Ivan Moscati

University of Insubria, Department of Economics, Varese, Italy

ivan.moscati@uninsubria.it

MEASURING UTILITY BY EXPERIMENTS AND AXIOMS IN ECONOMICS AND PSYCHOLOGY, 1955-1965 THE CASE OF SUPPES AND LUCE

This draft: December 24, 2014

Preliminary. Please, do not quote.

ASSA meeting, Boston, January 3-5, 2015

Session "Histories of behavioral economics"

1. Introduction

The axiomatic version of expected utility theory (EUT) put forward by John von Neumann and Oskar Morgenstern in their *Theory of Games and Economic Behavior* ([1944] 1953) generated an intense debate among utility theorists that lasted until the mid-1950s. In the course of the debate the assumptions implying EUT were clarified, and around 1952 EUT stabilized as the orthodox economic model for risky choices.[1] Building on ideas originally presented by Frank Ramsey (1931), in 1954 Leonard Jimmie Savage showed how EUT could be extended to the case in which the probabilities of uncertain outcomes are not objectively given but express the decision maker's subjective beliefs about the likelihood of the outcomes ([1954] 1971). This subjective extension strengthened the status of EUT as the orthodox economic model for risky choices.

Because the EUT axioms imply the existence of a cardinal utility function, i.e., of a function unique up to linear increasing transformations, the rise of EUT was associated with the rehabilitation of the concept of cardinal utility. This concept had been marginalized in the 1930s and early 1940s because in that period the majority of utility theorists supported a strict ordinal approach to utility analysis. EUT provided not only a theoretical justification for the use of cardinal utility, but suggested also a practicable way to measure it experimentally. In the course of the debate on EUT, Frederick Mosteller and Philip Nogee (1951) performed a first experiment in which they measured the (cardinal) utility of money of fifteen individuals on the basis of their choices between monetary gambles.

Finally, in the course of the debate over EUT, utility theorists developed a novel conception of utility measurement that reinforced the status of the rehabilitated cardinal utility.[2] In this conception, cardinal utility is not opposed to ordinal utility as an alternative and incompatible conception of a thing called "utility". Rather, ordinal and cardinal utility are conceived of as two equally legitimate ways of measuring utility, that is, of assigning numbers to the objects of choice in a manner that is convenient to predict choice behavior. In particular, ordinal utility numbers

[1] See in particular Friedman and Savage 1948, Marschak 1950, Malinvaud 1952, Samuelson 1952, and Herstein and Milnor 1953. On the debate originated by von Neumann and Morgestern's EUT and its outcomes, in Moscati 2014. Other reconstructions of that debate can be found in Mongin 1988, 2009, and 2014, Fishburn and Wakker 1995, Guala 2000, Giocoli 2003, Heukelom 2014.

[2] See Friedman and Savage 1952, Alchian 1953, Strotz 1953, Ellsberg 1954, and Friedman 1955.

are convenient to predict choices under certainty while, if one accepts EUT, cardinal utility numbers are convenient to predict choices under risk.

The debate on EUT and its outcomes paved the way for an intense research on individual decision making that had its peak from around 1955 to 1965, and centered on EUT, typically in the subjective version put forward by Savage, and cardinal utility. At the axiomatic level, various sets of axioms on preferences implying that utility is cardinal were put forward.[3] At the experimental level, a number of experiments to test EUT and measure utility were performed.[4] To this research contributed not only economists, such as Jacob Marschak, Gerard Debreu, Richard Quandt, Nicholas Georgescu-Roegen and Daniel Ellsberg, but also scholars with non-economic backgrounds who however adopted the economic framework of utility analysis to investigate decision making. Among these scholars were experimental psychologists Ward Edwards, Clyde Coombs, Sidney Siegel and Eugene Galanter, philosophers Patrick Suppes and Donald Davidson, and mathematical psychologist Duncan Luce.

The research on EUT and cardinal utility carried out in the period 1955-1965 is a part of the history of utility theory that has remained largely unexplored in the history of economics.[5] In this paper I attempt to partially fill this lacuna by reconstructing the work in utility analysis made by Patrick Suppes, Duncan Luce and their associates between 1955 and 1965.

The focus on Suppes and Luce has two major motivations. On the one hand, they abode by the expected utility framework and axiomatic method that during the debate on EUT had become orthodox in utility analysis. This allowed their papers to be published in *Econometrica* and other important economic outlets, their books to be reviewed in *American Economic Review* and other major economic journals, and their works to be cited or commented by some of the leading economists of the period such as Jacob Marschak, Kenneth Arrow, Gerard Debreu and Jimmie

3 See Marschak 1955 and 1960, Suppes and Winet 1955, Luce 1956, 1958 and 1959, Quandt 1956, Davidson and Suppes 1956, Luce and Raiffa 1957, Debreu 1958, 1959 and 1960, Georgescu-Roegen 1958, Suppes 1961.

4 See Edwards 1955, Siegel 1956, Papandreou and others 1957, Davidson, Suppes and Siegel 1957, Coombs and Komorita 1958, Davidson and Marschak 1959; Ellsberg 1961; Galanter 1962; Becker, DeGroot, and Marschak 1963 and 1964.

5 An exception is chapter 3 of Floris Heukelom’s book on the history of behavioral economics where the author discusses the work of Edwards and Coombs (Heukelom 2014, 71-95). In his review of experiments on individual decision making, also Colin Camerer (1995, 621-622) briefly discusses the 1955-1965 research in utility analysis.

Savage.[6] On the other hand, however, in dealing with utility analysis Suppes and Luce used ideas, concepts and methods coming from philosophy and psychology that were somehow unorthodox in economics. This mingling of orthodox and unorthodox elements situates their work in utility analysis really at the boundary between economics, psychology and, to a certain extent, philosophy, and makes it particularly relevant for the history of economics and its relationships with cognate disciplines.

Second, I found interesting the measurement-theoretic dimension of Suppes' and Luce's research on utility. In the course of their research, they became increasingly aware that their axiomatizations of utility could be seen as specific instances of a more general exercise, namely that of specifying axiomatically the conditions that make an object (not only utility) measurable in way rather than another. Beginning in the late 1950s, Suppes and Luce developed their axiomatic approach to measurement in a number of writings. This approach found its full-fledged expression in *Foundations of Measurement*, the book Suppes and Luce wrote in collaboration with David Krantz and Amos Tversky and whose first volume was published in 1971. By illuminating the connections between Suppes' and Luce's work in utility analysis and their theory of measurement the present paper contributes not only to the history of utility theory but also to the history of measurement theory.

2. Suppes' Axiomatizations of Cardinal Utility

Patrick Suppes (1922-2014) was a philosopher educated in the tradition of analytical philosophy who gave important contributions in fields as diverse as the philosophy of physics and probability, utility analysis, the general philosophy of science, the theory of measurement, logic and language theory, the theory of education, psychology and, more recently, neuroscience.[7] He published more than 30 books and 400 articles, and inevitably this paper deals only with a very limited part of Suppes research, namely that related to utility theory and centered on what was called the "Stanford Value Theory Project". Before arriving at this Project, it useful to outline Suppes's education and early academic career.

---

6 For comparison, none of the papers in utility analysis by psychologists Ward Edwards, Clyde Coombs, or Eugene Galanter were published in economics journals.

7 For an overview of Suppes' multifaceted scientific contributions until the late 1970s, see Bogdan 1979.

2.1. Axiomatizing Ratio Measurement

Suppes studied physics and meteorology in different universities, and eventually received a B.S. degree from the University of Chicago in 1943. After serving in the Army Air Force during the war, in 1947 he entered Columbia University as a graduate student in philosophy. There he was significantly influenced by philosopher and measurement theorist Ernest Nagel, and took courses in mathematical topics such as topology and group theory. With other Ph.D. students at Columbia, around 1948 he also organized an informal seminar on von Neumann and Morgenstern's theory of games. He graduated in June 1950 under Nagel's supervision, and in September of the same year he joined the Department of Philosophy at Stanford University, where he remained since then.[8]

In his first article, Suppes (1951) put forward a series of axioms that warrant the measurability of objects in traditional sense dating back to Aristotle (*Metaphysics*, bk. X, chap. 1) and Euclid (*Elements*, bk. V, definition 3). According to this traditional sense, measuring the property of an object (e.g., the length of a table) consists of comparing it with some other object that displays the same property and is taken as a unit (e.g., a meter-long ruler) and then assessing the numerical ratio between the unit and the object to be measured. In the current terminology of measurement theory, this traditional notion of measurement is usually called "ratio measurement". If we adopt this terminology, we can say that in his first article Suppes provided an axiomatization of ratio measurement.

He was not the first to do so. Fifty years earlier the German mathematician Otto Hölder ([1901] 1996) had laid down seven axioms on magnitudes and proved that, if magnitudes satisfy them, the ratio between any two magnitudes is well defined and one magnitude can be taken as a unit to measure the others. In 1931 Suppes' mentor Ernst Nagel had suggested a set of axioms different from Hölder's that should warrant the measurability of magnitudes in the ratio sense (Nagel 1931; see also Cohen and Nagel 1934). However, Nagel had not given any formal proof that his axioms actually deliver ratio measurability.[9] Building on Hölder and Nagel, Suppes considered a set of objects, a binary relation between these objects interpretable as the inequality relation ≤, and a binary function interpretable as the operation of addition +. He put forward

[8] The reconstruction of Suppes' studies and early career is based on Suppes 1979.

[9] More on Hölder's and Nagel's measurement theories in Michell 1999.

seven axioms concerning the set of objects, the relation ≤, and the operation + that are on the whole less restrictive than Hölder's seven axioms, and proved that they are nonetheless sufficient to warrant the measurability of the elements of the set in the ratio sense. [10]

Two aspects of Suppes' first article are relevant for the present paper. First, in a passage of his article Suppes (1951, 104, footnote 2) criticized Hölder for confusing the equivalence relation "=" with the logical relation of identity, and argued that, in their axiomatization of utility, von Neumann and Morgenstern had made a similar error by confusing the relation of indifference with that of identity.[11] This incidental comment shows that, since the early stages of his scientific career, Suppes was familiar not only with the axiomatic approaches to measurement in the tradition of Hölder, but also with von Neumann and Morgenstern's axiomatic approach to utility theory.

Second, in the 1951 article Suppes referred to ratio measurement as if it were the only possible form of measurement. Apparently, he was therefore not aware of the theory of measurement put forward by Harvard psychologist Stanley Smith Stevens in 1946, and according to which ratio measurement is only one possible form, or "scale", of measurement. While ratio measurement is associated with proportional transformations of measurement numbers (like the transformation of 1 yard into 0.9144 meters), other and less demanding scales are associated with larger families of transformations. In particular, what Stevens (1946, 678) called the "interval scale" is associated with linearly increasing transformations, while the "ordinal scale" is associated with monotonically increasing transformations.[12] Suppes' apparent unawareness of

---

[10] Suppes' seven axioms (1951, 164-165) are as follows: 1: ≤ is transitive; 2: + is closed in the set of objects K; 3: + satisfies the associative law; 4: if x, y and z are in K, and $x \leq y$, then $(x+z) \leq (y+z)$, i.e., "adding" the same element does not alter order; 5: if x and y are in K, and not $x \leq y$, then there is a z in K such that $x \leq y+z$ and $y+z \leq x$, that is, any element may be obtained by "summing" two other elements; 6: if x and y are in K, then not $x+y \leq x$, that is, "sum" is greater than the "summands"; 7: if x and y are in K and $x \leq y$, then there is a number n such that $y \leq nx$ (Archimedean property).

[11] In effect, von Neumann and Morgenstern's axioms concern indifference classes of lotteries rather than single lotteries, and this legitimates their use of the indifference relation. However, this feature of the von Neumann-Morgenstern axiomatic system was made fully clear by Edmond Malinvaud (1952) only after the publication of Suppes' 1951 article.

[12] More formally, ratio-scale measurement is associated with transformations f(x) of the form $f(x)=\alpha x$, where $\alpha>0$; interval-scale measurement is associated with transformations of the form $f(x)=\alpha x+\beta$, where $\alpha>0$; ordinal-scale measurement is associated with any transformation f(x) such that such $f'>0$.

Stevens' measurement theory is not surprising if we consider that Suppes' background was not in psychology but in philosophy, mathematics, and physics. The articles Suppes published between 1951 and 1954 also belonged to these latter disciplines, and more precisely focused on the axiomatic foundations of physics.[13]

2.2. The Stanford Value Theory Project

In the early 1950s two main factors contributed to shift Suppes' research interests toward economics, psychology and behavioral sciences in general. The first one was the influence of J.C.C. "Chen" McKinsey, Suppes' postdoctoral tutor at Stanford. McKinsey (1908-1953) was a logician who since the late 1940s had worked intensively on game theory at the RAND Corporation, a think tank located in Santa Monica, California, and created by the U.S. Air Force in 1946 with the goal of gathering civil scientists from different backgrounds and have them working on interdisciplinary research projects with possible military applications.[14] When McKinsey joined Stanford's Philosophy Department in 1951, he was competing his *Introduction to the Theory of Games* (McKinsey 1952), which became the first textbook in game theory.[15] Suppes' familiarity with game theory and decision analysis was further enhanced by the summer research position he had in early 1950s with David Blackwell and Meyer A. Girshick while they were writing their book *Theory of Games and Statistical Decisions* (1954), in which the tools of decision and game theory were employed to evaluate statistical procedures.

Another, more indirect but possibly more powerful factor that contributed to shift Suppes' research interests toward behavioral sciences was funding. In 1953 the Behavioral Science Division of the Ford Foundation awarded a considerable grant to Stanford's Philosophy Department for a study on "Value, Decision and Rationality". This study was later renamed the "Stanford Value Theory Project". Besides the Ford Foundation, in the years 1953-1955 the Department also

---

[13] See in particular McKinsey, Sugar, and Suppes 1953 and McKinsey and Suppes 1953.

[14] More on RAND and its role in the development of post-War War-II economics, in Leonard 2010.

[15] McKinsey had been forced to leave RAND in 1951 because his homosexuality was considered a security risk. More on McKinsey's scientific contributions and personality in Suppes 1979, Lepore 2004, and Burdman Feferman and Ferferman 2004, pp. 141-143. On security policies and attitudes toward homosexuality at RAND in the early 1950s, see Nasar 1998, especially pp. 185-186.

negotiated a series of contracts with the Office of Naval Research and other military agencies for research on decisions involving risk. It should also be noticed that in 1952 the Ford Foundation had appropriated a large grant to create at Stanford University a research center aimed at promoting interdisciplinary work in the social sciences. The Center for Advanced Study in the Behavioral Sciences inaugurated in 1954, and quickly become a focal site for interdisciplinary research in psychology, economics and related disciplines.[16]

In 1953, Suppes and McKinsey began working on the Stanford Value Theory Project, and embarked on the study another philosopher who had joined Stanford University in 1951, namely Donald Davidson. Davidson (1917-2003) is best known today for his influential works in the philosophy of mind and action, the philosophy of language, and epistemology. However, he published these works only from the early 1960s on. In the 1950s, he was very much busy with teaching and did not yet have a clear philosophical project. As he explained in a later interview: "Suppes and McKinsey took me under their wing […] because they thought this guy [i.e., him] really ought to get some stuff out" (Davidson in Lepore 2004).[17] Most of the research connected with the Stanford Value Theory Project was performed between 1953 and 1955, and appeared in print between 1955 and 1957. However, McKinsey contributed only to the first part of the project because in October 1953 he committed suicide.

2.3. Beyond Ratio Measurement

The first installment of the Stanford Value Theory Project was an article in which Davidson, McKinsey and Suppes (1955) discussed the notion of a "rational preference pattern" and the problem of measuring preferences. This article, titled "Outlines of a Formal Theory of Value. I" and published in *Philosophy of Science*, is usually cited in the economic literature for introducing into decision analysis the so-called "money pump argument". This argument supports the claim that preferences must be transitive by showing that an individual with intransitive preferences can be exploited and induced to pay money for nothing.[18] For our

[16] More on the creation of the Center for Advanced Study in the Behavioral Sciences and the funding of Stanford's Philosophy Department in the early 1950s in Solovey 2013 and Isaac 2013.

[17] More on Davidson's research in the 1950s in Isaac 2013 and Lepore 2004.

[18] See Davidson, McKinsey and Suppes 1955, 145-146. For a discussion of the money pump argument, see Anand 1993, and Sugden and Cubitt 2001.

narrative, however, the theory of measurement expounded in the first part of the article is more important.

Davidson, McKinsey and Suppes criticized the traditional identification of measurement with ratio measurement that was still widespread in a number of philosophical treatises of the period. They argued that "this ratio requirement is too rigid", as it also attested by the fact that a number of important physical magnitudes, such as longitude or calendar time, "do not satisfy it" (150). For the authors, the identification of measurement with ratio measurement has also "led to the erroneous view that no kind of measurement appropriate to physics is applicable to psychological phenomena" (151, footnote 8). Without citing Stevens, the authors then introduced a classification of measurement scales that closely resembles that introduced by the Harvard psychologist in 1946. They identified, in order of decreasing strength, the "absolute scale", which does not admit any arbitrary element and was not discussed by Stevens, and then the ratio, interval and ordinal scales, already discussed by Stevens.

Which scale can be applied to the objects of a given domain is to be determined in a precise way only by what Davidson, McKinsey and Suppes called a "coherent theory of measurement" (151). Such a theory should specify "axiomatically conditions imposed on a structure of empirically realizable operations and relations", and prove that "any structure satisfying the axioms is isomorphic to a numerical structure of a given kind." (151)

A couple of comments on the first part of Davidson, McKinsey and Suppes's article are here in order. From the viewpoint of this article, Suppes' 1951 paper on ratio measurement appears as a "coherent theory of measurement" for a particular empirical structure, endowed with specific operations and relations, and isomorphic to a specific numerical structure, namely that identified by proportional transformations. But this kind of exercise can be extended to other structures, operations and relations. In effect, the notion of "coherent theory of measurement" already delineates the approach to measurement that found its full-fledged realization sixteen years later in the *Foundations of Measurement* (1971). Secondly, although Davidson, McKinsey and Suppes's classification of measurement is very similar to Stevens', their insistence on the necessity of stating axiomatically the conditions for measurement cannot be found in the writings of the Harvard psychologist.

2.4. A Non-Orthodox Interpretation of EUT

After expressing their general views on measurement, Davidson, McKinsey and Suppes applied them to the measurement of preferences. They argued that preferences can be measured not only in the sense of an ordinal scale, but that there are also "substantial arguments to support the view that preference can be measured […] in the sense of an interval scale" (151), that is, in terms of a cardinal utility function. In particular they indicated two different systems of axioms concerning a set of riskless alternatives, the lotteries yielding these alternatives, as well as preference and indifference relations over alternatives and lotteries, and showed that both systems imply a cardinal utility representation of the preferences over the alternatives (152-157). Of these two axiomatic systems, here it suffices to say that they are modified versions of the systems put forward in the EUT literature of the early 1950s.

For us it is more important to highlight Davidson, McKinsey and Suppes's stance on the causal explanatory structure of EUT, which is different from the stance that had become standard in utility analysis after the conclusion of the debate on EUT. In the official view, the preferences over (risky) lotteries are primitive, and are not causally derived from the preferences over the (riskless) payoffs of the lotteries, and the fixed probabilities of the payoffs. In the standard view, EUT is about *representing* preferences over risky alternatives through the handy expected-utility formula, rather than *explaining* them on the basis of the more straightforward preferences over riskless alternatives. Davidson, McKinsey and Suppes did not share this official interpretation of EUT, and argued that the utility of lotteries should be causally explained from the utility of their riskless outcomes and, more specifically, from differences in utility between outcomes, rather than vice versa:

> Instead of dealing with probabilities and relative degrees of value [i.e., utility] simultaneously [as when preferences are defined directly over lotteries] it would seem far more natural to determine relative degrees of value first, and *then* modify these by the probabilities to yield decisions in uncertain situations. (1955, 158)

In his second article dealing with utility theory, Suppes investigated in more detail how cardinal utility could be derived from the comparison of utility differences between riskless outcomes.

2.5. Rediscovering Utility Differences

The possibility of obtaining cardinal utility by assuming that individuals are capable of ranking the utility differences between riskless alternatives, had been extensively discussed in the 1930s by Oskar Lange (1934), Henry Phelps Brown (1934), Roy Allen (1935), Franz Alt ([1936] 1971), Paul Samuelson (1938) and other economists. Most utility theorists, however, had remained skeptical about that possibility because the ranking of utility differences has no clear observable counterpart in terms of acts of choice, and therefore relies only introspection, which was not considered a reliable source of evidence (see e.g. Allen 1935, von Neumann and Morgenstern [1944] 1953, 24).[19]

The second installment of the Stanford Value Theory Project was an article co-authored by Suppes and his student Muriel Winet in which they proposed an axiomatization of cardinal utility based on the differences between riskless alternatives. The article, entitled "An Axiomatization of Utility Based on the Notion of Utility Differences", was published in July 1955 in the first volume of *Management Science*, a newly founded management journal that was open to studies in decision theory from different disciplines.[20]

In their work, Suppes and Winet (1955, 259) mentioned that the notion of utility differences had been discussed in economics, and cited Lange's 1934 article on the topic. They also took stance against the economists' opposition to introspection that, since the mid-1930s, had played a crucial role in the marginalization of utility differences in economic analysis. They claimed that in many areas of economic theory "there is little reason to be ashamed of direct appeals to introspection", and that there are sound arguments for justifying "the determination of utility differences by introspective methods" (261). They then affirmed that, despite the importance and legitimacy of utility differences, to the best of their knowledge "no adequate axiomatization for this difference notion has yet been given." (259) Evidently, they ignored that the Viennese

[19] On the 1930s debate on utility differences and cardinal utility, see Moscati 2013.

[20] In 1957 Winet completed her Ph.D. at Stanford with a dissertation on *Interval Measurement of Subjective Magnitudes with Subliminal Differences* (Wood Winet Gerlach 1957), a topic strictly related to that of her article with Suppes. Apparently, she did not pursue academic career further and did not published other papers.

mathematician Franz Alt ([1936] 1971) had provided a rigorous, and in fact very similar, axiomatization of cardinal utility almost twenty years earlier.[21]

Like Alt, Suppes and Winet considered two order relations – Q and R – over the elements of an abstract set K. Q is a standard, binary preference relation: xQy means that x is not preferred to y. R is a more tricky quaternary relation concerning "differences" or "intervals" between alternatives: (x,y)R(z,w) means that the interval between x and y is not greater than the interval between z and w. Suppes and Winet imposed eleven axioms on the set K, the relations Q and R. The conditions these axioms impose are very similar to the conditions defined by Alt's: completeness, transitivity and continuity of the two order relations, some form of additivity, and an Archimedean property.[22] Based on their eleven axioms, Suppes and Winet proved the intended representation theorem: the axioms imply the existence of a function u which is cardinal in nature, can be interpreted as a utility function, and is such that the interval between x and y is smaller than the interval between z and w if and only the utility difference between x and y is smaller than the utility difference between z and w (265-270).[23]

Suppes and Winet's article became a standard reference in the utility-theory literature of the period 1955-1960, and was often cited when a cardinal

---

[21] Alt's article had been published in German in the Austrian journal *Zeitschrift für Nationalökonomie*, that in 1936 was directed by Morgenstern. Although Morgenstern, Lange and Samuelson were aware of its existence, Alt's article became known to a larger Anglo-Saxon public only after it was mentioned by Joseph Schumpeter in his *History of Economic Analysis* (1954, 1063). More on Alt, his axiomatization of cardinal utility and the fortunes of his article, in Moscati 2013.

[22] Suppes and Winet's axioms are as follows. Axioms 1-4 require that both Q and R are complete and transitive. Axiom 5 imposes that only the "extension" of the interval between two elements x and y matter, and not the relative order of x and y; thus interval (x,y) is equivalent to interval (y,x). Axioms 6 means that any interval (x,y) can be bisected, i.e., that for any two elements x and y there exists a midpoint element t such that interval (x,t) is equivalent to interval (t,y). Axiom 7 states that, if two elements x and y are indifferent, then one can be substituted to the other without modifying the order relationships among intervals: if xQy, yQx, and (x,z)R(u,v), then (y,z)R(u,v). Axioms 8 requires that, if y is between x and z, then the interval between x and y is smaller than the interval between x and z. Axiom 9 is an additivity assumption: if interval (x,y) is smaller than interval (u,w), and interval (y,z) is smaller than interval (w,v), then the "sum" (x,z) of the two smaller intervals is smaller than the "sum" (u,v) of the two larger intervals. Axiom 10 imposes a continuity property: if interval (x,y) is strictly smaller than interval (u,v), then there is an element t between u and v such that interval (x,y) is still not greater than interval (u,t). Axiom 11 is an Archimedean assumption; it fundamentally states that each interval can be expressed as the sum of a finite sequence of smaller, equivalent intervals.

[23] More formally, xQy if and only if $u(x) \leq u(y)$, and (x,y)R(z,w), if and only if $|u(x)-u(y)| \leq |u(z)-u(w)|$.

representation of utility was associated with the possibility of comparing utility differences.[24] For Suppes, the article represented a further exercise in the elaboration of a "coherent theory of measurement", that is, in defining operations and relations over a certain set of objects (in this case the objects of preference) that make the set isomorphic to a specific numerical structure, which this time was identified by linear increasing transformations.

3. Suppes' Experimental Measurement of Utility

3.1. From Mosteller-Nogee to Davidson-Suppes-Siegel

As mentioned in the introduction, EUT suggests a practicable way to measure experimentally the decision maker's utility function. Imagine that an individual obeys the EUT axioms and is found to be indifferent between lottery [$500, 0.4; $1,000, 0.6], that is, $500 with probability 0.4 and $1,000 with probability 0.6, and $750 for sure. Then we can infer that for him u($750) is equal to 0.4×u($500)+0.6×u($1,000), whereby u is the cardinal function whose existence is implied by the EUT axioms. Since u is unique up to linear transformations of the form $\alpha x+\beta$, two points of it are arbitrary. Thus we can arbitrarily state that u($500)=0 and u($1,000)=1, and establish that for the individual u($750)=0.4×0+0.6×1=0.6.

In their pioneering experiment, Mosteller and Nogee (1951) had used EUT as a measurement device in the way just described and measured the utility function of fifteen individuals on the basis of their preferences between gambles where small amounts of real money could be won or lost. In November 1953, Davidson and Suppes began conceiving an experimental study to measure the utility of money that could improve on Mosteller and Nogee's.

In connection with this project, Davidson and Suppes elaborated an axiomatic model of decision making under uncertainty that they judged more suited for experimental analysis than the one used by Mosteller and Nogee. The model was later presented in an article published in *Econometrica* and titled "A Finitistic Axiomatization of Subjective Probability and Utility" (Davidson and Suppes 1956). Davidson and Suppes adopted the subjective approach to EUT of Ramsey and Savage, but opposed an important aspect of the Ramsey-Savage framework,

---

[24] See e.g. Luce and Raiffa 1957, Debreu 1958, Davidson and Marschak 1959, Arrow 1960, and Chipman 1960.

namely that the set of risky alternatives over which the individual's preferences are defined is infinite. They argued that "no one can ever compare an infinite list of alternatives" (264) and, accordingly, they developed a model that is "finitistic" in the sense that the set of outcomes and their risky combinations is finite.

However, neither Davidson nor Suppes had any experience in experimental investigation and therefore they involved in the project Sidney Sigel, then a Stanford Ph.D. student in psychology. Siegel (1916-1961) had just completed a doctoral dissertation in which he presented a measure of authoritarianism based on experimental techniques (Siegel 1954).[25] The Davidson-Siegel-Suppes experiment was performed in the spring of 1954 and involved nineteen male students at Stanford University who, like Mosteller and Nogee's subjects, had to choose between pair of gambles where small amounts of real money could be won or lost. In particular, in the Davidson-Siegel-Suppes experiment the gambles' payoffs ranged between from –35¢ to +50¢ and, in order to implement the finitistic character of the Davidson-Suppes model, fractions of cents were not allowed.

The results of the experiment were first presented in a 100-page technical report published in August 1955 (Davidson, Siegel, and Suppes 1955), and later in the book *Decision Making: An Experimental Approach* which was the last outcome of the Stanford Value Theory Project (Davidson, Suppes, and Siegel 1957).[26] As Daniel Ellsberg noticed in reviewing the book in the *American Economic Review*, it is not a systematic introduction to decision making, but "a long article, dealing fairly technically with problems connected with this particular set of experiments." (Ellsberg 1958)

Before presenting the utility measures obtained by Davidson, Siegel and Suppes, it is useful to illustrate certain features of their experimental design that

[25] More on Siegel's life and career in the memoir written by his wife Alberta (Engvall Siegel 1964) and in Innocenti 2010.

[26] The book also presents the findings of a second experiment that Davidson and Suppes performed early in 1955 without Siegel's collaboration. The experiment involved seven students in music at Stanford university, and aimed at measuring their utility for LP records of classical music on the basis of their choices between gambles having the records as prizes. The significance of this experiment, however, was marred by the high number of intransitive choices observed by Suppes and Davidson. The intransitivities, which were probably due to the fact most students perceived the LP records as too similar, made tricky to identify even ordinal utility functions for the records. See Davidson, Suppes, and Siegel 1957, 84-103.

respond to philosophical and psychological concerns quite extraneous to the economic theory of the period.

3.2. Gamble vs Gamble

In the Mosteller-Nogee experiment, subjects had to choose between a proper gamble involving the possibility of winning or losing money, and a sure outcome, corresponding to the *status quo* and associated with the refusal of playing the gamble. Davidson, Siegel and Suppes (1955, 61-62; 1957, 55-56) noticed that the possible existence of a specific utility for gambling, may it be positive or negative, could have distorted the utility measures obtained by Mosteller and Nogee. A positive utility for gambling would in fact led to overestimate the utility of the sure amount of money, while a negative utility for gambling would have the opposite effect.

In order to partially overcome this problem, in all decision situations but one Davidson, Siegel and Suppes let their experimental subjects choose between two proper gambles, both involving risk. In this way, they argued, the utilities for gambling associated with the two gambles should cancel out, and therefore the measures of the utility of money obtained from gambe-vs-gamble comparisons should be more precise than those obtained by Mosteller and Nogee.

3.3. ZOJ and ZEJ

Davidson, Siegel and Suppes considered gambles of the form "x cents of dollar if event E occurs, y cents of dollars if event E does not occur" – for brevity, [x¢, E; y¢, not-E]. In particular, they focused on gambles in which the events E and not-E are believed to be equally likely by the experimental subjects, that is, on events that have a subjective probability ∏ equal to one-half: ∏(E)= ∏(not-E)=0.5. Davidson, Siegel and Suppes concentrated on these fifty-fifty gambles because, they argued, experimental subjects understand them better than other gambles. Therefore utility measures obtained from choices over fifty-fifty gambles are more reliable than utility measures obtained from choices over more complex types of gambles like those used by Mosteller and Nogee (17).[27]

---

[27] In the first part of their experiment, Mosteller and Nogee (1951, 380) showed subjects a card with a series of five numbers called a "hand", like in the game of poker dice, e.g. 66431, and asked them whether they were willing to accept bets of the form: "If you roll five dices and beat 66431, you will receive 20 cents; if you do not beat it, you will lose the 5 cents you must risk to play." The experimental subjects were informed

However, how to find two mutually exclusive events that experimental subjects believe to be equally likely?

Davidson, Siegel and Suppes claimed that the events usually assumed to have this feature, such as "heads" or "tails" in tossing a coin, or "even number" or "odd number" in rolling a standard die, did not do. In the pilots of the experiment, in fact, the authors found that subjects seemed to believe one of those events more likely than the other. For instance, if a subject consistently prefers gamble [x¢, heads; –x¢, tails] to gamble [–x¢, heads; x¢, tail], this suggests that he considers heads more probable than tails. To overcame the problem, Davidson, Siegel and Suppes constructed special dies carrying a nonsense syllable, such as ZOJ, on three faces, an another nonsense syllable, such as ZEJ, on the other three faces. They found that subjects believed that the mutually exclusive events "ZOJ comes up when you toss the die" (E) and "ZEJ comes up when you toss the die" (not-E) were equally likely.

Thus the Davidson-Siegel-Suppes subjects participated to series a bets, in each of which they had to choose between pair of gambles of the form [x¢, ZOJ; y¢, ZEJ]. Indifference between gambles was ruled out, i.e., one gamble had to be selected. The ZEJ-ZOJ die was rolled, and the money was paid, only after the subjects had made all them choices between gambles.

3.4. Experimental Findings

Instead of measuring "directly" the utilities of certain amounts of money, Davidson, Siegel and Suppes looked for monetary amounts with certain specified utilities. They began considering amounts of money a=–4¢ and b=6¢, to which they arbitrarily assigned utility values –1 and +1: u(–4¢)=–1 and u(6¢)=+1. Using EUT as a measurement device, they then looked for monetary amounts c, d, f, g such that u(c)=–3, u(d)=+3, u(f)=–5 and u(g)=+5.

Since they did not allow for outcomes consisting of fractions of cents and ruled out indifference between gambles, they could determine the monetary amounts c, d, f and g only by approximation. For instance, they found that for Subject 1 amount c lies between –11¢ and –10¢.[28] In their procedure, the identification of

about the statistical probability that a five-dice rolling had to beat the displayed hand, but this probability was rarely one-half.

[28] Davidson, Siegel and Suppes proceeded as follows. They asked subjects to compare different pairs of gambles until they arrived at identifying two monetary amounts x and y such that (i) [6¢, ZOJ; x, ZEJ] is not preferred to [–4¢, ZOJ; –4¢, ZEJ], (ii) [–

successive monetary amounts relies on the identification of previous amounts, e.g. the identification of f relies on the identification of c. Therefore, the margin of approximation in identifying successive amounts increases, and the bounds within which the "true utility curve" lies widen. All this may result clearer by looking at Figure 1, which presents the estimated bounds for the utility curve for Subject 1 (62).

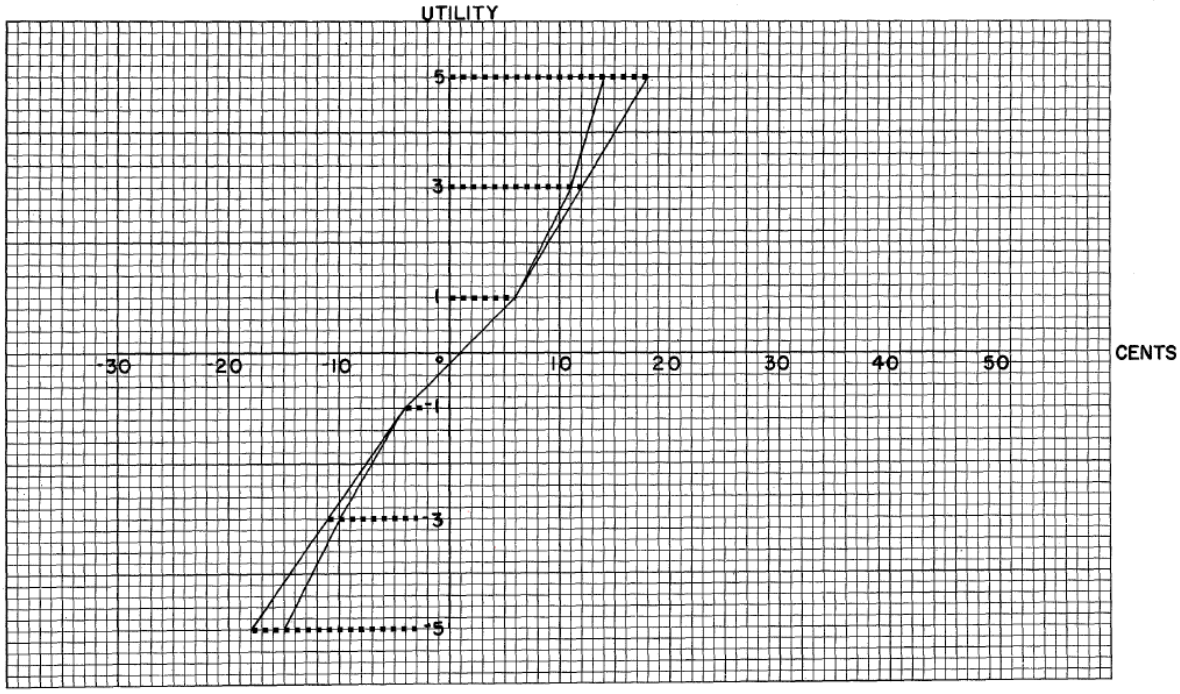

Figure 3. Bounds for the utility curve of Subject 1. *Source*: Davidson, Siegel and Suppes 1957, 63. Horizontal dotted lines were added to make the figure more readable.

The two continuous lines in the figure are the two bounds for the "true" utility curve. They are drawn for the six utility values −5, −3, −1, +1, +3, +5, which are connected by straight lines. As mentioned above, for u(c)=−3 the experimental evidence only allows to say that for Subject 1 the monetary amount c lies between −11¢ and −10¢. For u(f)=−5, the level of approximation increases, and it is only possible to say that for Subject 1 the monetary amount f lies between –

4¢, ZOJ; −4¢, ZEJ] is not preferred [6¢, ZOJ; y, ZEJ]. They found that for Subject 1, x=−11¢ and y=−10¢. If Subject 1 obeys EUT and [6¢, ZOJ; −11¢, ZEJ] is not preferred to [−4¢, ZOJ; −4¢, ZEJ], then u(6¢)+u(−11¢)≤u(−4¢)+u(−4¢). Since by construction u(−4¢)=−1 and u(6¢)=1, u(−11¢)≤−3. If [−4¢, ZOJ; −4¢, ZEJ] is not preferred to [6¢, ZOJ; −10¢, ZEJ], then u(−4¢)+u(−4¢)≤u(6¢)+u(−10¢), that is, −3≤u(−10¢). Since the utility function u is strictly increasing for money, 11¢≤c≤−10¢.

18¢ and −15¢. A similar analysis hold for monetary amounts d and g of the positive range of the utility curve.[29]

Davidson, Siegel and Suppes managed to measure in this approximate way the utility curves of fifteen of the fifteen subjects.[30] Their main experimental findings can be summarized as follows (see Davidson, Siegel and Suppes 1957, 62-72):

i) None of the fifteen subjects was uniformly risk adverse or risk loving; accordingly, none of the fifteen utility curves was convex or concave through its entire length;
ii) Ten subjects were risk loving about wins, and risk adverse with respect to losses; therefore, their utility curves displayed a trend similar to the curve of Subject 1's, i.e., concave for wins and convex for losses;
iii) Two subjects had the opposite attitude: they were risk adverse for wins and risk loving for losses; accordingly, their utility curves were convex for wins and concave for losses;
iv) The remaining three subjects appeared to be risk neutral; thus their utility curves were fundamentally linear through their length.

These findings do not hint at any clear empirical restriction on the shape of the subjects' utility functions and the corresponding attitudes toward risk. Therefore from an economic viewpoint they do not appear very useful. But the main goal of Davidson, Siegel and Suppes was not to determine the "typical" attitude toward risk of individuals, but to show that that "for some individuals and under appropriate circumstances it is possible to measure utility" (Davidson, Suppes, and Siegel 1957, 19). With respect to this goal, the experiment was not yet completed.

3.5. Verifying the Utility Measures

Davidson, Siegel and Suppes noticed that calculating and drawing a subject's utility function does not have, by itself, any clear theoretical significance:

---

[29] In particular, for Subject 1 11¢≤d≤12¢ and 14¢≤g≤18¢.

[30] For four subjects it was not possible to indentify, not even in an approximate way, monetary amounts x¢ whose utility was −5, −3, +3 or +5.

> The present model shares a defect common to many psychometric techniques of scaling, namely, no matter what responses the subject makes to the option presented him, a numerical utility function can be computed. (86)

For them the significance of an elicited utility function can be assessed only by testing its "predictive power" (86). Mosteller and Nogee had performed this test by using the utility functions elicited from choices among a first type of gamble to predict choices among a second and much more complex type of gamble.[31] In contrast, Davidson, Siegel and Suppes run a second experiment that took place after a period varying from a few days to several weeks after the first experiment. In the second experiment, they asked ten of the fifteen original subjects to choose among the same type of fifty-fifty gambles the subjects had faced in the first experiment. Using the approximation technique used in the first experiment, Davidson, Siegel and Suppes re-elicited the utility curve of each subject, and checked whether it was alike to the utility curve elicited in the first experiment. They found that for nine of the ten subjects the utility curves elicited in the two experiments were in fact very similar.[32]

3.6. The Interpretation of the Function u

A significant part of the debate on EUT whose outcomes have been summarized in the introduction concerned the interpretation of the function u. Can we really interpret the function u, which is elicited from preferences over risky alternatives, as expressing the individual's utility of riskless alternatives, in this case of riskless money? If we find that for a given individual the function u is, say, convex for wins, can we conclude that for this individual the marginal utility of monetary gains is decreasing?

The official answer to these questions that had emerged in utility theory during the debate on EUT was in the negative: the function u does not express the utility of riskless money or other riskless alternatives, and the curvature of u

[31] In the second part of their experiment, Mosteller and Nogee (1951, 382) asked experimental subjects to accept or reject bets having the following form: "You have the opportunity of betting or not betting 5 cents against this double offer: if you beat 22263 you will receive 20 cents; if you do not beat 22263 but do beat 66431, you will receive 3 cents; if you do not beat either, you will lose the 5 cents you must risk to play."

[32] Davidson, Siegel and Suppes (1957, 69) argued that the behavior of the remaining subject could be explained by the fact that he was a foreign student with some language difficulty.

cannot be used to determine whether the marginal utility of the riskless alternatives is increasing or decreasing. In the official view, the only economic meaning of the function u is that it can be used to predict choices.[33] For instance, if using EUT as a measurement device it is found that for an individual u($500)=0, u($750)=0.6, and u($2,000)=2.5, then he should prefer $750 for sure to lottery [$2,000, 0.2; $500, 0.8].[34]

The problem with this official view is that it obliterates any psychological meaning of the function u. The function u, as von Neumann and Morgenstern ([1944] 1953, 28) had written, reduces to be "that thing for which the calculus of mathematical expectations is legitimate". This was not the interpretation of Davidson, Siegel and Suppes. They saw utility as something that has a psychological reality and can be actually measured, at least "for some individuals and under appropriate circumstances".

It is true that in the second phase of their experiment Davidson, Siegel and Suppes checked for the predictive power of the utility measures obtained in the first phase. However, the second phase was not aimed at predicting new choices, but at verifying the validity and stability of the utility measures obtained in the first phase. This is why, contrary to Mosteller and Nogee, Davidson, Siegel and Suppes asked the subjects to choose among the same type of fifty-fifty lotteries in the two phases of their experiment.

3.7. Reception

The book by Davidson, Siegel and Suppes was positively reviewed by statistician Jack Kiefer (1959) in *Econometrica*. In the *American Economic Review*, Ellsberg (1958) praised Davidson, Siegel and Suppes for their effort to turn the utility notion into a genuine empirical variable and actually measure it. However, Ellsberg doubted that the utility measures the three researchers had obtained for monetary amounts ranging from –35¢ to +50¢ could be extrapolated to the larger monetary amounts "that would really be interesting to an economist" (1958, 1010). More generally, Ellsberg argued that, although the subjective expected-utility model used by Davidson, Siegel and Suppes can explain choice

[33] The official interpretation of the function u was clearly articulated by Friedman and Savage 1952, Ellsberg 1954, Savage [1954] 1971, and Luce and Raiffa 1957.

[34] The expected utility of [$2,000, 0.2; $500, 0.8] is 0.2×2,5+0.8×0=0.5, which is lower than u($750)=0.6.

behavior in some uncertain situations, "there are important classes of uncertain situations in which normal people will systematically *violate* [it]." (1010) [35]

In their widely read introduction to game and decision theory, Duncan Luce and Howard Raiffa (1957, 35) referred to the experimental study of Davidson, Siegel and Suppes as "the most elegant in the area" and considered its results as "very encouraging", but did not illustrate its details. In a subsequent review published in the *Journal of Philosophy*, Luce (1959b, 174) declared that the data obtained by Davidson, Siegel and Suppes were "probably the most satisfactory" in the field. However, he argued that these data, though "encouraging", were "still sufficiently fuzzy to be inconclusive." Moreover, the theoretical model underlying the Davidson-Siegel-Suppes experiment appeared to Luce "so special and complex" that it could hardly capture "the attention of non-specialists."

In a review article on choice theory published in *Econometrica*, Kenneth Arrow (1958, 12), who was a colleague of Suppes at Stanford, judged the results of the Mosteller-Nogee and Davidson-Siegel-Suppes experiments as generally consistent with EUT, but noticed that these results "can hardly be said to confirm it [EUT] in all the detailed applications one would like to make." The Davidson-Siegel-Suppes experiment was cited also by, among others, Davidson and Marschak (1959), Debreu (1960), and Marschak (1964).

These reviews and citations suggest that the experimental measurements of utility performed by Davidson, Siegel and Suppes were known to economists working in decision analysis. However, as noted by Ellsberg, Arrow, and Luce, those measurements seemed to suffer from what today we would call an "external validity problem": they appeared to be hardy extendable to situations different to the very specific one designed by Davidson, Siegel and Suppes, e.g., to situations involving larger amounts of money or risky alternatives not restricted to simple fifty-fifty gambles.

[35] In his review Ellsberg did not illustrate which uncertain situations he had in mind but, with insight, one cannot avoid thinking of the ambiguity situations he presented in his celebrated article on "Risk, Ambiguity, and the Savage Axioms" (Ellsberg 1961).

4. Luce's Psychologically Inspired Utility Analysis

4.1. From Mathematics to Utility Theory and Psychology

Duncan Luce (1925-2012) enrolled in the degree of aeronautical engineering at MIT in 1942, and graduated in 1945 after a period in the Navy during the war. In 1946 he returned to MIT for a Ph.D. in mathematics. During his graduate studies, Luce became interested in social psychology and the analysis of social networks. However, since no one in MIT's Department of Mathematics was interested in those topics and MIT did not have a psychology department, he wrote his doctoral dissertation on the mathematical theory of groups, graduating in 1950. From 1950 to 1953, Luce worked at MIT's Small Groups Laboratory, a research center devoted to the analysis of the interaction and communication between small groups of people. During his last year at the Laboratory, he began studying game theory as a possible tool for modeling group interactions.[36]

In 1953 Luce moved to Columbia University as an assistant professor of mathematical statistics and sociology. At Columbia Luce was involved in the Behavioral Models Project, which was funded by the Office of Naval Research and had among its goals the preparation of expositions of various applications of mathematics in behavioral science. To the project participated members of different departments at Columbia, such as Ernest Nagel from philosophy, Paul Lazersfeld from psychology, William Vickerey from economics, and Howard Raiffa from mathematical statistics. Raiffa, who served as the Project's chairman, had studied mathematics and statistics at the University of Michigan, and was familiar with game theory and mathematical psychology.[37] Within the activities of the Behavioral Models Project, Luce and Raiffa agreed to write together a short exposition of game theory, that in fact expanded into a 500-page book that was written between 1954 and 1956 and published in 1957 under the title *Games and Decisions* (Luce and Raiffa 1957).

---

[36] This reconstruction of Luce's studies and early career is based on Luce 1989.

[37] Raiffa (born 1924) received a B.Sc. in mathematics (1946), a M.Sc. in statistics (1947), and a Ph.D. in mathematics (1951), all from the University of Michigan. As a graduate student, he worked as a research assistant in a project sponsored by the Office of Naval Research (ONR) in the course of which he became interested in game theory. In particular, Raiffa wrote for the ONR an extended research report on non-zero sum games that eventually became his doctoral dissertation. After graduation and before moving to Columbia, Raiffa worked with Clyde Coombs at Michigan University on mathematical psychology and the theory of psychological measurement. More on Raiffa's life and career in Fienberg 2008.

Luce spent the academic year 1954-1955 at Stanford as a fellow at the Centre for Advanced Study in Behavioral Science. At the Centre, Luce drafted portions of *Games and Decisions* and made the acquaintance of Suppes, but this first encounter of theirs did not generate any common research project. More importantly, during that year at the Centre, Luce's research interests shifted away from social psychology and game theory and toward decision analysis: "I had become fascinated with von Neumann-Morgenstern theory of expected utility, with the Weber-Fechner problem of psychophysical scaling, and with their relation, if any." (Luce 1989, 249)

By "problem of psychophysical scaling", Luce refers to the attempts made since the 19th century by Ernst Weber, Gustav Fechner and other psychologists to measure sensations, such as the sensations of heaviness, brightness or loudness. In particular, Fechner took as the unit to measure sensations the "just-perceivable difference" of sensation, that is, the minimal discernible difference of sensation generated by a change in the physical stimulus.[38] The first product of Luce's new research interests was an article on "Semiorders and a Theory of Utility Discrimination", written when he was at the Centre for Advanced Study in Behavioral Science and published in the April 1956 issue of *Econometrica* (Luce 1956). In it Luce explored how the psychophysical notion of "just-perceivable difference" of sensation could be introduced into utility analysis.

4.2. Semiorders

Luce noticed that, if considered from a psychophysical perspective, the economic notion of "preference" is a specific type of sensation, which is generated by certain physical stimuli such as consumption goods, money, or gambles. The possible existence of "just-perceivable difference" of preference, however, is problematic for economic theory because it undermines the standard assumption that the indifference relation is transitive. To illustrate the issue, Luce (1956, 179) imagined a series of 401 cups of coffee, whereby the first cup contains 1 gram of sugar, the second cup 1.01 gram of sugar, the third cup 1.02 gram of sugar, etc. If we continue to add 0.01 gram of sugar to each subsequent cup in the series, the 401st cup will contain 5 grams of sugar. It appears plausible that an individual is indifferent between any two adjacent cups in the series (since

[38] On Fechner's attempts to measure sensations and the debate they generated, see Michell 1999.

they differ by only 0.01 gram of sugar), but is not indifferent between the first cup (1 gram) and the 401st cup (5 grams). In this case, the indifference relation is not transitive.

As Luce observed, this problem had been already pointed out by the Cambridge social anthropologist and economist Wallace E. Armstrong in a series of articles published in the *Economic Journal* and the *Oxford Economic Papers* since the late 1930s (Armstrong 1939, 1948, 1950, 1951). But while Armstrong's attitude was almost exclusively critical of standard utility analysis, Luce attempted to explore in an axiomatic way what kind of utility theory could be constructed by allowing for intransitive indifference relations. The "semiorders" in the title of his paper are in fact ordering structures in which the strict preference relation P is transitive while the indifference relation I is not.[39]

Luce showed how to construct a utility function u that assigns numbers to alternatives that are only semi-ordered. Without going into the details of the axiomatic model, we may mention here that an important role in Luce's construction is held by two functions (called "upper" and "lower" just noticeable difference functions) which characterize the change in utility necessary for indifference to become preference. In the last section of his article, Luce (188-190) considered the case in which the semi-ordered alternatives are lotteries, and defined some conditions under which a cardinal utility function representing the semiorder exists.

4.3 Utility Measurement in Games and Decisions

When in 1955 Luce returned to Columbia after his sabbatical at the Centre for Advanced Study in Behavioral Science, Raiffa left for the Centre and spent there the academic year 1955-56. Thus their co-authoring of *Games and Decisions* continued at distance. According to Raiffa, during the two years they worked on the project they were "face to face for five days" in total (Fienberg 2008, 142). For Luce, the distant collaboration in effect contributed to get the book finished: "I have always felt that we would never have written it had we been together,

[39] More precisely, let S be a set with elements a, b, c,… and P and I two binary relations defined over S. Luce (1956, 181) defined (P, I) as a semiorder over S, if for every elements a, b, c and d in S, the following axioms hold. Axiom 1: aPb, or bPa, or aIb (the semiorder is complete). Axiom 2: aIa (I is reflexive). Axiom 3: aPb, bIc, cPd, imply aPd. Axiom 4: aPb, bPc, bId imply not both aId and cId. Axiom 3 and 4 fundamentally require that an indifference interval never spans a preference interval. Axioms 1-4 imply that P is transitive but do not imply that I is.

because it would have been too easy to talk" (Luce 1989, 268). *Games and Decisions* was published in the second half of 1957, and is dedicated to the memory of von Neumann who had died in February of that year. Thanks to its clarity and the candid account of open problems it provides, the book quickly became a key reference for scholars working in game and decision theory and is still in print. Most of the book deals with game theory, but for our concerns the most important part of it is chapter 2 on "Utility Theory".

In that chapter Luce and Raiffa (1957, 23-32) presented their own axiomatization of expected utility theory and discussed some possible fallacies in the interpretation of the theory. In particular, they criticized the view – labeled as "Fallacy 1" – according to which alternative a is preferred to alternative b *because* the utility or the expected utility of a is larger than that of b. For Luce and Raiffa preferences come first, and the utility function is only "a convenient way *to represent* them." (32) They also criticized as "Fallacy 3" the view according to which the cardinal utility function drawing from the axioms of expected utility theory allows two compare utility differences between riskless alternatives. Luce and Raiffa did not oppose to the construction of a utility theory allowing for comparison of utility differences like that put forward by Suppes and Winet (1955). But they emphasized that expected utility theory "does not permit such comparisons" (32).

In the final part of the chapter on utility theory, Luce and Raiffa explicitly discussed the problems associated with the experimental measurement of utility. They admitted that "even under the most ideal and idealized experimental conditions" (36), it is extremely difficult to measure a person's utility function. With respect to the possibility of measuring utility outside the laboratory, they were even more pessimistic: "There is certainly no hope at all that it can be done under field conditions for situations of practical interest" (36). But if this the case, Luce and Raiffa asked, why had utility theorists devoted so many energies to discuss issues related to utility measurement and to attempt at measuring utility?

> If the theories built upon utility theory really demand such [utility] measurements, they are doomed practically; if they can be useful without making such measurement, then why go to the trouble of learning how? (36)

Luce and Raiffa's answer was threefold. First, they argued that even if utility were not measurable, utility theory could nonetheless provide useful conclusions: "As in the physical sciences, we would claim that a theory may very well postulate quantities which cannot be measured […], and yet that it will be possible to derive some conclusions from them which are of use" (36). If utility measurements could be made, the conclusions of utility theory would be richer, but this does not mean that "if the measurements cannot be made, nothing can be concluded." (36). Second, the attempts to measure utility in the idealized conditions of the laboratory are useful to test utility theory and, possibly, to ameliorate it:

> The main purpose [of laboratory measurements of utility] is to see if under any conditions, however limited, the postulates of the model can be confirmed and, if not, too see how they may be modified to accord better at least with those cases. (37)

Finally, if utility could be measured is some laboratory experiment, utility theorists would "fell less cavalier" (37) in postulating the existence of all the constructs required by utility analysis.

4.4. From Thurstone to Probabilistic Choice

One important post-Fechner development in the history of psychophysics was the probabilistic approach to measurement of sensations put forward by the American psychologist Louis Leon Thurstone in the 1920s and known as the method of "comparative judgment". In Thurstone's method, a single subject is confronted with pairs of stimuli, e.g. pairs of lights, and asked to rank them with respect to some dimension, e.g. brightness. Because of judgment errors, distraction, or variation in the senses' sensibility, the subject's ranking "is not fixed. It fluctuates." (Thurstone 1927, 274) This means that comparative judgment is in fact a random variable. As a consequence, when the subject is confronted more than once with the same pair of stimuli a and b, sometimes he will rank a over b, and sometimes b over a. Thurstone used the frequency with which a stimulus is ranked over another to rank the stimuli themselves: if light a is perceived to be brighter than light b more than fifty per cent of time, a is taken to be brighter than b. Based on the frequencies of comparative judgments on multiple pairs of stimuli, Thurstone claimed that it is possible to identify a unit of measurement for brightness and other sensations and measure them accordingly. Thurstone's claim, however, depends on a number of demanding statistical assumptions about the probabilistic process generating the comparative judgments.[40]

Beginning in the mid-1950s, Marschak and other economists applied Thurstone's probabilistic approach to utility theory.[41] In the probabilistic approach, a subject is said to prefer alternative a to alternative b if the probability that he chooses a over b – indicated as P(a,b) – is at least 0.5, i.e. if $P(a,b)\geq 0.5$. If P satisfies certain properties, then there exists a utility function u that represents the probabilistic preferences, in the sense that $u(a)\geq u(b)$ if and only if $P(a,b)\geq 0.5$. If further conditions such as the "strong stochastic transitivity" of P hold, then the utility function u is cardinal in nature, and the probability that a subject chooses an alternative over another is associated with the utility differences between the two alternatives, that is, $P(a,b)\geq P(c,d)$ if and only if $u(a)-u(b)\geq u(c)-u(d)$.[42]

4.5. Luce's Theory of Probabilistic Choice

While working on *Games and Decisions*, Luce became interested in the theory of probabilistic choice and began to elaborate his own version of it. Possibly during the winter of 1956-57, he conceived of the assumption that became the centerpiece of his theory, namely what he called the "Choice Axiom". As we will see in a moment, the Choice Axiom implies a number of significant restrictions on choice behavior. Building on the Axiom, in spring 1957 Luce developed his theory of probabilistic choice into a 100-page report that was bound with a red cover and distributed to some colleagues. In summer 1957, Suppes organized at Stanford University a six-week workshop on formal approaches to behavioral and social sciences, to which also Luce participated. Luce's report, or the "red menace" as it quickly began to be called, became a major focus of discussion at the workshop. Initially Suppes thought that the Choice Axiom was wrong, but then Luce succeeded in persuading him otherwise. The controversy over the Choice Axiom marked the beginning of the scientific collaboration and friendship between Suppes and Luce (Luce 1989, 250; Suppes 1979, 50).

---

[40] More on Thurstone's method of comparative judgments in Michell 1999.

[41] See in particular, Marschak 1955 and 1960, Quandt 1956, Georgescu-Roegen 1958, Debreu 1958 and 1960, Davidson and Marschak 1959, Block and Marschak 1960.

[42] Strong stochastic transitivity requires that if $P(a,b)\geq 0.5$ and $P(b,c)\geq 0.5$, then $P(a,c)\geq \max\,[P(a,b), P(b,c)]$, see e.g. Davidson and Marschak 1959, 240.

In the academic year 1957-58, Luce worked up his red-menace report. A first presentation of his theory, which however did not center on the Choice Axiom, was published in an *Econometrica* article entitled "A Probabilistic Theory of Utility" (Luce 1958). The revised report expounding the full-fledged theory was published the following year as a slim book that, predictably, had a red jacket and bore the title *Individual Choice Behavior. A Theoretical Analysis* (Luce 1959). Among others, Luce thanked Suppes, Marschak, Mosteller, and Stevens for critical comments on previous drafts of the work.

Luce's theory of probabilistic choice distinguishes from the theories put forward by Marschak and other economists because it draws almost completely from the Choice Axiom. In a simplified form, the Axiom states that, if x is an element of set R, which is in turn a subset of set S, than the probability $P_S(x)$ of choosing x from S, is equal to the probability $P_S(R)$ of choosing R from S times the probability $P_R(x)$ of choosing x from R. Formally, if $x \in R \subseteq S$, then $P_S(x)=P_S(R)P_R(x)$. For example, if a restaurant has a menu with three dishes – roast beef, steak, and fish – then the Choice Axiom states that the probability of choosing steak from the entire menu is equal to the probability of choosing a meat dish from the entire menu, times the probability of choosing steak over roast beef.[43]

In *Individual Choice Behavior* Luce explored the numerous and often surprising implications of the Choice Axiom in areas as different as psychophysics, the theory of learning and utility theory. In particular, Luce (34-36) suggested that the concept of just-noticeable difference of preferences, which is non-probabilistic in nature, can be easily obtained from a probabilistic model. It sufficient to define two alternatives as "different" if the probability of choosing one over the other exceeds some cutoff, e.g. if $P(x,y)>0.65$. Otherwise the two alternatives are not noticeably different, i.e., indifferent. In particular, Luce showed that, if the probabilistic preference relation P satisfies the Choice Axiom, then it generates a semiorder to which all results of his 1956 article can be applied.

[43] As Luce (1959, 7-11) pointed out, the Choice Axiom can be seen as an extension of the properties of conditional statistical probabilities to choice probabilities, or as a probabilistic version of Arrow's axiom of independence from irrelevant alternatives (Arrow 1951).

Furthermore, Luce (36-42) showed that the Choice Axiom implies that the choice probability P satisfies the strong stochastic transitivity and the other conditions that warrant the existence of a cardinal utility function that (i) represents the probabilistic preferences and (ii) is such that P(a,b)≥P(c,d) if and only if u(a)−u(b)≥u(c)−u(d). In other words, Luce showed that the probabilistic theories of choices that Marschak and others had derived from a larger set of axioms could be obtained from the Choice Axiom alone. Thus, instead of discussing whether strong stochastic transitivity and other axioms are descriptively or normatively acceptable, one could confine the discussion to the acceptability of the Choice Axiom alone.

4.6. Debreu against the Choice Axiom

The soundness of the Choice Axiom and the theory of choice Luce had built on it became more problematic when Debreu pointed out an implausible implication of the Axiom. In the review of *Individual Choice Behavior* published in the *American Economic Review*, Debreu (1960b) noticed that when some alternatives are more similar than others the Choice Axiom may be violated. The point can be illustrated using the restaurant example introduced above.

Imagine that, in comparing the three dishes on a pairwise basis the individual is probabilistically indifferent between them. P(roast beef, fish)=P(steak, fish)=½ because, say, the individual is indifferent between meat and fish. P(roast beef, steak) =½ because the individual is indifferent between the two types of meat. In this situation, the Choice Axiom implies that the probability of choosing each single dish from entire menu is equal to $^1/_3$. However, if the fundamental decision of the individual is whether to eat meat or fish, we would expect that the probability of choosing fish from the entire menu is still close to ½, while the probability of choosing roast beef (or steak) from the entire menu is around ¼.[44]

In a later essay written with Suppes (Luce and Suppes 1965, 337), Luce acknowledged that Debreu's argument was "convincing", and that the Choice Axiom does not hold for all decision situations. However, Luce stressed that

[44] In his review, Debreu (1960, 188) did not use dishes as examples but three recordings of classical music: a recording of the Debussy quarter, a recording of the eight symphony of Beethoven by a certain conductor, and a recording of the eight symphony of Beethoven by another conductor. A analogous point – based on a pony, a bicycle, and another bicycle endowed with a speedometer – was independently made by Savage in private correspondence with Luce; see Luce and Suppes 1965, 334.

Debreu's criticism applies not only to the Choice Axiom and the probabilistic theory of choice built on it, but to "all our current preference theories", that is, also to the more orthodox ones. In fact, Luce argued, Debreu criticism suggests that standard preference theories may easily fail if they do not include "some mathematical structure over the set of outcomes" that allow us to characterize outcomes that are more or less substitutable for one another (like roast beef and steak in the restaurant example), from outcomes that are significantly different (such as steak and fish) or are special cases of others. Thus, Debreu's argument points in fact at un-orthodox extensions of the economic theory of choice, which however neither he nor other economists pursued in the 1960s.

5. Conclusions [only a sketch]

Zenith of the interdisciplinary research by economists, psychologists, and philosophers, centered on EUT and cardinal utility: Stanford symposium on Mathematical Methods in the Social Sciences, 15-24 June, 1959. Proceedings edited by Arrow, Suppes, & Karlin (1960).

Papers by Arrow, Debreu, Marschak, Suppes, Luce.

1960-1970: decline of interdisciplinary exchanges.

Suppes and Luce: fewer or no publications in economics journals.

Explanation, Luce & Suppes, 1965, *Handbook of Mathematical Psychology* : "Psychological theories of preference have begun to acquire a richness and complexity that renders them largely useless as bases for economic theories." (253).

Impact of Suppes and Luce's work: not insignificant, but limited

Source: JSTOR business & economics, 2099 titles, 1960-1980, without self-citations:

Work	Davidson, McKinsey & Suppes 1955	Suppes & Winet 1955	Suppes & Davidson 1956	Davidson, Suppes, Siegel 1957	Luce 1956	Luce 1958	Luce 1959	**Total**
Topic	Formal theory of value	Utility differences	Finitistic axiomat.	*Decision making*	Semiorders	Probab. theory	*Individual choice behavior*	
# citations	**4**	**13**	**2**	**25**	**19**	**5**	**42**	**110**

Rise in citations after 1980 associated with rise of behavioral economics.

Conclusions: Suppes and Luce contributed to the economic research on decision analysis and brought into it some non-orthodox elements from philosophy and psychology:

- causal structure: from riskless to risky preferences
- justification of introspection and utility differences
- psychological concerns in experimental design
- imperfect discrimination and probabilistic choice from psychology

Suppes and Luces's work had some impact on economic decision theory but never became "orthodox" (e.g. no citations in Kreps 1988 or Mas-Colell 1995).
Reason: methodological divide: psychological theories "too complex for economic use".
This holds also for the rest of the interdisciplinary contributions of the 1955-1965 period.
After 1980: rise in citations for Suppes and Luce associated with rise of behavioral economics (e.g. Wakker 2010).

REFERENCES

Alchian, A. 1953. The Meaning of Utility Measurement. *American Economic Review* 43: 26–50.

Allen, R.G.D. 1935. A Note on the Determinateness of the Utility Function. *Review of Economic Studies* 2: 155–158.

Alt, F. 1936. Über die Messbarkeit des Nutzens. *Zeitschrift für Nationalökonomie*, 7: 161–169. English translation in J.S. Chipman et al. (Eds), *Preference, Utility and Demand*. New York: Harcourt Brace Jovanovich, 1971, pp. 424-431.

Anand, P. 1993. The Philosophy of Intransitive Preference. *Economic Journal* 103: 337-346.

Armstrong, W.E. 1939. The Determinateness of the Utility Function. *Economic Journal* 49: 453-467.

Armstrong, W.E. 1948. Uncertainty and the Utility Function. *Economic Journal* 58: 1-10.

Armstrong, W.E. 1950. A Note on the Theory of Consumer's Behavior. Oxford Economic Papers 2: 119-122.

Armstrong, W.E. 1951. Utility and the Theory of Welfare. *Oxford Economics Papers* 3: 259-271.

Arrow, K.J. 1951. *Social Choice and Individual Values*. New York: Wiley.

Arrow, K.J. 1958. Utilities, Attitudes, Choices: A Review Note. *Econometrica* 26: 1-23.

Arrow, K.J. 1960. The Work of Ragnar Frisch, Econometrician. *Econometrica* 28: 175-192.

Arrow, K.J., S. Karlin, and P. Suppes (eds.). 1960. *Mathematical Methods in the Social Sciences, 1959*. Stanford: Stanford University Press.

Becker, G. M., M. H. DeGroot, and J. Marschak. 1963. An Experimental Study of Some Stochastic Models for Wagers. *Behavioral Science* 8: 199-202.

Becker, G. M., M. H. DeGroot, and J. Marschak. 1964. Measuring Utility by a Single-Response Sequential Method. *Behavioral Science* 9:226–32.

Blackwell, D., and M.A. Girshick. 1954. *Theory of Games and Statistical Decisions*. New York: Wiley.

Block, H. D., and J. Marschak. 1960. Random Orderings and Stochastic Theories of Responses. In *Contributions to Probability and Statistics*, edited by I. Olkin et al., 97–132. Stanford: Stanford University Press.

Bogdan, R. (ed.). 1979. *Patrick Suppes*. Dordrecht: Reidel.

Burdman Feferman, A. and S. Feferman. 2008. *Alfred Tarski. Life and Logic*. New York: Cambridge University Press.

Camerer, C. 1995. Individual Decision making. In *The Handbook of Experimental Economics*, edited by J.H. Kagel and A.E. Roth, 587-703. Princeton: Princeton University press.

Chipman, J.S. 1960. The Foundations of Utility. *Econometrica* 28: 193-224.

Cohen, M.R., and E. Nagel. 1934. *An Introduction to Logic and Scientific Method*. New York: Harcourt Brace.

Coombs, C.H., and S. S. Komorita. 1958. Measuring Utility of Money through Decisions. *American Journal of Psychology* 71: 383-389

Cubitt, R.P., and R. Sugden. 2001. On Money Pumps. *Games and Economic Behavior* 37: 121-160.

Davidson, D. and P. Suppes. 1956. A Finitistic Axiomatization of Subjective Probability and Utility. *Econometrica* 24: 264-275.

Davidson, D., and J. Marschak. 1959. Experimental Tests of a Stochastic Decision Theory. In *Measurement: Definitions and Theories*, edited by C. W. Churchman and P. Ratoosh, 233–69. New York: Wiley.

Davidson, D., J.C.C. McKinsey and P. Suppes. 1955. Outlines of a Formal Theory of Value, I. *Philosophy of Science* 22: 140-160.

Davidson, D., S. Siegel, and P. Suppes. 1955. Some Experiments and Related Theory on the Measurement of Utility and Subjective Probability. Technical Report No. 1, August 15, 1955. Stanford University.

Davidson, D., P. Suppes and S. Siegel. *Decision Making: An Experimental Approach*. Stanford: Stanford University Press.

Debreu, G. 1958. Stochastic Choice and Cardinal Utility. *Econometrica* 26: 440-444.

Debreu, G. 1959. Cardinal Utility for Even-Chance Mixtures of Pairs of Sure Prospects. *Review of Economic Studies* 26: 174-177.

Debreu, G. 1960. Topological methods in cardinal Utility theory. In K.J. Arrow, S. Karlin, and P. Suppes (Eds.), *Mathematical Methods in the Social Sciences, 1959*, pp. 16-26. Stanford: Stanford University Press.

Debreu, G. 1960b. Review of *Individual Choice Behavior* by R.D. Luce. *American Economic Review* 50: 186-188.

Edwards, W. 1955. The prediction of Decisions among Bets. *Journal of Experimental Psychology* 50: 201-214.

Ellsberg, D. 1954. Classic and Current Notions of "Measurable Utility." *Economic Journal*, 64: 528-556.

Ellsberg, D. 1958. Review of *Decision Making* by D. Davidson, P. Suppes, S. Siegel. *American Economic Review* 48: 1009-1011.

Ellsberg, D. 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75:643–69.

Engvall Siegel, A. 1964. Sidney Siegel: A memoir. In S. Messick and A. H. Brayfield (eds), *Decision and Choice. Contribution of Sidney Siegel*, 1-24. New York: McGraw-Hill.

Fienberg, S.E. 2008. The Early Statistical Years: 1947-1967. A Conversation with Howard Raiffa. *Statistical Science* 23. 136-149.

Fishburn, P., and P. Wakker. 1995. The Invention of the Independence Condition for Preferences. *Management Science* 41: 1130-1144.

Friedman, M. 1955. What All is Utility?. *Economic Journal*, 65: 405-409.

Friedman, M., and L.J. Savage. 1948. The Utility Analysis of Choices Involving Risk. *Journal of Political Economy*, 56: 279-304.

Friedman, M., and L.J. Savage. 1952. The Expected-Utility Hypothesis and the Measurability of Utility. *Journal of Political Economy*, 60: 463-474.

Galanter, E. 1962. The Direct Measurement of Utility and Subjective Probability. *American Journal of Psychology* 75: 208-220.

Georgescu-Roegen, N. 1958. Threshold in Choice and the Theory of Demand. *Econometrica* 26: 157-168.

Giocoli, N. 2003. *Modeling Rational Agents*. Cheltenham: Elgar.

Guala, F. 2000. The Logic of Normative Falsification: Rationality and Experiments in Decision Theory. *Journal of Economic Methodology* 7: 59-93,

Herstein, I.N., and J. Milnor. 1953. An Axiomatic Approach to Measurable Utility. *Econometrica* 21: 291-297.

Heukelom, F. 2014. *Behavioral Economics. A History*. New York: Cambridge University Press.

Hölder, O. (1901) 1996. The Axioms of Quantity and the Theory of Measurement, Part I. *Journal of Mathematical Psychology*, 40: 235–252.

Innocenti, A. 2010. How a Psychologist Informed Economics: The Case of Sidney Siegel. *Journal of Economic Psychology* 31: 421–434.

Isaac, J. 2013. Donald Davidson and the Analytic Revolution in American Philosophy, 1940–1970. *Historical Journal* 56: 757-779.

Krantz, D.H, R.D. Luce, P. Suppes, and A. Tversky. 1971. *Foundations of Measurement*. Vol. 1: *Additive and Polynomial Representations*. San Diego: Academic Press.

Kreps, D. 1988. *Notes on the Theory of Choice*. Boulder, CO: Westview Press.

Lange, O. 1934. The Determinateness of the Utility Function. *Review of Economic Studies*, 1: 218-225.

Leonard, R. 2010. *Von Neumann, Morgenstern, and the Creation of Game Theory*. New York: Cambridge University Press.

Lepore E. 2004. Interview with Donald Davidson. In D. Davidson, *Problems of Rationality*, pp. 231–266. Oxford: Oxford University Press.

Luce D.R. 1956. Semiorders and a Theory of Utility Discrimination. *Econometrica* 24: 178-191.

Luce D.R. 1958. A Probabilistic Theory of Utility. *Econometrica* 26: 193-224.

Luce D.R. 1959. *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.

Luce D.R. 1959b. Review of *Decision Making* by D. Davidson, P. Suppes, S. Siegel. *Journal of Philosophy* 56: 173-177.

Luce D.R., 1989. R. Duncan Luce. In G. Lindzey (Ed.) *Psychology in Autobiography*, Vol. 8, pp. 245–289. Stanford: Stanford University Press.

Luce D.R., and H. Raiffa. 1957. *Games and Decisions*. New York: Wiley.

Luce, R. D., and P. Suppes. 1965. Preference, Utility, and Subjective Probability. In vol. 3 of *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush, and E. Galanter, 249–410. New York: Wiley.

Malinvaud, E. 1952. Note on von Neumann-Morgenstern's Strong Independence Axiom. *Econometrica* 20: 679.

Marschak, J. 1950. Rational Behavior, Uncertain Prospects, and Measurable Utility. *Econometrica*, 18: 111-141.

Marschak, J. 1955. Norms and Habits of Decision Making Under Certainty. In *Mathematical Models of Human Behavior*, pp. 45-53. Stamford, Conn.: Dunlap and Associates.

Marschak, J. 1960. Binary-Choice Constraints and Random Utility Indicators. In K.J. Arrow, S. Karlin, and P. Suppes (Eds.), *Mathematical Methods in the Social Sciences, 1959*, pp. 312-329. Stanford: Stanford University Press.

Mas-Colell, A., Whinston, M.D. and Green, J.R. 1995. *Microeconomic Theory*. New York: Oxford University Press.

McKinsey J.C.C., A.C. Sugar, and P. Suppes. 1953. Axiomatic Foundations of Classical Particle Mechanics. *Journal of Rational Mechanics and Analysis* 2: 253-272.

McKinsey J.C.C., and P. Suppes. 1954. Philosophy and the Axiomatic Foundations of Physics. *Proceedings of the Eleventh International Congress of Philosophy*, 6: 49-54.

McKinsey, J.C.C. 1952. *Introduction to the Theory of Games*. New York: McGraw-Hill.

Michell, J. 1999. *Measurement in Psychology*. Cambridge. Cambridge University Press.

Mongin P. 1988. Problèmes de Duhem en théorie de l'utilité espérée. *Fundamenta Scientiae* 9: 289-317.

Mongin P. 2009. Duhemian Themes in Expected Utility Theory. In A. Brenner and J. Gayon (eds.), *French Studies in the Philosophy of Science*, 303-357. Springer: New York.

Mongin P. 2014. Le paradoxe d'Allais: comment lui rendre sa signification perdue. *Revue économique* 65: 743-779.

Moscati 2013. How Cardinal Utility Entered Economic Analysis, 1909-1944. *European Journal of the History of Economic Thought* 20: 906-939.

Moscati 2014. The Debate on the Measurability of von Neumann and Morgenstern's Utility Function, 1944-1954. Paper presented at the seminar "Historical and Philosophical Perspectives on Economics", London School of Economics, November 2014.

Mosteller, F., and P. Nogee. 1951. An Experimental Measurement of Utility. *Journal of Political Economy* 59: 371-404.

Nagel, E. 1931. Measurement. *Erkenntnis* 2: 313–33.

Nasar, S. 1998. *A Beautiful Mind*. New York: Simon and Schuster.

Neumann, J. von, and O. Morgenstern. (1944) 1953. *Theory of Games and Economic Behavior*. 1st edition 1944; 2nd edition 1947; 3rd edition 1953. Princeton, N.J.: Princeton University Press.

Papandreou, A.G., O.H. Sauerlender, O.H. Brownlee, L. Hurwicz and W. Franklin. 1957. A Test of a Stochastic Theory of Choice. *University of California Publications in Economics* 16: 1–18.

Phelps Brown, E.H. 1934. Notes on the Determinateness of the Utility Function, I. *Review of Economic Studies* 2: 66–9.

Quandt, R.E. 1956. A Probabilistic Theory of Consumer Behavior. *Quarterly Journal of Economics* 70: 507-536.

Ramsey F.P. 1931. Truth and Probability. In F.P. Ramsey, *The Foundations of Mathematics and other Logical Essays*, pp. 156-198. New York: Harcourt, Brace and C.

Samuelson, P.A. 1938. The Numerical Representation of Ordered Classifications and the Concept of Utility. *Review of Economic Studies* 6: 65–70.

Samuelson, P.A. 1952. Probability, Utility, and the Independence Axiom. *Econometrica*, 20: 670-678-

Savage, L.J. (1954) 1971. *The Foundations of Statistics*. New York: Dover.

Schumpeter, J.A. 1954. *History of Economic Analysis*. New York: Oxford University Press.

Siegel, S. 1954. Certain Determinants and Correlates of Authoritarianism. *Genetic Psychology Monographs*, 49: 187–254.

Siegel, S. 1956. A Method for Obtaining an Ordered Metric Scale. *Psychometrika* 21: 207-216.

Solovey, M. 2013. *Shaky Foundations: The Politics-Patronage-Social Science Nexus in Cold War America*. New Brunswick, NJ; Rutgers University Press,

Stevens, S.S. 1946. On the Theory of Scales of Measurement. *Science*, 103: 677-680.

Strotz, R.H. 1953. Cardinal Utility. *American Economic Review, Papers and Proceedings*. 43, 384-397

Suppes, P. 1961. Behavioristic Foundations of Utility. *Econometrica* 29: 186-202.

Suppes, P. 1979. Self-profile. R.J. In Bogdan (Ed.), *Patrick Suppes*, pp. 3-56. Dordrecht: Reidel.

Suppes, P. A Set of Independent Axioms for Extensive Quantities. *Portugaliae Mathematica* 10: 163-172.

Suppes, P. and M. Winet. 1955. An Axiomatization of Utility Based on the Notion of Utility Differences. *Management Science* 1: 259-270.

Thurstone, L. 1927. A Law of Comparative Judgment. *Psychological Review* 34: 273-286.

Wakker, P.P. 2010. *Prospect Theory*. Cambridge: Cambridge University Press.

Wood Winet Gerlach, M. 1957. *Interval Measurement of Subjective Magnitudes with Subliminal Differences*. PhD Dissertation, Department of Philosophy, Stanford University.