Kahneman and Tversky and the making of behavioral economics

Heukelom, F.

Publication date
2009

Document Version
Final published version

2. Measuring decisions and measurement as decision in postwar psychology

1. Psychology and measurement in Michigan

From the 1950s to the 1970s the University of Michigan was the center of American psychology. It grew from seven faculty members in the late 1940s to some 225 faculty members in the second half of the 1960s [Krantz – Interview (2008), see also e.g. Peckham (2005), pp.245-266, Frantilla (1998)], and hosted the Institute of Social Research (ISR), Clyde Coombs' (1912 - 1988) Michigan Mathematical Psychology Program and Ward Edwards' (1927 - 2005) Engineering Psychology Laboratory and behavioral decision research. Over the years the ISR has received much attention in the literature [e.g. House et al. (2004), Bulmer (2001), Hyman (1991), Hollinger (1989)], while the history of Coombs' mathematical psychology and Edwards' behavioral decision research has not been fully explored. For a history of Daniel Kahneman and Amos Tversky's behavioral economics the important place to look might appear to be the ISR. Under the heading of the ISR, George Katona conducted his surveys on consumer confidence at the Survey Research Center (SRC), and even coined the label 'behavioral economics' to refer to this research. However, this chapter shows that the ISR is unimportant for the history of Kahneman and Tversky's behavioral economics, and that instead it was mathematical psychology and behavioral decision research that constituted the starting point for their subsequent collaboration.

To understand this, first the historical and organizational characteristics of the University of Michigan and its department of psychology need to be defined in more detail. Subsequently, because of its remarkable absence in the history of Kahneman and Tversky's behavioral economics, the ISR and its different centers need to be briefly discussed. After that, the third section deals with the relevant themes in mathematical psychology in the period roughly between 1950 and 1975. The fourth section describes the background and rise of behavioral decision research during the same period. Finally, the fifth section illustrates the close link between mathematical psychology, decision theory and behavioral decision research.

2. Psychology at the University of Michigan and the Institute for Social Research

David Krantz (1938-) and Robyn Dawes (1936-), two key actors in the Michigan Mathematical Psychology Program in the 1960s and 1970s, recall how the department of psychology at the University of Michigan grew tremendously during the postwar years.[4] In the immediate postwar years, before Coombs arrived in 1949, the department consisted of seven (voting) faculty members. As said, over the next two decades it expanded tremendously. Not all of these faculty members were full time employed by the department of psychology, although all could vote. By the late 1960s and 1970s the department of psychology employed roughly 60 full-time equivalents. Researchers had a part-time, or often even a zero-time contract with the department and held part time contracts with other institutions such as the ISR and the medical science departments. In fact, a considerable number of psychologists were working at the children's hospital, in the mental health program or in other medical science departments of the University of Michigan [Krantz – interview (2008)]. Still other psychologists were partly or wholly financed by external funds or grants. Coombs and his Mathematical Psychology Program, for instance, were financed through a grant from the National Institute of General Medical Science [Dawes – interview (2008)]. However, these multiple affiliations should not be seen as the result of vying for research funds among the psychologists. In fact, just the opposite was the case: there was enough money for nearly everyone to pursue their own ideas and interests in a general atmosphere of "live and let live" [Krantz – interview (2008), Dawes – interview (2008)]. Moreover, although the employer undoubtedly to some extent constrained the research, it was generally a non-binding way. Dawes, for instance, was employed for a year by the ISR and had an office in their building, but conducted very little work for them and continued working with Coombs and the mathematical psychologists [Dawes – interview (2008)].

These characteristics are important because it meant that if some psychologists, or groups of psychologists did not want to meet each other or discuss the merits of each other's work, they never had to because of the general availability of funds. It is in this light that the relationship between Coombs and Edwards should be seen. Both were strong, but very contrasting personalities who each had very different scientific programs, and the large number of people around and the general availability of funds ensured that they could conduct their own research programs without ever really having to confront one another. Furthermore, when two researchers with different backgrounds and research projects were interested in each

[4] Interview of the author with Krantz, Columbia University, New York, June 20, 2008. Interview of the author with Dawes, Carnegie Mellon University, Pittsburgh, June 23, 2008.

other's work, or interested in perhaps joining forces, there was little if any pressure to do so. Thus, Coombs and the other mathematical psychologists were aware that their work concerning the axioms of measurement was in one way or another related to the measurement methods used at the ISR, and vice versa the researchers at the ISR were equally aware of the work of Coombs and others [Krantz – interview (2008)]. But in day-to-day practice both groups simply pursued their own research agendas.

In the 1950s-1970s, the department of psychology was divided into ten fields of specialization: experimental, mathematical, physiological, personality, social, community, industrial organization, and the two largest, clinical and counseling psychology. Later, physiological psychology was relabeled biological psychology and mathematical psychology became part of experimental psychology, illustrating the close connection between both. But this classification was relatively loose and more a matter of classifying what people were doing than assigning them what to do. Coombs was only associated with mathematical psychology, but Edwards' Engineering Psychology Laboratory was associated with both mathematical and experimental psychology. Tversky too was associated with both specializations. Krantz was related to experimental, mathematical, and physiological psychology and Dawes to mathematical and clinical psychology. Thus, the department of psychology had an organization, both in terms of where the money came from and in terms of fields of specialization [Krantz – interview (2008)]. But, as a result of the large number of faculty members and the availability of funds, the organization in the 1960s was not tightly knit, so that everyone could more or less do what he or she wanted to do [Krantz – interview (2008), Dawes – interview (2008)].

Related to, but organizationally distinguished from the department of psychology were the centers organized under the Institute for Social Research (ISR). The Survey Research Center (SRC) was established by psychologist George Katona in 1946, who pioneered a social survey research on consumer sentiment. To finance the war, the American government had issued a large number of war bonds and with the end of the war in sight it wanted to know how likely it was that American consumers would maintain or liquidate these bonds. Because Katona felt that he could not immediately ask people what they would do with their money, he proposed starting with some general questions that would comfort the respondents and would get him or her to start thinking about their own budgets and future prospects. In these consumer confidence surveys Katona was the first to use the term 'behavioral economics,' as early as 1947 [Juster (2004), p.120]. Three years later, following the death of its founder Kurt Lewin (1890 - 1947) the Research Center for Group Dynamics (RCGD) was moved from the Massachusetts Institute of Technology (MIT) to Michigan. The two groups remained separate but were brought together under the newly-created Institute for Social Research (ISR). Since 1949 the ISR has been joined by other centers, and new centers have been created within the body of the ISR, such as the Center for Political Studies (CPS) and the Population Studies Center (PSC). In the 1960s, the ISR for a while contained the Center for Research and the Utilization of Scientific Knowledge (CRUSK), which later dissolved and disappeared. The scientists staffing the different centers of the ISR were social scientists and a few statisticians. Many were sociologists or political scientists, but the majority in the 1950s-1970s were the psychologists [Krantz – interview (2008)].

In order to protect its general university funds, the University of Michigan insisted upon creating of the SRC in 1946 that it was to be funded entirely through grants and contracts; a policy that was also applied to the ISR when it was created in 1949. This did not have any immediate financial implications as enough grants and contracts were available over the years. It did mean, however, that the ISR could not offer tenure to those it employed. There were always certain researchers who were the last to leave whenever funds ran out, but even these senior researchers and directors could never obtain tenure at the ISR [Krantz – interview (2008), Juster (2004), Hollinger (1989)].

The ISR and its research are remarkably *absent* in the main story of this thesis. Because of Katona's work on consumer confidence at the SCR and the term 'behavioral economics' that he created, Coombs' research and that of the mathematical psychologists is seemingly close to the psychological and social measurement of the ISR. One would imagine that there was some connection. Furthermore, Kahneman and Tversky's work during the 1960s and 1970s on human beings' perceptive and cognitive capacities, discussed in Chapters three and four, seems to be, at the very least, related to survey research on consumer confidence. In addition, one could point to the fact that Dawes was employed for a year by the ISR while working for Coombs' Mathematical Program. However, until Kahneman made a connection between his and Tversky's work and the economic and psychological survey work through his program of hedonistic psychology in the late 1990s, no link of any significance can be observed. The ISR and the research conducted at its centers are noteworthy because of their complete absence in the history of mathematical psychology, behavioral decision research, Kahneman and Tversky's collaboration and their behavioral economics.

The reason for this is that, although both the ISR and the mathematical psychologists and behavioral decision researchers were working on psychology and measurement, in fact the two groups conducted very different projects. The ISR worked on measuring actual social, psychological, and economic characteristics of the American population. In the social psychological tradition of Louis Leon Thurstone (1887 – 1955) and Kurt Lewin it measured the attitudes of the population to spending and saving, consumer confidence concerning the performance of the economy in the near future, and so on. Mathematical psychologists and behavioral decision researchers, on the other hand, investigated the underlying characteristics of the human being regarding decision making. In their research a measurement was understood to be a human decision between two stimuli, and was thus considered to be part of experimental psychology. In a general sense both groups were working in psychology and were concerned with measurement. But their actual research was only distantly related. Kahneman and Tversky's research grew out of mathematical psychology and behavioral decision research. Therefore, the ISR is not relevant to understanding the rest of this story.

3. Mathematical psychology

The tradition of using mathematics in the study of psychological phenomena goes back to Gustav Fechner (1860) and is closely related to experimental psychology. Fechner's psychophysics was a two-sided attempt to create a mathematical basis for a scientific field of psychology and to create a mathematical basis for (scientific) measurement. As measurement occurs through human observation, a theory of human observation is at the same time a theory of measurement, and a psychological theory of observation or perception [Heidelberger (1993, 2004), Daston and Galison (2007)]. As a basis for his psychophysics, Fechner posited the idea that the just noticeable difference (jnd) is constant across individuals. For instance, the smallest increment in the brightness of a light bulb glowing at a specific brightness, at a specific distance, in a specific environment, etc., Fechner supposed to be the same across individuals. However, jnd as a basis for psychophysics eventually fell victim to its own success in the 1920s after too many jnd's had been reported and the idea of one constant jnd for each stimulus across individuals could no longer be maintained [Gigerenzer (1987a), p.8].

Thurstone sought to save the psychophysical program in the 1930s by proposing frequency distributions instead of jnd's as a basis [Thurstone (1927a,b,c)]. Thurstone assumed that if you give two different stimuli to the individual (say two lights of different brightness) a large number of times, the relative frequency with which the individual judges the one to be larger than the other will reflect which of the two was the brighter. Moreover, and very important, when the order of objective values of the stimuli was independent of which individual perceived it, it was equally valid to ask a large number of individuals, instead of one individual a great number of times. If you wanted to know which of the two light bulbs was the brighter you could ask any individual, but if you wanted to know whether a Ferrari or a Bugatti is the more beautiful car, this method would be invalid as the order would differ across individuals. What one could ask, however, was whether drivers of a Saab consider the Ferrari or the Bugatti more beautiful, or whether Americans with a yearly income of over $20,000 have positive or negative expectations of future economic growth, or whether Protestants consider Catholics or Muslims more benevolent. These measurements were possible when one assumed that there is one preference of *the* Saab driver for either Ferrari or Bugatti, one preference of *the* Protestant for Catholics or Muslims when it comes to benevolence, and so on.

Similar to Fechner, Thurstone's theory was as much a psychological theory of human perception as it was a theory of scientific measurement. Thurstone developed his theory of measurement to facilitate his own research on attitude measurement. In 1928, he published a small book in which he reported the results obtained from having conducted an extensive investigation on religious attitudes, investigating for instance whether *the* Protestant has an attitude to the relative importance of work and leisure that is different from *the* Catholic [Chave and Thurstone (1928)]. In a one-time attempt to extend this work to economic demand theory, after holding discussions with Chicago economic colleague and friend Henry Schultz, Thurstone sought to construct the attitudes of the individual to different combinations of hats, shoes, and overcoats. The article was published in *The Journal of Social Psychology*, but Thurstone sought to connect experimental psychology and economics by labeling the curve that connected the different combinations of goods between which the individual was indifferent an "indifference function" [Thurstone (1931)]. Thurstone

(1931) was picked up by a few economists in the 1930s-1950s [Moscati (2007)], but was, to the best of the author's knowledge, ignored by experimental, social, and mathematical psychologists.

Thurstone's measurement program was not the only existing measurement program. In the 1940s and 1950s also the representational theory of measurement rose to prominence. The most important contributor to the representational theory of measurement at this time was Stanley S. Stevens (1906 – 1973). Stevens' program was strongly inspired by Bridgman's operationalism [Bridgman (1927)], and defined measurement as the operation of assigning numerals according to a rule. Stevens distinguished between different types of measurement, ranging from the mere assignment of numerals without any further restrictions such as in the number of players on a football team, to that of ratio-measurement, in which it had to make sense to add, subtract, multiply and divide the numerals. The main question Coombs, a student of Thurstone in the 1930s, and later mathematical psychologists were interested in was whether it was possible, and if so how, to combine Thurstone's measurement approach with the representational measurement tradition.

The term 'mathematical psychology' was coined by Thurstone in the 1930s but acquired common usage in the early 1950s following the creation of Coombs' Michigan Mathematical Psychology Program in 1949. The key importance of Thurstone is always mentioned when the origins of mathematical psychology are set out [e.g. Frederiksen and Gulliksen (1964), Laming (1973), Luce, Bush, Galanter (1963a), Tversky (1991), Stevens (1951)], but the driving force behind mathematical psychology as a separate field in psychology was Coombs. An important catalyst was a two-month summer institute in Santa Monica in the summer of 1952, organized by Coombs and mathematician Robert Thrall, not incidentally a summer institute that also played an important role in shaping the newly created field of behavioral decision research and equally important in the history of game theory as revealed by historians of economics [e.g. Dimand (2005), Weintraub (1992), Lee (2004)]. The Santa Monica conference brought a range of psychologists, economists and other scientists working on the mathematical and experimental investigation of decision making together and thus facilitated the start and progress of much prominent research. Leading mathematical psychologists from the late 1950s onwards include, besides Clyde Coombs, David Krantz, and Amos Tversky, R. Duncan Luce (1925-), Patrick Suppes (1922-), and William Estes (1919-).

The contributions made to the field increased so much that in 1964 the *Journal for Mathematical Psychology* was founded.[5] This gave self-proclaimed mathematical psychologists a more solid basis. However, it had not yet become a society. In 1975 the board of editors of the *Journal for Mathematical Psychology* discussed the possibility of a merger with the Psychometric Society and its journal *Psychometrika*. This effort was due to the financial mismanagement of *Psychometrika* and the general desire of both groups to secure their financial future by combining conferences, journal administration and so forth. But, in addition, it was argued by individual members and the board of editors of both the *Journal for Mathematical Psychology* and *Psychometrika* that also content-wise the merger might be beneficial. In the end, two proposals were put forward for voting in the two groups, one in which the two would be completely merged into one society with two journals and one in which two divisions would exist, each having their own journal under the umbrella of one overarching society. But although Coombs, Krantz, and Tversky had all indicated to Luce, one of the editors of the *Journal for Mathematical Psychology* that they would vote in favor of a merger, both proposals were rejected. In response, the editors of the *Journal for Mathematical Psychology* proposed in 1976 to create the Society for Mathematical Psychology.[6] This proposal was accepted and the Society was officially founded in 1977.

Mathematical psychologists defined their field not on the basis of a particular understanding of psychological phenomena, but instead on the basis of a method of investigation of psychological phenomena. The field was characterized as "the attempt to use mathematical methods to investigate psychological problems," and it was thus, "not defined in terms of content but rather in terms of an approach" [Coombs, Dawes, and Tversky (1970), p.1]. It signified "not the study of a particular type of behaviour or the delineation of some new class of psychological phenomena but, rather, the application of new techniques to traditional psychological problems" [Laming (1973), p.1]. Mathematical psychology was defined rather broadly as an attempt to use theories and techniques from the field of mathematics to represent and investigate psychological phenomena. As a result, all research that applied mathematics to what could be considered psychological phenomena in principle fell

[5] The founding committee consisted of Richard C. Atkinson, Robert R. Bush, William Estes, R. Duncan Luce and Patrick Suppes. This paragraph draws on letters and minutes from the archive of Luce in Harvard University.
[6] The journal editors were Wlliam Batchelder, William Estes, B.F. Green, and R. Duncan Luce.

under the heading of mathematical psychology. This is illustrated by the three-volume *Handbook of Mathematical Psychology* (1963-1965) that started its exposition of what mathematical psychology is with a list of thirty-nine "Basic References in Mathematical Psychology."[7] Mathematical psychology aimed to synthesize all mathematical approaches to individual human behavior.

The scope of this list of basic references turned out to be more wishful thinking than an actual reflection of research conducted by mathematical psychologists. The inclusion of economist Kenneth Arrow and political scientist Herbert Simon suggested a synthesis that did not exist. Mathematical psychology was supposed to include all mathematical reasoning related to human behavior, but in day-to-day practice it was almost exclusively focused on psychophysics, measurement theory and decision theory [Gigerenzer and Murray (1987), Coombs et al. (1970)]. Mathematical psychology of the 1950s, 1960s and 1970s was about the mathematics of measurement theory and, directly related, about the mathematics of decision theory. Decision theory will be discussed in more detail below. But before that it is necessary to devote a few words to the measurement theory of mathematical psychology.

The theory of measurement developed by the mathematical psychologists was, as said, inspired by both Thurstone's and Stevens' theories on measurement. Moreover, the effort to set up a mathematical psychology program by Coombs was principally influenced by Thurstone. Yet, after a while the work on measurement of mathematical psychologists drifted away from Thurstone and towards Stevens. The self-perceived task of the mathematical psychologists became to develop further the mathematical structure of Stevens' view of measurement. The single most important publication on measurement of the mathematical psychologists were the three volumes of *Foundations of Measurement* (1971, 1989, 1990), a co-production of Krantz, Luce, Suppes and Tversky. It became the standard work on the representational theory of measurement in psychology.

In the summer of 1965, at the end of a three-week measurement workshop held at the University of Michigan, the already established scholars and long-time

[7] These basic references include among others Arrow's *Social Choice and Individual Values* (1951), N.R. Campbell's *Foundations of Science* (1957), Chomsky's *Syntactic Structures* (1957), Guilford's *Psychometric Methods* (1954), Luce and Raiffa's *Games and Decisions* (1957), Simon's *Models of Man* (1957), Stevens' *Handbook of Experimental Psychology* (1951), and Thurstone's *Multiple Factor Analysis* (1947), and *The Measurement of Values* (1959).

friends Luce and Suppes invited the "then two brightest young people working in the area" [Luce's letter to Hamada, June 23, 1986] to write a book on measurement that would summarize and synthesize all the recent work done on measurement in mathematical psychology. Despite the gap between the publication of the first volume and volumes two and three, most of the three volumes of *Foundations of Measurement* was written in the late 1960s.[8]

The main author of the first volume was Krantz, who consequently was also made its first author. The editor and first author of the second volume was Suppes, whereas the third volume was edited by Luce. The main initiator and contact person throughout the whole project was Luce. Luce and Tukey (1964), the very first article published in the *Journal for Mathematical Psychology*, formed the basis for much of the measurement work in mathematical psychology, and hence also formed an important basis for *Foundations of Measurement*. Interestingly, the authors discovered along the way that much of what they were doing had been done before by mathematician and economist Gérard Debreu [e.g. Debreu (1954, 1958, 1959a, 1959b, 1960)]. But Debreu had taken a topological approach that was difficult to understand for economists and psychologists [Krantz – interview (2008)]. The reference to Debreu is intriguing because it illustrates that economists and psychologists were working on the same phenomenon, but understood it differently. For mathematical economist Debreu his work was on utility theory, and for the mathematical psychologists it was about measurement.[9]

In the first two sentences of the first chapter of the first volume of *Foundations of Measurement* the authors stated their belief in the representational theory of measurement and the object of their book explicitly: "When measuring some attribute of a class of objects or events," they argued, "we associate numbers (or other familiar mathematical entities, such as vectors) with the objects in such a way that the properties of the attributes are faithfully represented as numerical properties. In this book we investigate various systems of formal properties of attributes that lead to measurement in this sense" [Krantz et al. (1971), p.1]. Foundations of Measurement thus referred to the mathematical properties used in the numerical

[8] This paragraph draws on the interview with Krantz and letters from Luce's archive in Harvard.

[9] Also historians of economics have focused only on the economic interpretation of Debreu's work. For instance, Weintraub and Mirowski (1994) note that "Debreu is best read as providing a handbook for the working economic theorist of the neoclassical components of economic theory. In retrospect, it is hard to read *Theory of Value* [1959b] as anything else" (p.266).

structure in the representational theory of measurement. The first chapter puts forth what were called the three basic procedures of measurement: 1) ordinal measurement, 2) counting of units, and 3) solving inequalities. It only differs from the approach set out by Stevens (1939, 1951) in that it was more mathematically refined and sophisticated. The remainder of the book is based on these three procedures. This view of measurement served as an important component in decision theory and behavioral decision research, as set out below, but it also illustrates which approach mathematical psychologists took towards the world they investigated. I have followed the example of the measurement of length as it is used in *Foundations of Measurement* (1971). The same example was employed in Stevens (1939, 1951), and Bridgman (1927), but using less mathematical formalization.[10]

In ordinal measurement the only thing that is required for measuring the length of different rods is that numbers be assigned to rods of different lengths in a consistent manner. If one labels the different rods *a, b* etc, and considers the assignment of numbers to denote the length as a function of the rods, the only thing that is required for ordinal measurement is that "$a \succ b$ *if and only if* $f(a) > f(b)$" [Krantz et al. (1971), p.2], in which the difference between $\succ$ and $>$ is the difference between the empirical and the numerical structure. That is, the numerical structure $f(a) > f(b)$ can be mapped onto the empirical or natural structure $a \succ b$. A mathematical relation, here an inequality, comes to represent the relation between two natural objects, of their relative lengths in this case. Hence, if we have assigned any number to the first rod, and the second rod exceeds the length of the first rod, the only thing required in ordinal measurement is that we assign it a larger number. This is the most general and unconstrained procedure of measurement that can be applied to any attribute of any object; provided that the empirical comparison can be made and that the sensitivity of the comparison process exceeds the disparities of the objects measured.

The procedure of counting of units, which is the second procedure in *Foundations of Measurement*, is an extension of ordinal measurement that allows for a comparison to be made of the lengths of the rods. If we wish to not only represent

[10] What I present here is a relatively brief sketch of one specific approach within the representational theory of measurement. For a methodological discussion of measurement in general and the representational theory of measurement in particular see Boumans (2004, forthcoming). For a thorough exposition of the history of measurement theory in nineteenth century experimental psychology and of the link of this psychological literature to interwar logical positivism see Heidelberger (1993, 2004). For a discussion of postwar measurement theory of the *Foundations of Measurement*, and its link to logical positivism/empiricism and Stevens, see Michel (1999, 2007).

that $a \succ b$, but also that, say, the length of rod *a* exceeds twice the length of rods of length *b*, hence $a \succ b \circ b'$, this is the procedure of measurement we require, where $\circ$ is the notation for + in the empirical structure and b' is employed to distinguish in the empirical structure between two rods of the same size. With respect to ordinal measurement, a number of extra assumptions are needed in order to establish this procedure for the counting of units. For example, to make the representation for the addition of $b \circ b'$ mathematically possible, we have to assume that two rods of lengths *b* can be represented by $2f(b)$. For the third procedure, that of solving inequalities, it requires in addition that the different distances between numbers in the numerical structure are meaningful representations for properties of the empirical structure. For instance, the numerical representation $2a + 5b = 3c$ needs to be regarded as a meaningful representation of the empirical structure.

The example illustrates that the representational theory of measurement in mathematical psychology started from mathematics and logic [Michell (2007)]. The fundamental assumption in this view of measurement is that if the scientist wants to measure, he or she needs the appropriate mathematical system. Thus, it assumed that the phenomena he or she wants to measure are clearly defined. If the scientist wants to measure length, temperature, wealth, or utility, what he or she needs to do is specify mathematically all the characteristics used in the measurement procedure *and* in the empirical system he or she wants to measure, and then afterwards apply this to the observations. When, for instance, transitivity is a mathematical requirement or a characteristic of the measurement system the scientist wants to use to measure temperature, he or she needs to start from the observation or assumption that the natural phenomenon of temperature has transitive properties. In other words, if the numerical structure that he or she uses to measure temperature has the property that it is transitive, the measurements of temperature are interpreted as transitive. This is equally true for situations where the human being is used as a measurement instrument. If the psychologist wants to measure the human perception of utilities through human beings using a measurement framework that employs transitivity, he or she needs to assume that human perception of utilities has transitive properties. Ideally one first discussed whether transitivity made sense in the case of temperature, religious attitudes or utility, but if this stage was forgotten the mathematical framework used would determine how the world was understood.

As said, from its inception measurement theory has been linked to psychophysics and experimental psychology. Heidelberger (1993, 2004) shows that from the start Fechner's psychophysics was as much a psychological theory relating objective stimulus to subjective sensation as a theory of measurement. Fechner devised his psychophysical system as a scientific foundation of measurement. It provided a scientific theory for the human body as a measurement device [Heidelberger (1993, 2004), see also Michell (2007)]. For the mathematical psychologists of the postwar period this link between psychophysics as a psychological theory and as a theory of measurement still served as the basis for their work.

It appeared to mathematical psychologists that mathematical psychology transcended the distinction between psychology and economics. As said, the basic references in mathematical psychology from the *Handbook of Mathematical Psychology* included the works of economists such as Arrow, Howard Raiffa, and Frederick Mosteller, who were considered to be important contributors to economics as well. In addition, the *Handbook* included publications written by non-economists which were considered to be important by economists for the field of economics, such as Savage and Simon. It also contained a book that was co-authored by a psychologist and an economist, Luce and Raiffa's *Games and Decisions* (1957). In addition, the summer institute in Santa Monica in 1952 provided an important impetus for both mathematical psychology and economics.

Yet, to conclude from this that in mathematical psychology economics and psychology indeed were one and the same thing, and hence unified would be a mistake. The list of basic references used in mathematical psychology contained many more books that were unfamiliar to economists than it did books that were familiar.[11] The 1952 Santa Monica conference, immediately mentioned when the history of mathematical psychology is touched upon, is important for mathematical psychology because it was organized by a mathematical psychologist, Coombs, and afterwards proved to have been the beginning of a rapid rise in mathematical psychological research. The mathematical psychologists did not make a link to the field of economics in relation to the Santa Monica summer institute.

[11] Examples include Osgood's *Method and Theory in Experimental Psychology* (1953), and Rosenbith's *Sensory Communication* (1961).

Mathematical psychology's view of economics can be further illustrated by their discussion of what economists would immediately recognize as an economic book, von Neumann and Morgenstern (1944)'s *Theory of Games and Economic Behavior,* in Coombs *et al.* (1970) *Mathematical Psychology, An Elementary Introduction*. The *Theory of Games and Economic Behavior*, Coombs *et al*. argued, is the most important modern contribution to utility theory, that is, the theory that derives from the philosophical-psychological theory of utilitarianism. It is a mathematical refinement of what is a philosophical or psychological theory. The book does, of course, have "Economic Behavior" in its title but to Coombs *et al.* economic behavior was a subset of behavior, just as social, religious, political or any other kind of behavior, and thus part of psychology. Mathematical psychologists drew on sources that economics also relied on, but they employed these sources in a different way than did economists. A similar case is Krantz, Luce, Suppes, and Tversky's reference to Debreu work as measurement theory, as mentioned above.

4. Decision theory and behavioral decision research

4.1 Decision theory

The main question that Coombs had started his mathematical psychology program with was how could Thurstone's measurement theory be brought in line with the new representational theory of measurement, a process that culminated in *Foundations of Measurement*, an axiomatic interpretation of the representational theory of measurement that has little to do anymore with Thurstone. But mathematical psychology maintained the link with Thurstone and psychophysics in general by continuing to emphasize the two-sided role of their approach as being both a theory of measurement and a psychological theory of human behavior. Furthermore, with respect to their theory of human behavior the mathematical psychologists brought their theories in line with the recent developments in theories of human behavior. The new theory they incorporated was decision theory. The thus modernized two-sided theory of psychophysics was described as follows at the beginning of Chapter eight, *Foundations of Measurement I*:

> Unlike most theories of measurement, which may have both physical and behavioral interpretations, the theory of expected utility is devoted explicitly to the problem of making decisions when their consequences are uncertain. It is probably the most familiar example of a theory of measurement in the social sciences. [Krantz et al. (1971), p.369]

In mathematical psychology the originally two-sided psychophysical theory of just noticeable differences had been abandoned, but the idea of one theory serving both as a theory of measurement and as a theory of human behavior had been maintained. No economist, perhaps with the exception of Francis Edgeworth, would have understood utility theory as a theory of measurement. But for the mathematical psychologists the representational theory of measurement and the theory of expected utility theory, or decision theory, were two sides of the same psychological coin.

Decision theory studied which decision an individual should make when he or she is faced with uncertain or incomplete information. Decision theory's revival in the twentieth century was principally due to Leonard Savage. It goes back to the second half of the seventeenth century when mathematicians and other scholars started to investigate how to calculate mathematically the optimal decision in uncertain situations. The starting point is prosaically represented by the figure of the Chevalier de Méré, a notorious gentleman-gambler at the court of Louis XIV, who asked mathematicians Blaise Pascal and Pierre de Fermat to solve a number of gambling problems. The mathematics that came out of these and similar questions was probability theory and rational choice theory [Hacking (1975), Daston (1988)]. Eighteenth-century probability theory gave rise to nineteenth-century statistics and came to pervade every corner of scientific and daily life [Daston (1983,1988), Porter (1986,1994)], and it is therefore no exaggeration to characterize this development as "probabilistic revolution" [Krüger, Daston, and Heidelberger (1987), Krüger, Gigerenzer, and Morgan (1987)].

A major problem confronting probability theory was what became known as the 'St. Petersburg Paradox,' invented by Nicholas Bernoulli in 1713. Bernoulli demonstrated that gambles could be constructed for which probability theory computed a maximum willingness to pay that was clearly at odds with intuition.[12] The most famous solution to the St. Petersburg Paradox was offered by his cousin Daniel Bernoulli in 1738 [Bernoulli (1954)]. Daniel Bernoulli distinguished between wealth and "moral wealth," in which moral wealth depends on wealth logarithmically.[13] Up until the early twentieth century the literature on mathematical theory of decision making under uncertainty consisted mainly of attempts to solve this and similar paradoxes [Edwards (1954), p.380].

Between the 1920s and the 1950s a number of ideas were introduced that thoroughly reshaped the way decision theorists, as they were now labeled, thought about decision making under uncertainty.[14] Authors such as Bruno de Finetti [e.g. de Finetti (1937,1949, 1951)] and Frank Ramsey [Ramsey (1931)] introduced the idea that probability theory could not only be applied to objective uncertainties out there in the world, such as the probability that a coin falls heads and the probability that the sun rises tomorrow, but also to subjective probabilities, that is uncertainties inside the individual of the sort 'how uncertain am I that it will rain tomorrow?,' or 'how certain am I that this second-hand car will last at least two years?' In a related development, authors such as John Maynard Keynes [Keynes (1921)] and Rudolf Carnap [Carnap (1950)] extended the theory of logic to include uncertain propositions, that is propositions with a degree of probability that is less than 1.[15] In this logical probability approach, uncertainty stems from the subject's personal belief in the occurrence of an event. The difference between objective and subjective probability is that objective probability is a probability obtained on the basis of available information and mathematical theory, a probability that is the same for everyone. Subjective probability, on the other hand, is a number attached to the personal belief of an individual. Subjective probabilities of the same event may thus differ across individuals.

The distinction between the two was not unproblematic and this is still not the case, for it is difficult to determine where to draw the line between the two. Statistical

[12] The St. Petersburg paradox has given rise to a vast array of literature. An overview of the different sides to the debate that have developed over the past 250 years can be found in Jorland (1987).
[13] As a synonym for moral wealth the original Latin text used the term "emolumentum," which in the English translation of 1954 is translated as "utility," upon the advice of Savage. See also Teira (2006).
[14] This paragraph briefly indicates a few points in a large literature. Useful overviews are Hajek (2007), von Plato (1994), and Eriksson and Hajek (forthcoming).
[15] This research can be traced back to nineteenth-century authors such as George Boole and Augustus De Morgan [e.g. Maas (2005), pp.111-122, MacHale (1985)].

data, the basis for objective probability, is information observed by human beings and can thus equally be considered input for a subjective probability. Moreover, all of the calculations for objective probability are always conducted by human beings, and can therefore also be considered as subjective probabilities instead of objective probabilities. Adherents of the so-called subjectivist or Bayesian school argued precisely this: that statistics is simply the extension of the process of human belief formation to a more formal domain. This *ipso facto* meant that the whole of statistics is a process of human decision making under uncertainty, albeit a process which is scrutinized more rigorously and recorded more formally.

In other words, the subjectivist probability theory commenced by de Finetti and Ramsey, and the logical probability approach of Keynes, Carnap and others made statistics a part of decision theory. Thus, Wald's influential *Statistical Decision Functions* (1950) stated on the first page that "[a] statistical decision problem arises when we are faced with a set of alternative decisions, one of which must be made, and the degree of preference for the various possible decisions depends on the unknown distribution $F(x)$ of X" [Wald (1950), p.2, see also Fishburn (1964)]. Decision theory was no longer only about which decision we as human beings should make given our preferences and the objective probability of different states of the world, but it was also about which conclusion should be inferred by statisticians from statistical data. Decision theory had incorporated statistics and was now an all-encompassing theory of human decision making under uncertainty.

Another new development was initiated by von Neumann and Morgenstern's *Theory of Games and Economic Behavior* (1944). In the course of constructing game theory, von Neumann and Morgenstern introduced the concept of stochastic preference, which can be found either in a weak or a strong form [see Tversky (1969) for the distinction between the two]. Stochastic preference embodies the idea that an individual who has only a very small preference for A as opposed to B, may not always correctly perceive this small difference and may mistakenly choose B. The difference is so small that he or she cannot consciously perceive it and considers him- or herself to be indifferent towards A and B. However, if the choice is repeated a large number of times, he or she will nevertheless choose A more often than B. Therefore, this individual is said to stochastically prefer A to B.

Stochastic preference eliminated the concept of indifference. Even if the individual has an infinitely small preference for A as opposed to B, this preference would show up if the choice was repeated often enough. The individual is unaware of his or her preference for A as opposed to B and considers him- or herself to be indifferent, but he or she is not, and the mathematics therefore needs to model him or her as such. In a similar way, stochastic preference dealt the final blow to experimental psychology's just noticeable differences. Just noticeable differences, as previously mentioned, were introduced by Fechner as the lowest difference in stimulus, including that between preferences, which an individual could observe. With stochastic preference, the concept of just noticeable difference had become obsolete. The experimenter could now give the subject the same choice a large number of times and from the outcome it could be inferred which of the two options he or she preferred, even if the individual him- or herself claimed to be indifferent. Stochastic preference allowed going below just noticeable differences, and thus rendered it obsolete as a starting point for psychophysics.[16]

Furthermore, von Neumann and Morgenstern (1944) cut short the discussion on what exactly utility is and how it should be measured: "We [..] assume that the aim of all participants in the economic system, consumers as well as entrepreneurs, is money, or equivalently a single monetary commodity. This is supposed to be unrealistically divisible and substitutable, freely transferable and *identical*, even in the quantitative sense, with whatever "satisfaction" or "utility" is desired by each participant [von Neumann and Morgenstern (1944), p.8, emphasis added]. With regard to the unit of analysis of decision theory, von Neumann and Morgenstern thus effectively turned the clock back to before Daniel Bernoulli, when the rational decision depended on the absolute, objective value of money. For von Neumann and Morgenstern, the agents in decision problems wanted to maximize their monetary income, not their Bernoullian utility. However, von Neumann and Morgenstern (1944) labeled this money 'utility.'

The different aspects and new ideas were organized under the heading of one theory by Leonard "Jimmy" Savage (1917-1971) in his *The Foundations of Statistics* (1954). Savage divided decision theory into two realms, a normative realm and what he labeled, an "empirical" realm, a reference to the "empirical" domain of measurement theory, as discussed above. In the normative realm, rational human beings investigated how decision making under uncertainty should be done, and

[16] Note the similarity with Thurstone's psychophysical theory of measurement as discussed above. See also Gigerenzer (1987a,b).

established rational principles for this behavior. In the empirical domain, scientists investigated whether people in everyday life behave according to the principles of the normative theory. For the research in the empirical domain Savage had the experimental psychologists in mind, but as a mathematician Savage himself stuck to developing the normative theory. The investigation of normative or rational decision behavior was considered a deductive science, an investigation that was best done in the comfort of the armchair. But it was, according to Savage, not just mathematicians who had contributed or could contribute to developing this normative theory. Important contributions had been made by economists and philosophers. Savage thus considered economics and philosophy to be deductive armchair sciences just like mathematics.

The purpose of *The Foundations of Statistics* was to bring together two themes in Western thought that go back to ancient Greece: inductive inference and reasoning [Savage (1954), p.1]. The formal investigation of reasoning is logic. Until the end of the first half of the twentieth century when Savage introduced his position, logic was only concerned with certain propositions; the purpose of Savage's book, and the logical probability and subjective probability tradition in which it stood, was to extend logic to uncertain propositions.[17] As inductive inference typically leads to uncertain propositions, in the sense that the probability that one's inference is correct is never 1, such an extension united reasoning and inductive inference. The result was what we call statistics. "Decisions made in the face of uncertainty pervade the life of every individual and organization," Savage argued, and,"[i]t may be said to be the purpose of this book, and indeed of statistics generally, to discuss the implications of reasoning for the making of decisions" [Savage (1954), p.6].

Savage's theory investigated what a rational person does in the face of uncertainty. Rationality to Savage is a theory of reasoning, either formalized or not. For certain propositions, it is generally accepted that this theory is logic. That is, the axioms of logic are widely accepted as describing and providing rules for reasoning about certain propositions. For the extension of logic to uncertainty Savage presented in the book, this was less clear, as the theory still had to be developed. That is, Savage contended that it was as yet not clear whether the axioms he presented were indeed

[17] I here forgo discussion of Savage's historical introduction. As above, one may object by pointing to for instance George Boole and Augustrus De Morgan. Savage, however, does not discuss the period in between the Bernoullis and de Finetti, Andrey Kolmogoroff and Ramsey.

the best description, and provided the best rules, for reasoning under uncertainty. The reader must subsequently thus verify for him- or herself the axioms Savage presented.

How could this be achieved? As these axioms had to do with reasoning, the reader should verify them by reasoning. In fact, Savage was cautious when seeking to convince the reader of his approach. “I am about to build up a highly idealized theory of the behavior of a “rational” person with respect to decisions,” he wrote. But,”[i]n doing so I will, of course, have to ask you to agree with me that such and such maxims of behavior are “rational.” [..] So, when certain [i.e. some - FH] maxims are presented for your consideration, you must ask yourself whether you try to behave in accordance with them, or, to put it differently, how you would react if you noticed yourself violating them” [Savage (1954), p.7]. Like the axioms of logic, Savage’s “maxims of behavior” were axioms of decision making that all rational individuals should agree upon. They were independent of any preferences or beliefs and derived from an introspection that comes before experience of any kind.

Savage’s theory did not say anything about whether people in the real world actually behave according to his theory. It is important to distinguish this from the previous point. On the one hand, Savage’s readers, philosophers, mathematicians, economists, psychologists and any other rational individuals needed to investigate through introspective reasoning, whether they agreed with the new axioms for decision making under uncertainty, as they had done for over two thousand years with the axioms of logic proper. If these axioms were agreed upon, they could then be used as rules for sound reasoning, just like the rules of logic had been used as rules for sound reasoning. On the other hand, when established and agreed upon, the question could be posed whether people in their everyday decision making under uncertainty would behave in accordance with the new axioms. It should be stressed that such an exercise could only be undertaken when the rules of reasoning had been established, at the very least by the scientists conducting the empirical investigation. In other words, in such an empirical investigation into real-life decision making under uncertainty, the rules or axioms themselves were agreed to be true, and could not be experimentally scrutinized. It would nevertheless be fruitful to conduct empirical investigation in sciences, such as psychology, which were concerned with actual decision behavior by people in the real world, and not so much with the theory of reasoning itself. In order to clarify this point Savage conceptually distinguished between the already-mentioned normative and empirical realms [Savage (1954), pp.19-20].

What conclusion needed to be drawn when a subject in an experimental setting was observed to make a decision that violated the rules governing the theory of reasoning? First of all, the theory could only be applied experimentally to subjects that can reason. Roughly, this included all normal and healthy adults; it was not useful to ask a subject that cannot reason to make a rational decision. Children, the mentally disabled and animals were therefore excluded from experimental investigation. But when subjects capable of reasoning were observed making decisions that violated the axioms, such decisions were deemed irrational decisions, or simply errors. To Savage these errors were the result of failed or too little reasoning. The individual had made a mistake in his or her reasoning or had not given it enough thought. When the subject would think further or when his or her error would be explained, he or she would recognize his or her mistake and correct his or her behavior. Savage noted that "[t]here is, of course, an important sense in which preferences, being entirely subjective, cannot be in error; but in a different, more subtle sense they can be. [..] A man buying a car for $2,134.50 is tempted to order it with a radio installed, which will bring the total price to $2,228.41, feeling that the difference is trifling. But, when he reflects that, if he already had the car, he certainly would not spend $93.85 for a radio for it, he realizes that he has made an error" [Savage (1954), p.103].

How should the empirical testing of the theory be done? The central issue here was that the information needed to make a rational decision should be the same for the experimenter and experimental subject alike. The reason was that if the experimenter was not sure that the experimental subject used the exact same information as input for his or her decision, the experimenter could never establish whether the subject was making the correct decision, or an error. For instance, if the experimental subject believes that the deck of cards of the experimenter has been shuffled unfairly, while the experimenter knows that it has been shuffled fairly, the subject could make a decision that is rational given his or her own belief, but which is irrational given the experimenter's belief. In the case of decision making under uncertainty, it should be completely clear what the uncertainty of the inference was, and what the value of the decision was. In other words, the probabilities and utilities of the different decisions involved should be clear.

To conceptually clarify, Savage invented the term ‘small world,’ as opposed to ‘grand world’ in which we live most of the time, for situations in which probabilities and utilities are clearly defined. A small world is a decision situation in which all the probabilities, utilities and consequences of the different options are clear to both experimenter and experimental subject. Therefore, “[i]t will be noticed that the small-world states are in fact events in the grand world, that indeed they constitute a partition of the grand world” [Savage (1954), p.84]. For instance, when the subject is asked to choose between five dollars for certain or a six in ten chance of winning ten dollars, this is a small world situation. The uncertainty and value of each decision are defined and clear to everyone. However, when the subject is asked to choose between a ten-year old Mercedes and a three-year old Toyota, we are in a large world decision situation. Both the value of the different options as well as the probabilities of all kinds of uncertainties associated with the two options is unclear and dissimilar for both the experimenter and experimental subject.

The value of the different options was to be measured in utilities. On the interpretation of the theory of utility, Savage fully sided with von Neumann and Morgenstern. Echoing the *Theory of Games and Economic Behavior*, Savage suggested that economists and others had been somewhat led astray in constructing complicated theories of utility, as a result of the previously-mentioned paper by Bernoulli. “For a long period,” Savage argued,” economists accepted Bernoulli’s idea of moral wealth as the measurement of a person’s well-being apart from any consideration of probability, though “utility” rather than “moral wealth” has been the popular name for this concept among English-speaking economists.” As a result, “[e]conomists were for a time enthusiastic about the principle of diminishing marginal utility, and they saw what they believed to be reflections of it in many aspects of everyday life” [Savage (1954), p.95]. However, thanks to von Neumann and Morgenstern we were now back on the right track and able to measure choice-options by using a money scale of utility. Utility equals money and is nothing more than a convenient measurement scale of preferences. “A function U that [..] arithmetizes the relation of preferences among acts will be called utility. [..] I have chosen to use the name “utility” in preference to any other, in spite of some unfortunate connotations this name has in connection with economic theory, because it was adopted by von

Neumann and Morgenstern when they revived the concept to which it refers, in a most stimulating way" [Savage (1954), p.95].

Savage's decision theory may sound rather theoretical and anything but applicable to everyday life, but normative decision theory could also be applied to questions in the world outside science's ivory tower. Indeed, the whole purpose of decision theory was to help us to make better decisions. During World War Two in the United States, the application of decision theory to everyday problems developed under the rubric of operations research [e.g. Klein (2000)]. Operations research aimed to gather all information available relating to a particular problem and then use decision theory to calculate the optimal decision. "Operations research makes the claim that, by pitting the forces of research against large-scale problems, the decision maker (manager, president, general, etc.) will be freed to devote his time to other tasks" [Fishburn (1964), p.4]. Although based on deductive introspective reasoning, decision theory was explicitly meant to be applied to real world decision problems. In turn, behavioral decision research was closely related to operations research. It was a newly-created field in psychology that, like operations research, sought to apply decision theory to real-world problems.

4.2 Behavioral decision research

The founding father of the empirical investigation of decision theory in psychology was Edwards, who in 1958 joined the University of Michigan [Philips and von Winterfeldt (2006)]. In Michigan Edwards founded the Engineering Psychology Laboratory to study and improve human decision making [Fryback (2005)]. Edwards was strongly influenced by von Neumann and Morgenstern's and Stevens' work on measurement, and was one of the first promoters of Savage's normative-empirical decision making program. Edwards admired Savage as one admires a genius, something Edwards shared with others who had read *The Foundations of Statistics,* such as Luce, Tversky and Krantz [Krantz – interview (2008)]. Edwards' 1954 article on the historical background of decision making research, "The Theory of Decision

Making," and its 1961 follow-up, "Behavioral Decision Theory," created the field of behavioral decision research.[18] Behavioral decision research was dominated by Edwards until the early 1970s. From that moment on a number of his students started to develop their own interpretations. The most successful were Slovic and Lichtenstein, who developed a constructed preferences approach that drew connections with Simon [e.g. Slovic and Lichtenstein (1971, 1973, 1983)]; and Kahneman and Tversky, who created their Heuristics and Biases approach.

Edwards and his behavioral decision research adopted the framework set out by Savage, and understood Savage's distinction between a normative and an empirical domain to be the same as experimental psychology's distinction between normative and descriptive. Decision theory was understood as providing a theoretical framework for the objective stimuli that the subject is presented with in his case of decision making under uncertainty. The self-assigned task of behavioral decision researchers was to investigate experimentally which decision subjects make with respect to this objective stimulus. In the traditional framework, experimental psychology investigated individuals' subjective perception of objective values, such as weight or brightness differences. In behavioral decision research the weights and light bulbs were replaced with the utilities and probabilities of decision theory. Given the objective values of the utilities and probabilities, decision theory determined the objective decision. Behavioral decision research then investigated experimentally which decision the subject actually made. In this way Savage's decision theory and his distinction between a normative and an empirical domain were integrated into the experimental psychological framework, in which decision theory determined the objective benchmark with which the subject's subjective decision was compared. The distinction between the normative and the descriptive was often and clearly made by behavioral decision researchers. Here is an example:

> Decision theory is the study of how decisions are or ought to be made. Thus it has two faces: descriptive and normative. Descriptive decision theory attempts

[18] Different names for Edwards' program and its offspring exist. Behavioral decision research, behavioral decision theory, and behavioral decision making are all used to refer to the same psychological program. It is not clear when and how these terms exactly originated; although behavioral decision theory has been around at least since Edwards published his second overview article in 1961. Behavioral decision research, the most commonly used label seems to have originated in the 1970s, but has been applied in retrospect to the research of the 1960s also. To avoid confusion I use the term behavioral decision research in this thesis.

to describe and explain how actual choices are made. It is concerned with the study of variables that determine choice behavior in various contexts. As such, it is a proper branch of psychology. Normative decision theory is concerned with optimal rather than actual choices. Its main function is to prescribe which decision should be made, given the goals of the decision maker and the information available to him. Its results have a prescriptive nature. They assert that if an individual wishes to maximize his expected gain, for example, then he should follow a specified course of action. As such normative decision theory is a purely deductive discipline.
[Coombs, Dawes, and Tversky (1970), p.114]

It was in this regard that Edwards was interested in economics, as exemplified by his extensive and knowledgeable discussion of economics in Edwards (1954). Like Savage, Edwards understood economics as a normative, deductive theory of human decision making, and he discussed it on an equal footing with statistics, mathematics and philosophy. Thus, Edwards noted that economics is an "armchair" science [Edwards (1954), p.14], not because he denounced economics, but because he understood economics to be an armchair science just as mathematics, statistics, and philosophy. In his classification of the field of decision theory as a) the theory of riskless choice, b) the application of the theory of riskless choice to welfare economics, c) the theory of risky choices, d) transitivity in decision making, and e) the theory of games and of statistical decision function, economics is predominantly about a) and b). Von Neumann and Morgenstern's *Theory of Games and Economic Behavior* (1944) was understood to be a deductive, armchair science as well: Game theory as a mathematical theory "can be viewed as a branch of normative decision theory" [Coombs, Dawes, and Tversky (1970), p.202].

Edwards' discussion of "economic man" should also be read in this light. Economic man for Edwards is someone who makes his choices according to the normative theory, making it therefore a normative concept. If you ask what economic man would do in a certain decision problem, you ask what the normative solution is. At the same time, economic man as the embodiment of the normative theory, forms a hypothesis about actual decision making that can be tested: "if economic man is a model for real men, then real men should always exhibit transitivity of real choices. Transitivity is an assumption, but it is directly testable. So are the other properties of economic man as a model for real men" [Edwards (1954), p.16]. But transitivity and the other properties of economic man were also the assumptions of measurement theory, as set out before. As a result, mathematical psychologists moved smoothly from measurement theory, to decision theory, signal detection theory and back.[19] So did Edwards and his behavioral decision research.

Savage and other decision theorists investigated the normative decision theories, and it was the task of psychologists, according to Edwards, to investigate the descriptive part and in turn to see how well human beings in their actual everyday decision making behave according to the normative principles set out by decision theory. What was at least just as important for Edwards, however, was the question how human decision making could be improved. The research conducted and favored by Edwards was explicitly called "engineering psychology." The perceived relevance of this research is illustrated by Edwards' comments on his visit to the North American Aerospace Defense Command in the mid 1960s [Phillips and von Winterfeldt (2006), p.5]. In this command center an enormous amount of information was gathered and decisions made by the personnel had potentially enormous consequences. Therefore, it was of utmost importance not only to know how people made decisions on the basis of uncertain information, but also to find out how the decision system could be organized such that the best decision could be made. In light of future developments in the field to be made by Kahneman and Tversky, it should be noted here that Edwards and other behavioral decision theorists did not consider the human being to be an inapt or limited decision maker in the sense of not understanding the divine rules of decision theory. For Edwards, the starting point was that the human being is very capable of making complicated decisions in situations based on uncertain information. It is just that there is only so much a single human being can do. For that reason, human beings may sometimes deviate from what is normatively the right decision, and therefore it may be useful to think about how to help human beings decide when, for whatever reason, the decision making process is especially difficult or especially important.

Edwards and behavioral decision research evaluated decisions in terms of utility and extensively referred to economists and their use of the concept of utility.

[19] Signal detection theory (SDT) is a branch of psychophysics that investigates the individual's ability to distinguish between signal and noise. In other words, it investigates decision making under noisy conditions. See e.g. Green and Swets (1964).

Nevertheless, Edwards and behavioral decision research did not understand utility in the same way as economists. For behavioral decision research utility was merely a new concept for an already existing idea in experimental psychology, that of valence. “The notion of utility is very similar to the Lewinian notion of valence. Lewin conceives of valence as the attractiveness of an object or activity to a person. Thus, psychologists might consider the experimental study of utilities to be the experimental study of valence, and therefore an attempt at quantifying parts of the Lewinian theoretical schema” [Edwards (1954), p.25, see also Frijda (1986)]. Valence measures the intrinsic attractiveness or averseness of an individual to a certain event, object or situation. Thus, if an individual is more attracted to Islam than to Christianity, Islam has a higher valence. In addition, emotions can be classified in terms of valence. Anger and fear are emotions with a negative valence, joy has a positive valence. By equating utility with valence, Edwards and behavioral decision research understood utility to be a general measurement of an individual’s attitude towards events, objects and situations. As a result, an individual preferring ten to eight dollars, was psychologically in the same situation as an individual preferring Islam to Christianity.

In behavioral decision research, the behavior of the experimental subjects was evaluated in terms of the normative benchmarks. The human being was considered to be a mechanism that reasons logically and applies Bayesian statistics. In other words, the individual was considered to be a logician and Bayesian statistician of some sort. The purpose of behavioral decision research, then, was to figure out whether this human being is a good logician and Bayesian statistician. This particular type of understanding of human behavior was neatly summarized in a paper by Rapoport and Tversky. “[The behavioral decision research] approach to the study of choice behavior,” they argued,” is based on the comparison between the normative solution of a decision problem and the observed solution employed by subjects.” As a consequence, “man is viewed as an intuitive statistician who acts in order to maximize some specified criteria while operating on the basis of probabilistic information” [Rapoport and Tversky (1970), p.118].

Edwards and the developing behavioral decision research approach created a program that took decision theory as provided by mathematicians, economists and philosophers, and especially Savage, as point of departure. It compared actual human decision making with respect to this norm, measuring the decisions made in terms of “utility,” and looked for ways to improve human decision making. But behavioral decision researchers recognized that matters were a little more complicated than they usually portrayed them. In their introduction to *Decision Making* (1967), for instance, Edwards and Tversky noted that “the distinction between what an organism should do and what it does do is slippery” (p.8). The problematic distinction between the normative and the descriptive was a recurring theme, although it was far outnumbered by the instances in which the distinction was standardly used. The problem was that in Savage’s decision theory the normative and the descriptive were closely related. The normative rules were rules that every healthy adult should agree with when thinking them carefully through. The normative decision theory was as much a prescription for optimal behavior as it was a description of an adult’s behavior who has carefully thought through which decision to make. In experimental psychology, however, the distinction was much stronger. In experimental psychology, the descriptive value of the stimulus, the sensation, was supposed to deviate from the objective norm. Thus, when decision theory was integrated into the experimental framework, the normative-descriptive distinction of decision theory risked becoming a much stronger and much more absolute distinction than it was meant to be. This was unproblematic as long as the experiments showed that most of the time subjects indeed did make their decisions according to the norms of decision theory, and it was what Edwards and his behavioral decision researchers expected to find and actually did find. However, when the experiments indicated that there might be systematic differences between the norms of decision theory and actually observed behavior, an idea that gradually developed during the 1960s (treated in the third chapter), it did become problematic.

Throughout his career Edwards wanted to maintain the initial decision theoretical understanding of the close connection between the normative and the descriptive. Until the early 1970s, his disciples in behavioral decision research kept this perspective as well. Normative theory described human behavior in situations where we really want to behave as best as we can, for instance in cases where the stakes are high. Normative theory was thus to some extent descriptive. Moreover, “[d]ecision theory may be viewed as primarily an analysis of the environment; that is, an orderly summary of those features of the environment that control behaviour.” Therefore, “[s]uch a description of the environment, combined with the simple assumptions about behaviour tendencies that the organism brings to that environment, may yield an effective description of behaviour” [Edwards and Tversky (1967), p.8]. Although the distinction between normative and descriptive was used all the time, it was at the same time clear that the two sides were closely connected and that perhaps it was not possible even to distinguish between the two.

Decision theory and behavioral decision research in the 1950s and 1960s both considered themselves to be directly related to economics and used extensive amounts of economics. It is especially Edwards' evidently extensive knowledge of economics [e.g. Edwards (1954, 1961)] which tempts the reader to conclude that here we have a case in which psychology and economics were truly integrated into one research project. But the way in which Savage, Edwards and others talked about economics does not resonate with the way in which economists spoke about economics. Such different prominent economists as Lionel Robbins, Paul Samuelson, and Milton Friedman would not have agreed to be engaged in constructing a normative theory of decision making.

Some of this incommensurability showed up in psychologists' assessment of economics. In his discussion of Samuelson's economics, Edwards was somewhat puzzled that "[i]f preference is operationally defined as choice, then it seems unthinkable that this requirement can ever be empirically violated" [Edwards (1954), p.15]. Moreover, the interpretation of utility in terms of Lewinian valence appears, if perhaps not entirely incompatible, not exactly what economists had in mind when they use the concept of utility. Thus, although the frequent references to economics in decision theory and behavioral decision research suggest otherwise, economists and psychologists understood their disciplines and the relationship between them in fundamentally different ways. Quite a few theories and concepts traveled from economics to psychology. But the way in which these theories and concepts were used in psychology was not something economists would have recognized as belonging to their field.

5. "Measurement theory in psychology is behavior theory"

Mathematical psychology was directly related to decision theory and behavioral decision research. Mathematical psychology applied mathematics to the investigation of psychological phenomena, and as both decision theory and behavioral decision research used a great deal of mathematics, a natural and direct link existed between the two. How to formulate mathematically how people should behave and how people actually do behave in situations under uncertainty, were research questions that belonged to mathematical psychology as well as to decision theory and behavioral decision research. Hence, the same scientist could naturally be perceived as being a contributor to these different fields at the same time. Tversky, Luce, and Suppes serve as examples.

But the link between mathematical psychology, decision theory, and behavioral decision research also went much further than the mere use of mathematics. Mathematical psychology's representational theory of measurement and behavioral decision research's experimental investigation of human decision making started from different perspectives, but were partly about the same subject: normative decision behavior. Mathematical psychology's representational theory of measurement used the human body as a measurement device. In the case of utilities and probabilities, for instance, the human being was used to measure human perception of utilities and probabilities, human perception of risk averseness, and human perception of loss averseness. But in order to make this a valid procedure it must be assumed that the human being as a measurement device functions consistently. Furthermore, the representational theory of measurement's definition of consistency was: according to the normative rules of decision theory. The assumption needed to be made was that the human measurement instrument behaved according to the normative decision theory.

Behavioral decision research, on the other hand, compared behavior of individuals in its experiments with the norms of decision theory, for which it used the representational theory of measurement. The two fitted neatly together. Assuming that subjects behave according to the normative rules, the mathematical psychologists set up measurement frameworks that measured the perception of utilities, risk averseness and so on. Assuming that, in general, subjects behave according to the normative rules, behavioral decision researchers investigated under which circumstances subjects made mistakes. Mathematical psychologists provided behavioral decision research with a solid theory of measurement, and behavioral decision research informed mathematical psychologists under which circumstances its human measurement instrument was less accurate.

To illustrate further why for mathematical psychologists "measurement theory in psychology is behavior theory," [Coombs (1983), p.36] it is useful to ask how experimental psychologists measured the phenomena they were interested in. How did they measure the attitude of religious people who go to church twice a day? How did they measure the perception of "rape" in terms of good versus bad? How did they measure the perception of a probability of 0.01%? How did they measure the relative utility of receiving a certain ten dollars as opposed to a 0.8 chance of receiving fifteen dollars? The answer, as already indicated, is that they measured all these psychological phenomena through the human being. “In psychological measurement, the individual is the measuring device; he plays the role of the pan balance, the meter stick, or the thermometer” [Coombs (1983), p.36]. The psychologist used individuals to measure the value of psychological phenomena of *the* individual. This could be the human being in general, it could be the member of a culture, and it could even be the individual itself. One could, for instance, use individuals as a measurement instrument to measure the individual’s risk averseness. “Psychological measurement theory is concerned with the empirical regularities in [the individual’s] behavior that justify numerical assignments to the stimuli he is responding to and/or justify numerical assignments to him” [Coombs (1983), p.36].

However, to “justify numerical assignments” to stimuli and to “justify numerical assignments” to the individual on the basis of “empirical regularities in this behavior” the psychologists needed to understand that behavior. They, in other words, needed a theory describing human behavior. The psychologists needed to understand how humans function to be able to use them as measurement instruments, just as the physicist needs to understand how the thermometer works in order to use it as an instrument. But in the case of the human being as a measurement instrument, this could not be just any understanding; it needed to be a rational understanding. Work done by Heidelberger (1993, 2004) points us to the fact that in nineteenth-century German experimental psychology, the human being functioned as a measurement device. We now see that post World War Two work regarding the representational theory of measurement and in behavioral decision research showed that in order for the human being to function as a measurement instrument the human being needed to be understood as behaving rationally. The psychologist needed to have a psychophysical or decision theoretical explanation of the individual’s response towards different stimuli in terms of *rationality*. In the case of decision making on the basis of utilities and probabilities, that theory of rationality was decision theory. Decision theory explained how an individual would rationally respond to different stimuli, and thus informed the psychologist which numeral to assign to the different stimuli. To make the link between measurement theory and decision theory one had to assume that the individual that is used as the measurement instrument behaves rationally.

What happens if we find out that this individual in fact does not always behave rationally? This question did not come up seriously until the late 1960s and will be dealt with extensively in later chapters. However, from the above we can see what happens. If individuals are found to behave irrationally, this is problematic for decision theory because it means that decision theory does not provide a good description of human behavior. As long as the deviations from decision theory are random this is not too problematic. It would be the same problem as knowing that some or even all of the thermometers do not measure exactly right but that they measure correctly on average. However, when individuals are found to deviate systematically from the norms of decision theory, it becomes a serious problem. It not only means that decision theory is not a good description of actual, rational human behavior, it also implies that measurement theory is based on flawed assumptions. For instance, if the psychologist wants to measure what the relative value of two uncertain outcomes is and assumes that people have decided rationally, he or she simply asks a few people which of the two they prefer and thus measures which of the two has the highest expected value. But if it now turns out that human beings systematically deviate from rational behavior, the psychologist cannot infer from their choices, i.e. from the measurement, which of the two options has the higher expected value.

6. Conclusion

Mathematical psychology continued experimental psychology's focus on mathematization and measurement. In the postwar period it aligned itself with the re-appearance of decision theory in the work of Savage, and with the empirical investigation of decision theory in Edwards' behavioral decision research. This alliance proved that in order to use the human being as a measurement instrument in psychology, it needed to be assumed that the individual makes its decisions rationally. Behavioral decision research was related to mathematical psychology's measurement theory in its use of measurement theory. Behavioral decision research compared experimentally actual human decision behavior with the norms of decision theory, with the explicit purpose of engineering solutions for situations in which decision making is particularly difficult, or the individual is prone to make mistakes. The three intertwined developments of mathematical psychology, decision theory, and behavioral decision research together constituted a scientific program of human decision behavior revolving around a set of axioms that determine rational or normative decision behavior. Furthermore, a comparison of human behavior to this normative benchmark could be made within a descriptive domain by means of experimental investigation.

It is tempting to conclude from the many references made to economics in mathematical psychology, decision theory and behavioral decision research that the three were connected to economics. And to some extent this is true. Mathematical psychology did incorporate economic texts, and Edwards, and to a minor extent Savage, based their research on extensive discussions of economics. But mathematical psychology, decision theory, and behavioral decision research used the economic literature for their own purposes, and they did this in ways that were at odds with economic practice.