can be accommodated through the use of a well-chosen subjective probability function. Modeling choice as the result of an actual transformation of stated objective knowledge of probabilities is a way to constrain the subjective probability function within reasonable bounds. In so doing, it changes the epistemic status of the objective set of probabilities in the theory. They now represent some shared view about the objective values in the world.

Here we can identify a specific issue relating modeling strategies to the very possibility of building a theory of subjective probabilities. Lichtenstein does not claim that this modeling of choice under uncertainty through a compensatory model (SEU) is the only possibility to interpret choice data. Indeed, she also advances another, non-compensatory model of choice and another possible view about subjective probabilities. As just said, when two bets have the same EV, V and Sk, *Ss* bid more for the bet with the largest least likely amount. But there is also another phenomenon to be taken into account: When two bets have the same EV and Sk but differ in V, *Ss* bid more for the smaller V and do not seem to account for the larger least likely amount. This suggests that *Ss* are following a lexicographic decision process: "These findings suggest that *Ss* approach bets with a lexicographic ordering of relevant variables: EV is of primary importance. When deciding between bets of equal EV, *Ss* prefer low V. When both EV and V are constant, *Ss* attend to the least likely amount, avoiding a small chance of large loss." (ibid, 168) Along this interpretation of the way *Ss* process information and make their choice among three-outcome bets, the notion of probability cannot be addressed, since it would entail comparing values of probabilities within different informational environments. This lexicographic model is perfectly compatible with the idea that *Ss* do not twist probabilities.

The main outcome of the research done after Edwards (1954c; 1961) on the psychology of judgement and decision making is a complete revision of the notion of probability that was established along the opposition between objective and subjective probabilities. Research done by Lichtenstein, Slovic, Tversky and Kahneman and several other cognitive psychologists (Slovic and Lichtenstein, 1971) leads to a simple conclusion: In experiments with explicit objective probabilities or a priori stable priors elicited from *Ss*, these probabilities are always reframed by *Ss* in performing the task of the experiment. The celebrated opposition between objective probabilities à la Von Neumann-Morgenstern and subjective probabilities à la Savage is thus irrelevant. Instead, cognitive psychologists insisted on the influence of the structure of the problem-solving situation on the building of subjective probabilities.

Now, one distinction that deserves experimental scrutiny is whether and under what conditions a system of probabilities—as an expression of degrees of beliefs regarding the occurrence of events—deviates or not from objective values or from some reasonable subjective values and whether it is coherent. On this issue, the contributions of our four protagonists lead to a multiplicity of statements that need ordering. It is all the more necessary to clarify the meaning of their findings since they convey important conclusions regarding the possibility or not of improving choice environments and coming closer to some "optimal" choice.

As was shown in the previous sections, the heuristic and biases approach is conceived of as a general basis for thinking about a wide variety of situations where people have to make judgments about a probability or a likelihood. This does not necessarily imply the calculation of a numerical value (e.g., the probability that a child will get a cold if he goes to school without a coat on a cold winter day). Most often, it is just a mental operation of categorizing things according to some stereotypes. This wide set of judgments simply reflects the fact that we hold beliefs about the functioning of the world and the likelihood of unique events. The main question then is about the meaning attributed to a probability when no stable set of events is conceivable as the representative sample against which a judgment of representativeness can be elaborated. On some occasions, Tversky and Kahneman address this issue. By the end of "Availability", they uphold that

> "Although the 'true' probability of a unique event is unknowable, the reliance on heuristics such as availability or representativeness, biases subjective probabilities in knowable ways. A psychological analysis of the heuristics that a person uses in judging the probability of an event may tell us whether his judgment is likely to be too high or too low. We believe that such analyses could be used to reduce the prevalence of errors in human judgment under uncertainty." (Tversky and Kahneman, 1973, 231)

In other words, even though the whole apparatus to think about probability judgments is based on the assumption that *Ss* always use subjective probabilities, there is a substratum of non-biased sets of subjective probabilities against which some systematic bias could be identified. One may wonder how these views rely or not on the possibility of obtaining measurements of subjective probabilities. It may be that such a calculation (even an interval estimate) depends on the kind of task that *Ss* have to perform (for instance, judging a likelihood *vs* making a prediction). In any case, Kahneman and Tversky are prudent and do not claim to abide by some axiomatic theory of probabilities:

> "We use the term 'subjective probability' to denote any estimate of the probability of an event, which is given by a subject, or inferred from his behavior. These estimates are not assumed to satisfy any axioms or consistency requirements. We use the term 'objective probability' to denote values calculated, on the basis of stated assumptions, according to the laws of the probability calculus. It should be evident that this terminology is noncommittal with respect to any philosophical view of probability." (Kahneman and Tversky, 1972, 431)"

This assumes, as is the case in all the experiments discussed previously, that the features of the world—the outcomes of judgments and decisions as they will be experienced—are not modified by the subjective probabilities, as in any substantive account of rational behavior. This is in tune with certain remarks regarding situations of strategic behavior: On the chessboard or on the battlefield, "the player may tend to regard his opponent's strategy as relatively constant and independent of his own moves." (Tversky and Kahneman, 1973, 230)

We have seen on some occasions that it is an important point in the presentation of the various heuristics that the direction of the bias can be assessed with certainty. Still, there is at least one open question, which has to do with situations where some interplay between different heuristics is assumed. For instance, when discussing the application of availability to everyday situations, Tversky and Kahneman mix representativeness and availability and assume that they complement each other in a given direction to the effect of leading to distorted judgments of probability (e.g. the clinician establishing a diagnosis on the basis of recollected cases). Also, in the case of a scenario production, availability and anchoring lead to some path dependency, reducing people's ability to change their views : "The production of a compelling scenario is likely to constrain future thinking." (Tversky and Kahneman, 1973, 230).

On the same issue, some other statements by Slovic point to the fact that the relationships between subjective likelihoods or probabilities and some objective knowledge about probability is not known. "The availability of instances is affected by factors such as recency and imaginability, which may, but need not, bear any relation to the event's probability" (Slovic, 1972, 5) To Slovic, the heuristics of anchoring and adjustment needs a better understanding of adjustment mechanisms, and he seems to assume that the final value of a subjective probability calculation must be appraised against some estimated objective (interval) value in order to assess that an S systematically makes insufficient adjustments. Slovic points to two likely explanations. Either people get tired of adjusting or the anchor "takes on a special salience and people feel that there is less risk in making estimates close to it than in deviating far from it." (Slovic, 1972, 11)

In the conclusive section of "Heuristics and Biases", Tversky and Kahneman return to the meaning of subjective probabilities, and their stance is now explicitly against the axiomatic view of a measurable subjective probability in decision theory (Savage, 1954). Decision theory builds a set of subjective probabilities conditional on some normative axioms.

> "A person bets on Team A rather than on Team B because he believes

> that Team A is more likely to win. He does not infer the belief that Team A is more likely to win from the fact that he has chosen to bet on it. Thus, in reality, subjective probabilities determine preferences among bets and are not derived from them as in the axiomatic theory of rational decision." (Tversky and Kahneman, 1974, 1130)

Behind this criticism at the very foundations of SEU, the real issue is linked to this a priori assumption that subjective probabilities are an independent set of data in tune with a system of beliefs.[83]

In a lecture summarizing the heuristics and biases approach, Tversky delivers his opinion about the limits of the axiomatic approach, pointing to the impossibility for any human being to control for the coherence of his whole system of beliefs and for the experimenter to come to grips with this sytem.

> I do not believe that the coherence, or the internal consistency, of a given set of probability judgments is the only criterion for their adequacy. The judgments may be consistent among themselves but incompatible with other beliefs held by the individual. Consider a person whose subjective probabilities for all possible outcomes of a coin-tossing game reflect the gambler's fallacy. That is, his estimate of the probability of "tails" on any toss increases with the number of consecutive "heads" that preceded that toss. The judgments of such a person could be internally consistent and therefore acceptable as adequate subjective probabilities according to the criterion of the formal theory. These probabilities, however, are incompatible with the belief that a coin has no memory and is therefore incapable of generating sequential dependencies. For judged probabilities to be considered adequate, or rational, internal consistency is not enough. The judgments must be compatible with the entire web of beliefs held by the individual, and not only consistent among themselves. Compatibility among beliefs is the essence of rational judgment. Unfortunately, there can be no simple formal procedure for assessing the compatibility of a set of probability judgments with the judge's total system of beliefs." (Tversky, 1974, 157)

[83] Another challenge to the SEU model concerns the assumed independence between desirability of an outcome (payoffs) and subjective probabilities. Emotions, personal biases towards pessimism or optimism, could well play their role in an *S*' subjective probability function in the face of a decision under uncertainty. And these feelings could well be attached to some features of the payoffs. Slovic (1966c) elaborated upon an experiment to the effect of discriminating between various possible kinds of profile of subjective biasing of probabilities (partial or complete optimism or pessimism, feeling that extreme events will actually happen). Broadly, the experiment consists in asking *Ss* to look at sequential drawings with replacement of chips in a bag and to evaluate after each sequence of drawings the probability that the bag contains 30%, 40%, 50%, 60% or 70% of red chips (actually the bag contained 50% of red chips). Different systems of valuations associated with the experimental design allow to test for various possible profiles of twists of subjective probabilities attributed to each possible event. The experiment leads to conclude that *Ss* are subjectively biased by the value of the rewards, in a coherent way. However, there is a variety of profiles between Ss and for the same Ss over different trials.

Broadly, Tversky is pointing to the fact that the whole system of beliefs to be encompassed in order to assess the rationality of a single agent is beyond the theoretician's reach. However, maybe as an anticipation of his future work with Kahneman on the subjective expected utility model (Kahneman and Tversky, 1979), Tversky (1974, 157-158) recognizes that the empirical-descriptive approach of psychology and the formal-normative approach of decision theory are both concerned with improving human judgment (Tversky, 1974, 157). He calls for "a deeper theoretical analysis of subjective probability [that] will hopefully lead to the development of practical procedures whereby judged probabilities are modified or corrected to achieve a higher degree of compatibility with all these types of knowledge." (ibid., 158)

In sum, the heuristics and biases approach in its formative years contains contrasting views regarding probability calculations. On the one hand, it offers several perspectives to consider intuitive probability calculations and states of mind about likelihood of events that are connected with a theory of beliefs and knowledge. On the other hand, it contains also several statements about the possibility to operationalize theoretical results in order to reduce biases in probability calculations. The choice of the adequate model (either compensatory or not) is a practical matter.

6.2 Heuristics vs biases

From this reconstruction of the early history of the heuristic and biases approach, it seems important to dwell on the distinction between heuristics and biases.

The notion of a "bias" as a departure from a normal or expected behavior or judgment was sometimes mentioned in the 1960s, and the very first instance of a heuristic as a general pattern of information processing was not at first distinguished from a list of biases associated with it. This is quite understandable, since the building of a heuristic assumes that a wide variety of observed behaviors or judgments can be encompassed and explained within a single general theoretical structure of information processing. In the case of a heuristic, the structure is made of a set of information at the disposal of *Ss* within a controlled environment and a task to be performed based on this structure: judging a likelihood, selecting a preferred option, ranking events, calculating a probability or a probability interval, valuing or choosing a bet. A clear distinction between a heuristic and a bias could not be conceived based on only very few experiments. It was necessary for Tversky and Kahneman to accumulate enough experimental results (also reflecting a diversity of tasks and structures) to organize them within broad categories and to make possible the distinction between biases and heuristics.[84]

The distinction between a heuristic and a bias also became clearer as Tversky and Kahneman were able to make a distinction between heuristics themselves. "Representativeness" discusses the notion of heuristic only in the last two pages (451-452). There, the conjecture that the use of a heuristic can be expanded to general cases when no objective material base for calculations is available goes together with the identification of another heuristic, availability, and with a broad view about the main features that characterize each of them. Representativeness is viewed as covering

[84] In their first co-authored article the distinction was not made."The representation hypothesis describes a cognitive or perceptual bias" (Tversky and Kahneman, 1971, 110).

situations where *Ss* have to cope with an excess of information to compare an item with a reference population. Judgment then proceeds by elimination and filtering of information to reduce the cost of dealing with too much information. Judgment is connotative. As regards availability, it relies on the operation of building a set of elements stimulated by the item to be judged itself. The operation of construction of the relevant set derives from specific features of the item to be compared; it is thus denotative.

> "The representativeness heuristic is more likely to be employed when events are characterized in terms of their general properties; whereas the availability heuristic is more likely to be employed when events are more naturally thought in terms of specific occurrences." (Kahneman and Tversky, 1972, 452)

This also establishes the possibility of using both heuristics in situations where both the general characteristics and specific characteristics of an item are important. The series of articles by Tversky and Kahneman shows a permanent effort to broaden the meaning of heuristics and to establish them as a meta-theoretical structure in cognitive psychology. It is a condition for establishing biases as a characterisation of the outcomes associated with a heuristic when this heuristic applies within a *specific* structure of information and to a *specific* task. So, whereas in 1972 the status of the "law of small numbers" as a bias or as a representation hypothesis was not stabilized, two years later, the organization of the theory as a set of heuristics leading to biases dictated the organization of "Heuristics and biases" (1974). In this article, previous descriptions of biases associated with specific experiments are now given a name and classified as expressions of heuristics.

- Biases associated with representativeness are: insensitivity to prior probability; insensitivity to sample size; misconception of chance; insensitivity to predictability; illusion of validity; misconception of regression.

- Biases associated with availability are: biases due to the retrievability of instances; biases due to the effectiveness of a search set; biases of imaginability; illusory correlation.

- Biases associated with anchoring and adjustment are: insufficient adjustment; biases in the evaluation of conjunctive and disjunctive events; anchoring in the assessment of subjective probability distributions.

As regards the heuristic of anchoring, some remarks are in order. Definitely, this heuristic was developed to complement the analysis of cognitive mechanisms, and cases of pure anchoring and (mal)adjustment effects are not the most interesting cases, even though they may plague some of our daily decisions. From its first theorizing by Slovic and Lichtenstein (1968), anchoring was thought of as related to the properties of the problem structure, essentially the features of the gap between the display of information and the task to be performed. In Lichtenstein and Slovic's contributions, this gap is a contrast between the several dimensions of the cues (probabilities and payoffs) and the single dimension of the response task (bidding a selling price or buying price). From this initial gap, some anchoring effect can happen, but some adjustment mechanisms are also at play and can be analyzed with the help of representativeness or availability. The extent to which one of these heuristics is necessary to complement the analysis of the adjustment process depends also, in part, on the psychological costs for the *S*. In this respect, anchoring, more than other heuristics, has the specific function of "eas[ing] the strain of integrating information." (Slovic, 1972, 10), it is a strategy by which our mind makes further calculations easier in order to arrive at a reasonable answer, it eases "the strain that information processing places upon our memory." (ibid., 11). In an effort to put some order into the description of the heuristic of anchoring, Slovic proposed to account for some salient properties of the problem-structure, namely "concreteness" and "compatibility". Concreteness expresses the idea that a judge or decision maker "tends to use only the information that is explicitly displayed in the stimulus object and will use it only in the form in which it is displayed. Information that has to be stored in memory, inferred from the explicit display, or transformed tends to be discounted or ignored." (Slovic, 1972, 9). Compatibility accounts for the fact that *Ss* are naturally oriented toward using information of the same dimension as that associated with the task to be performed (see also Slovic and MacPhillamy (1974)). Being two ways to ease the process of information processing, concreteness and compatibility stand as additional elements for thinking about the anchoring process.

Is it possible to account for this distinction between heuristics and biases in the modeling of rational agents in economics? It is another aspect of the history of behavioral economics to understand whether the questions economists address can account for the broad distinctions between heuristics or whether the relevant facts of cognitive psychology are rather some well-identified facts.

What is at stake for a history of behavioral economics in insisting on the distinction between heuristics and biases? Precisely, as with the case of the theory of probability, I surmise that this distinction is not easily transferable within the framework of the theory of rational choice. On this point also, a history of behavioral economics should appraise the way this distinction was addressed, adapted, or transformed, as the paradigm of behavioral economics developed in the 1980s and 1990s. Without entering more precisely into this history, it is a fact that the broad theoretical framework associated with the use of heuristics has been downplayed while, over time, the identification of new biases or "effects" has exploded, in close association with an effort to measure these effects. So much so that Kahneman considered that the distinction between heuristics and biases was even sometimes "moot" and at the same time asked: "Too many biases?"(Kahneman, 1991, 143). Identifying biases and measuring them assumes that the status of the rationality assumption and the possibility of optimal behavior has been solved.

6.3 Rationality

As a last word on the Lichtenstein-Slovic-Tversky-Kahneman joint contribution to the theory of judgment and decision during this foundation period, we are now able to characterize how the issue of rational behavior was transformed and to identify several attitudes toward the idea of improving individuals' decisions.

The concept of "rationality" is seldom discussed for its own sake in the literature that we have reviewed. However, the whole system that emerges from the contributions of our four protagonists constitutes a basis of thoughts and experimental results which does not ignore the idea of rationality, nor the stated purpose of improving the way people make decisions in order to better take into account their own preferences and values.

In my view, the research carried out by Lichtenstein, Slovic, Tversky and Kahneman set the stage for a broad questioning of rationality that goes beyond the standard approach in several respects.

- First, it questions the meaning of preferences and the possibility to have true preferences in complex environments.
- Second, if ever such true preferences could be assumed in some contexts, it questions the legitimacy of markets and monetary environments to elicit these preferences.
- Third, it rejects all kinds of sophisticated optimization abilities (computing abilities).
- Fourth, it draws attention to the psychological cost of dealing with a wealth of heterogeneous information.
- And, last but not least, it draws attention to *Ss* feelings about the manipulability of their choice environment.

The picture of the agent that emerges from the heuristics and biases approach is that of a "judge" who, in making decisions, has to struggle with his own imperfections, his frailties toward a manipulable environment, and his endless concerns about the coherence of his system of beliefs.

The official and precautionary statement often repeated by the authors insists on the ecological rationality of heuristics:

> "Although these heuristics sometimes lead to biases, they are highly economical and usually effective. In general, representative outcomes are more likely than non-representative outcomes, frequent occurrences are more available than infrequent occurrences, and adjustment is a useful estimation procedure." (Tversky, 1974, 157)

Beyond this statement, there are scattered mentions of the difficulties facing individual judges in order to cope with all the aspects of a rational choice.

First, as we have seen, "Intransitivity" questions the possibility of finding coherent preferences, even when the response mode is a simple choice between alternatives. This difficulty arises in specific contexts, but it is only compatible with few dimensions or characteristics in the alternatives.

Second, "Preference reversal" questions the possibility—even in the case of true values or true preferences—to elicit the preferences that represent them if the method used focuses on monetary valuations. At least a task that favors using the whole information is likely to lead to a more faithful representation of preferences. Comparing preferences with the set of reachable baskets of goods or any kind of outcome expressed in monetary terms jeopardizes the very possibility of valuing the final states without twisting our preferences.

Third, the heuristic and biases approach rejects the idea of calculation and optimization, since these calculations will be highly dependent on the framing of the problem and the response mode chosen. Probabilities, even when "objective" (being explicitly given or taken as a personal base rate) are not given the same weight according to the task performed and the display of information. So, the main outcome of the heuristic and biases approach is that due to all kinds of heuristics, *Ss* are not able to make optimal decisions: "the failure of one's decisions to appropriately reflect his personal values can be considered one of the most fundamental aspects of non-optimal decision making." (Slovic, 1972, 7)

What is the exact meaning of non-optimality here? Does it assume some known optimal judgment or decision that is based on well-identified and stable values and preferences? It seems to be a matter of circumstances. Most often, it is assumed that biases operate in some definite direction and that some decision making technology or some choice architecture might improve choices. Even in this case, the exact "amount" of a bias and "how much" decision-making can be improved is not known for sure. For instance, Slovic (1972, 8) and Tversky (1974) wonder about the possibility of assessing what is an optimal decision or what are true preferences. When it is considered that decisions are biased in some determinate direction, even if true intrinsic values are not fully known to *Ss*, it is expected that some debiasing might be possible, and that this debiasing will not bear on the perceived system of preferences, thus helping *Ss* make better decisions.

Another very important issue about rationality departs from the usual treatments in economics. It has to do with the way agents think about their environment and "nature", i.e., about the manipulability or not of their environment. In economics, the stochastic aspects of the environment are assumed to be objective data known by all agents. In contrast with this, as we have shown, cognitive psychologists wonder about the attitude of *Ss* toward the problem-solving situation they are facing and the kind of confidence or attention needed in information processing. This somewhat anticipates later discussion by Kahenman about the functioning of generic modes of cognitive function System 1 (intuitive) and System 2 (deliberative) (Kahneman, 2003).

A fundamental assumption is that the very fact of adopting a systematic method to make a decision means that we are confident that the problem-solving situation we are facing is not devised to take advantage of this method. As Tversky puts it in "Intransitivity", in using a method to make a decision, we assume "that the world is not designed to take advantage of our approximation methods." (Tversky, 1969, 46) In the same vein, as already pointed out, Slovic and Lichtenstein assert that *Ss* use strategies that avoid using too much cognitive effort if they think that the environment is not manipulated in such a way as to take advantage of their strategies (Lichtenstein and Slovic, 1971, 55). Given what has just been said, it is thus necessary to reflect on the potential effects of a choice architecture on *Ss'* beliefs about their choice environment. As far as the issue of rationality is concerned, Lichtenstein, Slovic, Tversky and Kahneman question the usual boundaries between a parametric decision and a strategic decision. These boundaries are no longer fixed a priori, they are part of the problem-situation and must be handled carefully both in experiments and in the generalizations we make of it. How this issue is addressed through the development of behavioral economics needs to be documented. It is linked with—but not reducible to— hot issues about ecological rationality, and the possibility that individuals adapt their representations of their environment through learning (Gigerenzer et al., 1988; Gigerenzer, 1994).

Elements for thinking about circumstances that prevent optimization and ecological rationality arguments were already present in the early 1970s. Tversky, in his lecture before the Royal Statistical Society (1974), probably went the farthest in thinking about the normative consequences of the new representation of individual agency promoted by the heuristic and biases approach. Even though the standard apparatus of rational choice was rejected, the heuristic and biases approach set out new criteria to appraise rationality and departures from rationality. The notions of optimization and preferences were replaced with heuristics and learning. The most important thing is whether judges (*Ss*), within their environment, are likely to learn from past experience. In our daily experiences, events we are experiencing are not coded and categorized in such a way that we can gain intuitions about statistical laws. Human beings—whatever the reason for this—do not tend to interpret all their judgments as instances of a common subset of like judgments characterized by the same random structure. "The failure to realize that judgmental operations are repetitive—even when they apply to unique events—is a major obstacle for effective learning." (Tversky, 1974, 7) Consequently, Tversky called for the addition of new variables into the theory of decision under uncertainty, namely, characteristics about the nature of uncertainty (a pure chance event vs an event with the same probability but involving skill, or a risky event vs an ambiguous event) (ibid., 158). Taking account of all due characteristics of agents' systems of beliefs, he called for the development of "practical procedures whereby judged probabilities are modified or corrected to achieve a higher degree of compatibility with all these types of knowledge" (ibid., 158-159)

As a conclusion, it seems reasonable to assert that the heuristics and biases approach led to a representation of rationality enriched with individuals' representation of the problem-situation they are facing. Hence, the normative features of any model of decision making, the identification of biases, should be supported by a cautious examination of the way *Ss* build for themselves a representation of their environment.

7 A final word: the L-S-T-K nexus

The motivation for the present study was that within a short time span (mid 1960s - mid 1970s), Sarah Lichtenstein, Paul Slovic, Amos Tversky and Daniel Kahneman developed new theoretical perspectives on human judgments and choices in various experimental contexts. Their approach proceeded from a criticism of the standard models of rational behavior proposed by economists and decision scientists. On the way, these psychologists also abandoned or transformed some theoretical and methodological tenets of their own discipline, notably as regards the practice of measurement and the assumption of the existence of certain stable underlying psychological continua of feelings that human beings use to evaluate their environment. They also set the stage for various modeling strategies that can accommodate compensatory and non compensatory representations of human behavior. However, as it stood in the middle of the 1970s, the heuristics and biases approach did not claim to provide a new, fully-fledge theory of rational behavior that could be opposed to the *homo economicus* paradigm. However, the heuristics and biases approach called into question the meaning of a number of concepts and views that structured the theory of rational behavior in economics ; and all these views were based on similar experimental and methodological principles that shaped their unity. Henceforth, the conditions for a comparison between the economic theory of rational choice and the cognitive psychology of choice were met. After the seminal contributions of Lichtenstein, Slovic, Tversky and Kahneman, the history of behavioral economics could start.

And indeed, it was not long before the first results in cognitive psychology would find their way into economics. As is well known, Kahneman and Tversky's prospect theory was to play a driving role in convincing economists of the potential of cognitive psychology for the study of economic behavior (Kahneman and Tversky, 1979). But already even before this landmark contribution, the very first experimental results on preferences obtained by Tversky, Lichtenstein and Slovic were already mentioned by specialists of experimental economics. Charles R. Plott, an experimental economist from Caltech, regarded Tversky's and Lichtenstein and Slovic's cycles of intransitivity and preference reversals as "items of great curiosity" (Plott, 1976, 519). With David M. Grether, he carried out an experiment to show that the causes of intransitivity or preference reversals are certain neglected economic effects (income effects, misspecified incentives). Contrary to their expectations, the preference reversal phenomenon turned out to be robust (Grether and Plott, 1979).[85]

All the elements pointed out in this reshaping of the notion of rationality and of the connected notions of subjective probability and individuals' perception about their environment (manipulability vs randmoness; possibility to learn about this environment from from everyday experience), set the stage for a wealth of research agenda. With a view to meliorating individual choice, the wealth of ideas brought forward by Lichtenstein, Slovic and Tversky contained all the necessary elements for further debates on the proper use of decision technologies, education and regulation of choice environments in order to improve decision making. This is what I propose to name the Lichtenstein-Slovic-Tversky-Kahneman (LSTK) nexus.

As the behavioral economics paradigm developed[86] psychologists would, in parallel, nourish the field with new discoveries and theoretical perspectives and update of old ones. Most often, Lichtstenstein, Slovic, Tversky and Kahneman were key.

[85]In the face of persistent attempts to downplay or rebut the preference reversal phenomenon, Slovic and Lichtenstein would come back to the subject, suggesting that economists stop fighting the phenomenon of reversals and try to learn something from them (Slovic and Lichtenstein, 1983).

[86]A non exhaustive list of economists involved Colin Camerer, Werner De Bondt, Peter Diamond, Richard T. Carson, Jack Knetsch, David Laibson, John A. List, Matthew Rabin, Richard H. Thaler, George Loewenstein, Richard Zeckhauser

Kahneman and Tversky made essential contributions not only to the analysis of decision under risk (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), but also to the anlysis of framing and loss aversion (Tversky and Kahneman, 1989, 1991), Slovic identified "affect heuristics" (Finucane et al., 2000), Lichtenstein and Slovic explored the construction of preferences (Lichtenstein and Slovic, 2006), while Kahneman highlighted the distinction between System 1 and 2 (Kahneman, 2011) (first proposed by Stanovich and West (2000)) and reassessed heuristics through the idea of "attribute substitution" (Kahneman et al., 2002). This would not have been possible without the LSTK nexus.

References

Abelson, R. P. and Tukey, J. W. (1959). Efficient conversion of non-metric information into metric information. In *Proceedings of the Social Statistics Section of the American Statistical Association*, pages 226–230.

Adams, E. and Fagot, R. (1959). A model of riskless choice. *Behavioral Science*, 4(1):1–10.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, pages 503–546.

Alpert, M. and Raiffa, H. (1969). Report on the training of probability assessors. *Unpublished manuscript, Harvard University.*

Alt, F. (1971). *On the Measurability of Utility [Über die Messbarkeit des Nutzens, 1934]*, pages 424–431. Harcourt Brace Jovanovich Ltd.

Angner, E. (2020). *A course in behavioral economics*. Bloomsbury Publishing.

Angner, E. and Loewenstein, G. (2007). Behavioral economics. *Handbook of the philosophy of science: Philosophy of economic*, pages 641–690.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational behavior and human performance*, 9(3):396–406.

Bar-Hillel, M. (2024). Daniel kahneman: The psychologist who won a nobel prize in economics for replacing homo economicus with flesh-and-blood humans. *Proceedings of the National Academy of Sciences*, 121(27):e2409646121.

Beals, R., Krantz, D. H., and Tversky, A. (1968). Foundations of multidimensional scaling. *Psychological review*, 75(2):127.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232.

Berg, N. (2010). Behavioral economics. *21st century economics: A reference handbook*, 2:861–872.

Block, H. D. and Marschak, J. (1960). Random orderings and stochastic theories of responses. In *Contributions to probability and statistics*, pages 97–132. Stanford University Press.

Boumans, M. (2012). Measurement in economics. *Philosophy of Economics*, 13:395–423.

Boumans, M. (2015). *Science outside the laboratory: Measurement in field science and economics*. Oxford University Press.

Brady, M. E. and Arthmar, R. (2016). The Keynes-Knight and the de Finetti-Savage's approaches to probability: an economic interpretation. *History of economic Ideas*, 24.

Bruner, J. S. and Potter, M. C. (1964). Interference in visual recognition. *Science*, 144(3617):424–425.

Camerer, C. F. and Loewenstein, G. (2004). *Advances in behavioral economics*, chapter Behavioral economics: Past, present, future, pages 3–51. Princeton University Press.

Carroll, J. D. and Green, P. E. (1995). Psychometric methods in marketing research: Part i, conjoint analysis. *Journal of marketing Research*, 32(4):385–391.

Cartwright, E. (2018). *Behavioral economics*. Routledge.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior*, 6(1):151–155.

Chapman, L. J. and Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *journal of Abnormal Psychology*, 72(3):193.

Chapman, L. J. and Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of abnormal psychology*, 74(3):271.

Chipman, J. S. (1958). Stochastic choice and subjective-probability (abstract). *Econometrica*, 26(4):613–613.

Chipman, J. S. (1960). *Decisions, values and groups*, chapter Stochastic choice and subjective probability, pages 70–95. Pergamon Press.

Cohen, J., Chesnick, E., and Haran, D. (1972). A confirmation of the inertial-ψ effect in sequential choice and decision. *British Journal of Psychology*, 63(1):41–46.

Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological review*, 57(3):145–158.

Coombs, C. H. (1958). On the use of inconsistency of preferences in psychological measurement. *Journal of Experimental Psychology*, 55(1):1–7.

Coombs, C. H. (1959). Inconsistency of preference as a measure of psychological distance. *Measurement: Definitions and theories. New York: Wiley*, pages 221–232.

Coombs, C. H. (1964). *A theory of data.* Wiley.

Coombs, C. H., Dawes, R. M., and Tversky, A. (1970). *Mathematical psychology: An elementary introduction.* Prentice-Hall.

Coombs, C. H. and Pruitt, D. G. (1960). Components of risk in decision making: Probability and variance preferences. *Journal of Experimental Psychology*, 60(5):265.

Coombs, C. H. and Pruitt, D. G. (1961). Some characteristics of choice behavior in risky situations. *Annals of the New York Academy of Science*, 89(5):784–794.

Davidson, D. and Marschak, J. (1959). Experimental tests of a stochastic decision theory. In Churchmann, C. W. and Ratoosh, P., editors, *Measurement: Definitions and theories*, pages 233–269. John Wiley.

Davidson, D., McKinsey, J. C. C., and Suppes, P. (1955). Outlines of a formal theory of value, I. *Philosophy of science*, 22(2):140–160.

Davidson, D. and Suppes, P. (1956). A finitistic axiomatization of subjective probability and utility. *Econometrica: Journal of the Econometric Society*, pages 264–275.

Davidson, D., Suppes, P., and Siegel, S. (1957). *Decision making; an experimental approach.* Stanford University Press.

Davis, J. M. (1958). The transitivity of preferences. *Behavioral science*, 3(1):26–33.

Debreu, G. (1959). *Mathematical Methods in the Social Sciences, 1959*, chapter Topological methods in cardinal utility theory. Stanford University Press.

Diamond, P. and Vartiainen, H. (2012). Introduction. In Diamond, P. and Vartiainen, H., editors, *Behavioral economics and its applications*, pages 1–6. Princeton University Press, Princeton, NJ.

Edwards, W. (1953). Probability-preferences in gambling. *The American Journal of Psychology*, 66(3):349–364.

Edwards, W. (1954a). Probability-preferences among bets with differing expected values. *The American Journal of Psychology*, 67(1):56–67.

Edwards, W. (1954b). The reliability of probability-preferences. *The American Journal of Psychology*, 67(1):68–95.

Edwards, W. (1954c). The theory of decision making. *Psychological bulletin*, 51(4):380–417.

Edwards, W. (1955). The prediction of decisions among bets. *Journal of experimental psychology*, 50(3):201.

Edwards, W. (1961). Behavioral decision theory. *Annual review of psychology*, 12(1):473–498.

Edwards, W. (1962a). Dynamic decision theory and probabilistic information processings. *Human factors*, 4(2):59–74.

Edwards, W. (1962b). Utility, subjective probability, their interaction, and variance preferences. *Journal of Conflict Resolution*, 6(1):42–51.

Edwards, W. (1968). Conservatism in human information processing. *Formal representation of human judgment.*

Edwards, W. (1969). A bibliography of research on behavioral decision processes to 1968. Memorandum report 7, Human Performance Center, University of Michigan.

Edwards, W. (1971). Bayesian and regression models of human information processing–a myopic perspective. *Organizational Behavior and Human Performance*, 6:639–648.

Edwards, W., Lindman, H., and Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological review*, 70(3):193.

Edwards, W. and Slovic, P. (1964). Information-seeking to reduce the risk of decisions. In *Predecisional Processes in Decision Making: Proceedings of a Symposium.* Tech Report USAF AMRL-TDR-64-77, 1964, 5-20.

Edwards, W. and Tversky, A. (1967). Decision making: selected readings. *Penguin modern psychology*, 8.

Feduzi, A., Zappia, C., and Runde, J. (2017). De Finetti and Savage on the normative relevance of imprecise reasoning: A reply to Arthmar and Brady. *History of economic ideas*, 25:211–223.

Finucane, M. L., Alhakami, A., Slovic, P., and Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of behavioral decision making*, 13(1):1–17.

Fisher, I. (1906). *The nature of capital and income.* Macmillan and Cie.

Fryback, D. (2005). Ward Edwards: Father of behavioral decision theory. *Medical decision making*, 25(4):468–470.

Georgescu-Roegen, N. (1936). The pure theory of consumers behavior. *The Quarterly Journal of Economics*, 50(4):545–593.

Georgescu-Roegen, N. (1958). Threshold in choice and the theory of demand. *Econometrica*, 26(1):157–168.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In Wright, G. and Ayton, P., editors, *Subjective probability*, pages 129–161. John Wiley & Sons.

Gigerenzer, G., Hell, W., and Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3):513–525.

Goldberg, L. R. (1968). Simple models or simple processes? some research on clinical judgments. *American Psychologist*, 23(7):483.

Grether, D. M. and Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The american economic review*, 69(4):623–638.

Griswold, B. J. and Luce, R. D. (1962). Choices among uncertain outcomes: A test of a decomposition and two assumptions of transitivity. *The American journal of psychology*, 75(1):35–44.

Gulliksen, H. (1956). Measurement of subjective values. *ETS Research Bulletin Series*, 21(3):1–28.

Hattwick, R. E. (1989). Behavioral economics: An overview. *Journal of Business and Psychology*, 4:141–154.

Heukelom, F. (2009). *Kahnemann and Tversky and the making of behavioral economics*. Rozenberg Publishers.

Heukelom, F. (2011). Building and defining behavioral economics. In *Research in the History of Economic Thought and Methodology*, pages 1–29. Emerald Group Publishing Limited.

Heukelom, F. (2012). Three explanations for the Kahneman-Tversky programme of the 1970s. *The European Journal of the History of Economic Thought*, 19(5):797–828.

Heukelom, F. (2014). *Behavioral economics: A history*. Cambridge University Press.

Hicks, J. R. and Allen, R. G. (1934a). A reconsideration of the theory of value, part I. *Economica*, 1(1):52–76.

Hicks, J. R. and Allen, R. G. (1934b). A reconsideration of the theory of value, part II. a mathematical theory of individual demand functions. *Economica*, 1(2):196–219.

Kahneman, D. (1991). Article commentary: Judgment and decision making: A personal view. *Psychological science*, 2(3):142–145.

Kahneman, D. (2002). biographical. https://www.nobelprize.org/prizes/economic-sciences/2002/kahneman/biographical/.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5):1449–1475.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., Frederick, S., et al. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, 49(49-81):74.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3):430–454.

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

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Krantz, D., Luce, D., Suppes, P., and Tversky, A. (1971). *Foundations of measurement, Vol. I: Additive and polynomial representations*. Academic Press.

Krantz, D. H. (1964). *The scaling of small and large color differences*. PhD thesis, University of Pennsylvania.

Krantz, D. H. (1972). Measurement structures and psychological laws: Measurement of psychological variables is closely linked to the testing of qualitative psychological laws. *Science*, 175(4029):1427–1435.

Krantz, D. H. and Tversky, A. (1971). Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, 78(2):151.

Lenfant, J.-S. (2006). Complementarity and demand theory: From the 1920s to the 1940s. *History of Political Economy*, 38(Suppl 1):48–85.

Lenfant, J.-S. (2012). Indifference curves and the ordinalist revolution. *History of political economy*, 44(1):113–155.

Lenfant, J.-S. (2018). Probabilising the consumer: Georgescu-Roegen, Marschak and Quandt on the modelling of the consumer in the 1950s. *The European Journal of the History of Economic Thought*, 25(1):36–72.

Levine, D. K. (2012). *Is behavioral economics doomed?: The ordinary versus the extraordinary*. Open Book Publishers.

Lewis, M. (2016). *The undoing project: A friendship that changed the world*. Penguin UK.

Lichtenstein, S. (1965). Bases for preferences among three-outcome bets. *Journal of Experimental Psychology*, 69(2):162.

Lichtenstein, S. and Newman, J. R. (1967). Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic science*, 9:563–564.

Lichtenstein, S. and Slovic, P. (1970). Reversals of preference between bids and choices in gambling decisions. *Oregon Research Institute - Research Bulletin*, 10(7):1–28.

Lichtenstein, S. and Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, 89(1):46–55.

Lichtenstein, S. and Slovic, P. (1973). Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of experimental psychology*, 101(1):16–20.

Lichtenstein, S. and Slovic, P. (2006). *The construction of preference: An overview*. Cambridge University Press.

Lichtenstein, S., Slovic, P., and Zink, D. (1969). Effect of instruction in expected value on optimality of gambling decisions. *Journal of Experimental Psychology*, 79(2p1):236.

Lindman, H. R. (1971). Inconsistent preferences among gambles. *Journal of Experimental Psychology*, 89(2):390–397.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, pages 178–191.

Luce, R. D. (1959). *Individual Choice Behavior. A theoretical Analysis*. John Wiley & Sons Inc.

Luce, R. D. and Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of mathematical psychology*, 1(1):1–27.

Marschak, J. (1955). Norms and habits of decision making under certainty. *Mathematical models of human behavior*, pages 45–53.

May, K. O. (1954). Intransitivity, utility, and the aggregation of preference patterns. *Econometrica*, pages 1–13.

Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological bulletin*, 100(3):398.

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*, volume 53. Cambridge University Press.

Michell, J. (2007). Representational theory of measurement. *Measurement in economics: A handbook*, pages 19–39.

Michell, J. (2021). Representational measurement theory: Is its number up? *Theory & Psychology*, 31(1):3–23.

Morrison Jr, H. W. (1962). *Intransitivity of paired comparison choices*. University of Michigan Press.

Moscati, I. (2007). Early experiments in consumer demand theory: 1930-1970. *History of political economy*, 39(3):359–401.

Moscati, I. (2018). *Measuring utility: From the marginal revolution to behavioral economics*. Oxford University Press.

Mosteller, F. and Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy*, 59(5):371–404.

Nagatsu, M. (2015). Behavioral economics, history of. *International encyclopedia of the social & behavioral sciences*, 2:443–449.

Papandreou, A. G. (1953). An experimental test of an axiom in the theory of choice (abstract). *Econometrica*, 21:477.

Papandreou, A. G., Sauerlender, O. H., Brownlee, O. H., and Hurwicz, L. (1957). A test of a stochastic theory of choice. *University of California Publications in Economics*, 16:1–18.

Peterson, C. R. and Beach, L. R. (1967). Man as an intutuitive statistician. *Psychological Bulletin*, 68(1):29–46.

Phillips, L. D. and von Winterfeldt, D. (2006). Reflections on the contributions of ward edwards to decision analysis and behavioral research. Operational Research Group, LSE.

Plott, C. R. (1976). Axiomatic social choice theory: An overview and interpretation. *American Journal of Political Science*, pages 511–596.

Pollatsek, A. (1966). On the relation between subjectively expected utility theory and the psychophysics of gambling. *Michigan Mathematical Psychology Program Technical Report, MMPP*, pages 1–66.

Preston, M. G. and Baratta, P. (1948). An experimental study of the auction-value of an uncertain outcome. *The American journal of psychology*, 61(2):183–193.

Quandt, R. E. (1956). A probabilistic theory of consumer behavior. *Quarterly Journal of Economics*, 78(4):507–36.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17):61–71.

Samuelson, P. A. (1950). The problem of integrability in utility theory. *Economica*, 17(68):355–385.

Samuelson, P. A. (1965). *Foundations of Economic Analysis*. Harvard University press.

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical association*, 46(253):55–67.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Sent, E.-M. (2004). Behavioral economics: How psychology made its (limited) way back into economics. *History of political economy*, 36(4):735–760.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The journal of the acoustical society of America*, 36(12):2346–2353.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 69(1):99–118.

Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *The American economic review*, 49(3):253–283.

Simon, H. A. (1976). From substantive to procedural rationality. In Kastelein, T. J., Kuipers, S. K., Nijenhuis, W. A., and Wagenaar, G. R., editors, *25 years of economic theory*, pages 65–86. Springer.

Slovic, P. (1962). Convergent validation of risk taking measures. *The Journal of Abnormal and Social Psychology*, 65(1):68.

Slovic, P. (1966a). Cue-consistency and cue-utilization in judgment. *The American Journal of Psychology*, 79(3):427–434.

Slovic, P. (1966b). Risk-taking in children: Age and sex differences. *Child Development*, pages 169–176.

Slovic, P. (1966c). Value as a determiner of subjective probability. *IEEE Transactions on Human Factors in Electronics*, 7(1):22–28.

Slovic, P. (1969a). Analyzing the expert judge: A descriptive study of a stockbroker's decision process. *Journal of Applied Psychology*, 53(4):255.

Slovic, P. (1969b). Structural determinants of cue utilization in judgment. *Oregon Research Institute*.

Slovic, P. (1972). From Shakespeare to Simon: Speculations–and some evidence–about man's ability to process information. *Oregon Research Bulletin*, 12(2).

Slovic, P. and Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of experimental psychology*, 78(3):1–18.

Slovic, P. and Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational behavior and human performance*, 6(6):649–744.

Slovic, P. and Lichtenstein, S. (1983). Preference reversals: A broader perspective. *The American Economic Review*, 73(4):596–605.

Slovic, P., Lichtenstein, S., and Edwards, W. (1965). Boredom-induced changes in preferences among bets. *The American journal of psychology*, 78(2):208–217.

Slovic, P. and MacPhillamy, D. (1974). Dimensional commensurability and cue utilization in comparative judgment. *Organizational Behavior and Human Performance*, 11(2):172–194.

Slovic, P. and Tversky, A. (1974). Who accepts savage's axiom? *Behavioral science*, 19(6):368–373.

Slutsky, E. E. (1915). Sulla teoria del bilancio del consumatore. *Giornale degli Economisti*, 51(1):1–26.

Staël von Holstein, C.-A. S. (1971). Two techniques for assessment of subjective probability distributions—an experimental study. *Acta Psychologica*, 35(6):478–494.

Stanovich, K. E. and West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, 23(5):701–717.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103:677–680.

Thaler, R. H. (2016). Behavioral economics: Past, present, and future. *American economic review*, 106(7):1577–1600.

Thrall, R. M., Coombs, C. H., and Davis, R. L., editors (1954). *Decision processes*. John Wiley & Sons, Inc.

Thurstone, L. L. (1927a). Equally often noticed differences. *Journal of Educational Psychology*, 18(5):289–93.

Thurstone, L. L. (1927b). A law of comparative judgment. *Psychological review*, 34(4):273–286.

Thurstone, L. L. (1927c). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, 21(4):384–400.

Thurstone, L. L. (1927d). Psychophysical analysis. *The American journal of psychology*, 38(3):368–389.

Thurstone, L. L. (1931). The indifference function. *The Journal of Social Psychology*, 2(2):139–167.

Thurstone, L. L. and Jones, L. V. (1957). The rational origin for measuring subjective values. *Journal of the American Statistical Association*, 52(280):458–471.

Tomer, J. F. (2007). What is behavioral economics? *The Journal of Socio-Economics*, 36(3):463–479.

Trendler, G. (2019). Conjoint measurement undone. *Theory & Psychology*, 29(1):100–128.

Truc, A. (2022a). Forty years of behavioral economics. *The European Journal of the History of Economic Thought*, 29(3):393–437.

Truc, A. (2022b). Interdisciplinary influences in behavioral economics: a bibliometric analysis of cross-disciplinary citations. *Journal of Economic Methodology*, 29(3):217–251.

Tune, G. (1964). Response preferences: A review of some relevant literature. *Psychological bulletin*, 61(4):286–302.

Tversky, A. (1964). *Additive choice structures*. University of Michigan.

Tversky, A. (1967a). Additivity, utility, and subjective probability. *Journal of Mathematical psychology*, 4(2):175–201.

Tversky, A. (1967b). A general theory of polynomial conjoint measurement. *Journal of Mathematical Psychology*, 4(1):1–20.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review*, 76(1):31.

Tversky, A. (1972a). Choice by elimination. *Journal of mathematical psychology*, 9(4):341–367.

Tversky, A. (1972b). Elimination by aspects: A theory of choice. *Psychological review*, 79(4):281.

Tversky, A. (1974). Assessing uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2):148–159.

Tversky, A. and Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, 71(5):680–683.

Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2):105.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Tversky, A. and Kahneman, D. (1989). Rational choice and the framing of decisions. In *Multiple criteria decision making and risk analysis using microcomputers*, pages 81–126. Springer.

Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, 106(4):1039–1061.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5:297–323.

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

Tversky, A. and Krantz, D. H. (1969). Similarity of schematic faces: a test of. *Perception & Psychophysics*, 5:124–128.

Tversky, A. and Russo, J. E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical psychology*, 6(1):1–12.

Tversky, A., Sattath, S., and Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological review*, 95(3):371.

Tversky, A., Slovic, P., and Kahneman, D. (1990). The causes of preference reversal. *The American Economic Review*, pages 204–217.

Weber, R. and Dawes, R. (2005). Behavioral economics. *The handbook of economic sociology*, 2:90–108.

White, G., Slovic, P., and Kunreuther, H. (1974). Decision processes, rationality, and adjustment to natural hazards. *Natural Hazards: Local, National and Global.*

Wilkinson, N. and Klaes, M. (2017). *An introduction to behavioral economics.* Bloomsbury Publishing.

Winkler, R. L. (1967). The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical association*, 62(319):776–800.

DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2024
GREDEG Working Papers Released in 2024

2024-01	Davide Antonioli, Alberto Marzucchi, Francesco Rentocchini & Simone Vannuccini *Robot Adoption and Product Innovation*
2024-02	Frédéric Marty *Valorisation des droits audiovisuels du football et équilibre économique des clubs professionnels : impacts d'une concurrence croissante inter-sports et intra-sport pour la Ligue 1 de football*
2024-03	Mathieu Chevrier, Brice Corgnet, Eric Guerci & Julie Rosaz *Algorithm Credulity: Human and Algorithmic Advice in Prediction Experiments*
2024-04	Mathieu Chevrier & Vincent Teixeira *Algorithm Delegation and Responsibility: Shifting Blame to the Programmer?*
2024-05	Maxime Menuet *Natural Resources, Civil Conflicts, and Economic Growth*
2024-06	Harald Hagemann *Hayek's Austrian Theory of the Business Cycle*
2024-07	Rami Kacem, Abir Khribich & Damien Bazin *Investigating the Nonlinear Relationship between Social Development and Renewable Energy Consumption: A Nonlinear Autoregressive Distributed Lag (ARDL) Based Method*
2024-08	Abir Khribich, Rami Kacem & Damien Bazin *The Determinants of Renewable Energy Consumption: Which Factors are Most Important?*
2024-09	Abir Khribich, Rami Kacem & Damien Bazin *Assessing Technical Efficiency in Renewable Energy Consumption: A Stochastic Frontier Analysis with Scenario-Based Simulations*
2024-10	Gianluca Pallante, Mattia Guerini, Mauro Napoletano & Andrea Roventini *Robust-less-fragile: Tackling Systemic Risk and Financial Contagion in a Macro Agent-Based Model*
2024-11	Sandye Gloria *Exploring the Foundations of Complexity Economics: Unveiling the Interplay of Ontological, Epistemological, Methodological, and Conceptual Aspects*
2024-12	Frédéric Marty *L'Intelligence Artificielle générative et actifs concurrentiels critiques : discussion de l'essentialité des données*
2024-13	Leonardo Ciambezi *Left for Dead? The Wage Phillips Curve and the Composition of Unemployment*
2024-14	Sophie Pommet, Sylvie Rochhia & Dominique Torre *Short-Term Rental Platforms Contrasted Effects on Neighborhoods: The Case of French Riviera Urban Destinations*
2024-15	Katia Caldari & Muriel Dal Pont Legrand *Economic Expertise at War. A Brief History of the Institutionalization of French Economic Expertise (1936-1946)*
2024-16	Michela Chessa & Benjamin Prissé *The Evaluation of Creativity*

2024-17 Supratim Das Gupta, Marco Baudino & Saikat Sarkar
Does the Environmental Kuznets Curve Hold across Sectors? Evidence from Developing and Emerging Economies

2024-18 Patrice Bougette, Oliver Budzinski & Frédéric Marty
Ex-ante versus Ex-post in Competition Law Enforcement: Blurred Boundaries and Economic Rationale

2024-19 Alexandre Truc & Muriel Dal Pont Legrand
Agent-Based Models: Impact and Interdisciplinary Influences in Economics

2024-20 Sanjit Dhami & Paolo Zeppini
Green Technology Adoption under Uncertainty, Increasing Returns, and Complex Adaptive Dynamics

2024-21 Adelina Zeqiri, Issam Mejri & Adel Ben Youssef
The Metaverse and Virtual Reality in Tourism and Hospitality 5.0: A Bibliometric Study and a Research Agenda

2024-22 Adel Ben Youssef, Mounir Dahmani & Mohamed Wael Ben Khaled
Pathways for Low-Carbon Energy Transition in the MENA Region: A Neo-Institutional Perspective

2024-23 Dorian Jullien & Alexandre Truc
Towards a History of Behavioral and Experimental Economics in France

2024-24 François Claveau, Jacob Hamel-Mottiez, Alexandre Truc & Conrad Heilmann
The Economics of JEM: Evidence for Estrangement

2024-25 Alexandre Truc, François Claveau, Catherine Herfeld & Vincent Larivière
Gender Homogeneity in Philosophy and Methodology of Economics: Evidence from Publication Patterns

2024-26 Flora Bellone, Arnaud Persenda & Paolo Zeppini
The Rise of China in the Global Production Network: What Can Autocatalytic Sets Teach Us?

2024-27 Jeanne Mouton & Benoit Rottembourg
Auditing the Ranking Strategy of a Marketplace 's Algorithm in the Frame of Competition Law Commitments with Surrogate Models: The Amazon 's Buy Box Case

2024-28 Patrice Bougette & Frédéric Marty
Transport and Competition Law

2024-29 Frédéric Marty
La rationalité économique de la distinction ex post / ex ante en matière de concurrence et de régulation dans le secteur numérique

2024-30 Masaki Aoyagi & Maxime Menuet
Incentive Compatible Information Disclosure

2024-31 Jean-Sébastien Lenfant
The Lichtenstein-Slovic-Tversky-Kahneman Nexus. A Prehistory of Behavioral Economics (1969-1974)