R.D. Luce: ALGEBRAIC SYSTEMS OF MEASUREMENT

(summarized by J. de Leeuw)

1. Extensive Measurement

The major problem of the foundations of measurement is to find axiomatic systems that permit the construction of homomorphic mappings of a given empirical relational system, which satisfies the axioms, into an appropriate numerical system, which also satisfies the axioms. In these lectures, the numerical system will always be a subset of the set of real numbers, Re. Of course, such empirical relational systems are of scientific interest only if there is at least one interpretation for which the axioms are (approximate) empirical laws. In those cases, the numerical representation summarizes these laws in a way that it is easy both to remember and to make valid deductions. The simplest structure for which measurement-theoretical considerations are possible is the system $\langle A, \succsim \rangle$, where A is an arbitrary set and $\succsim$ is some ordering relation. We shall suppose that $\succsim$ is a weak order, i.e. $\succsim$ satisfies

i) for all $x \in A$, $x \succsim x$ (reflexivity)

ii) for all $x, y, z \in A$, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$ (transitivity)

iii) for all $x, y \in A$, either $x \succsim y$ or $y \succsim x$ or both (connectedness)

In the usual way, define the strict ordering $\succ$ by

$$x \succ y \text{ iff } x \succsim y \text{ and not } (y \succsim x),$$

and the indifference relation $\sim$ by

$$x \sim y \text{ iff } x \succsim y \text{ and } y \succsim x.$$

It is easily shown that $\succ$ is a strict simple order and that $\sim$ is an equivalence relation when $\succsim$ is a weak order. The representation theorem for a weak ordering answers the question: under what conditions is there a homomorphic mapping of A into a subset of Re? To formulate the answer, we need the following definition: a subset B of a set A is called <u>order-dense</u> in A if for all $x, y \in A$ and $\notin B$, there exists an element $b \in B$, such that $x \succsim b \succsim y$. Then the answer is given by the

Cantor-Birkhoff theorem:

Theorem 1.1: <u>Suppose that $\langle A, \succsim \rangle$ is a weak ordered structure. There exists a function f, that maps A into Re monotonically</u>, i.e.,

$$x \succsim y \ \underline{\text{iff}} \ f(x) \geq f(y), \quad \text{for all } x, y \in A,$$

<u>iff A contains a countable order-dense subset.</u>

A sketch of a proof of this representation theorem can be found in Birkhoff (1967, p. 200; it is not quite correct: B must include the end points of all gaps). The most familiar example of a countable order-dense subset is, of course, the set of rational numbers considered as a subset of the reals. The uniqueness theorem for the case under consideration is as follows:

Theorem 1.2: <u>If f and f' are both homomorphisms of $\langle A, \succsim \rangle$ into the reals, then there is a strictly monotonic increasing numerical function ψ such that $f = \psi(f')$, i.e., the representation forms an ordinal scale.</u>

This uniqueness theorem shows that a numerical representation of $\langle A, \succsim \rangle$ has a considerable lack of invariance. Scientifically this is an obvious disadvantage (it renders classical analytic techniques nearly useless) and that makes systems of this simplicity of little interest. In general, however, the data include more information than just a weak ordering of an abstract set. By using this additional information, we attempt to strengthen the invariance of the representation.

Our first example of added structure is the introduction of a binary operation, written as o. The theory of such systems $\langle A, \succsim, o \rangle$ is called extensive measurement. The empirical situation is familiar from and important for physics. We have a set of objects A; these objects can be compared with each other, and they can be concatenated. Examples are length, mass, and time. In the measurement of mass we can put two different objects x and y in the pans of a pan balance (in a vacuum) and establish, by noting which, if either, pan drops, whether $x \succ y$, $y \succ x$ or $x \sim y$. Moreover, we can put two different objects x and y in the same pan and study their combined effect, xoy.

Although the direct significance of extensive measurement for psychology is limited, the mathematics involved is fundamental for all other measurement systems.

In classical theories of extensive measurement, it is assumed that the system is closed under the binary operation o, i.e.

$$\text{if } x,y \in A, \text{ then } xoy,\ yox \in A.$$

In practice, however, unrestricted concatenation causes trouble (it would, for example, result in the ultimate destruction of any pan balance). Observe, moreover, that in probability theory we have $p(A \cup B) = p(A) + p(B)$ when, and only when, A and B are disjoint events. This means that union of disjoint sets is very much like concatenation, but clearly unrestricted concatenation (unions) is not acceptable. To overcome these objections, we add to the system a set (relation) B that formulates the restrictions on concatenation. B is a subset of $A \times A$. Verbally, we interpret $(x,y) \in B$ to mean that x and y can be concatenated.

The axioms for (generalized) extensive measurement are the following:

i) $\langle A, \succcurlyeq \rangle$ is a weakly ordered set

ii) $B \subseteq A \times A$ and $B \neq \emptyset$

iii) $o : B \rightarrow A$

iv) if $(x,y) \in B$ and $x \succcurlyeq x'$ and $y \succcurlyeq y'$, then $(x',y') \in B$

v) if $(x,y) \in B$, $(xoy,z) \in B$, then $(y,z) \in B$ and $(x,yoz) \in B$ and $(xoy)oz \sim xo(yoz)$

vi) if $x \succcurlyeq y$ and $(x,z) \in B$, then $xoz \succcurlyeq yoz$ and $zox \succcurlyeq zoy$

Observe, that, by the third axiom, the system is closed under o iff $B = A \times A$. Axiom iv) forces a certain structure on the system, one that is plausible both for probability and for mass. In the latter case, it says: if two weights don't break the balance, then two lesser weights won't either. Axiom v) asserts that the operation o is associative provided that the relevant elements can be combined at all, and axiom vi) shows that the ordering is compatible with the operation o. A system $\langle A,B,\succcurlyeq,o \rangle$ that satisfies axioms i)-vi) is called a <u>weakly-ordered local semigroup</u>. Such a semigroup is called <u>positive</u> if, in addition to i)-vi), we have

vii) for all $(x,y) \in B$, $xoy \succ x$ and $xoy \succ y$.

It is called solvable if in addition

viii) if $x > y$, then there exists a $z \in A$, with $(y,z) \in B$ and $x \sim y \circ z$.

Finally, in most measurement systems, we need an Archimedean axiom. This axiom is named after a property of the real numbers: if I^+ is the set of positive integers and if α, β are positive reals, then the set $\{n \mid n \in I^+ \text{ and } \beta \geqslant n\alpha\}$ is finite. To use a similar axiom in our context we need the notion of n copies of an element of A. We define nx by induction:

a) $1x = x$,
b) if $(n-1)x \in A$ and $((n-1)x, x) \in B$, then $nx = (n-1)x \circ x$,

and we state the Archimedean axiom as follows:

ix) for all $x, y \in A$, the set $\{n \mid n \in I^+$, nx is defined and $y \geqslant nx\}$ is finite.

Now we can state the representation and uniqueness theorems for extensive measurement with restricted concatenation.

Theorem 1.3: If $\langle A, B, \geqslant, \circ \rangle$ is a positive, solvable, Archimedean, weakly-ordered local semigroup, then there exists a function f from A into Re^+, such that

i) $x \geqslant y$ iff $f(x) \geqslant f(y)$,

ii) if $(x,y) \in B$, then $f(x \circ y) = f(x) + f(y)$,

iii) if f' is another function that satisfies i) and ii), then there exists a positive number α such that for all nonmaximal elements x in A, $f(x) = \alpha f'(x)$. (An element x of A is maximal if $x \geqslant y$ for all $y \in A$.)

Axioms i)-ix) can be classed in two groups. In the first, we have those that are necessary in terms of the representation. They simply follow from the fact that a representation with the properties mentioned in the theorem exists. This group includes axioms i), v), vi), vii) and ix). The second group of non-necessary properties are called structural conditions; they are only sufficient and not necessary for the representation to exist. This group includes the axioms ii), iii), and iv) that describe the structure of the set B, and the solvability axiom viii) that asserts that certain equations can be solved.

Whether or not these axioms can and/or must be tested empirically is a subtle problem. In physics the solvability axiom viii), the positivity axiom vii), and the weak order axiom i) are assumed to hold for idealized measuring instruments and an idealized set of objects. Violations are ascribed to friction of the pan balance and other imperfections of the empirical situation. In psychology, the violations of axiom i) may be more serious, because we do not always have a clear idea of what "ideal" would mean. Intransitivity of preference and, especially, of indifference are common phenomena. One would like to have a suitable statistical model to test hypotheses and to assess the seriousness of these violations, but in the area of fundamental measurement problems no satisfactory statistical procedures are now available.

Theorem 1.3 is a generalization of a classical theorem of Hölder: An Archimedean simply ordered group is isomorphic to a subset of the additive reals. In this case o is assumed to be a closed group operation, i.e. $B = A \times A$, o is associative, and identity and inverses exist. The existence of inverses makes the solvability axiom viii) unnecessary, since $x = y\text{o}(y^{-1}\text{o}x)$. We retain the important "compatibility" axiom vi) and also the Archimedean axiom ix).

A proof of Theorem 1.3 follows these lines: for $x,y \in A$, let $N(x,y)$ be the largest integer for which both $N(x,y)x$ is defined and $y \geq N(x,y)x$. Such an integer exists by the Archimedean axiom. We distinguish two cases. In the first, A has a least element, x_0, relative to the given order $\geq$. It is easily shown that $y \sim N(x_0,y)x_0$ or, in words, y can be exactly reached by concatenating a finite number of copies of x_0. In this case, set $f(y) = N(x_0,y)$, and it can be shown without great difficulty to have the desired representation properties. In the second case, we assume that A has no least element. With x fixed and $y,z \in A$, consider the ratio $N(x,y)/N(x,z)$. The numerator tells how many copies of x are approximately equal to y, and the denominator tells the same thing for z. If we take x smaller and smaller, which is possible since, by hypothesis, there is no least element, the approximations to y and z become better and better. In fact, it can be proved (by standard inequality techniques) that the relevant limit exists, and we define

$$\frac{f(y)}{f(x)} = \lim_{x\downarrow} \frac{N(x,y)}{N(x,z)} .$$

The resulting mapping f is then shown to have the asserted properties, with axiom vi) playing a most important role. An important feature of this constructive proof is the use of a standard series, which consists of the set of integral multiples of a certain "small" element, to approximate other elements. Whenever we use Hölder-type methods of proving representation theorems, such standard series arise. Moreover they provide a practical constructive method for finding numerical representations.

Another practical method to obtain representations from a finite sample of data uses results concerning systems of linear inequalities. If there is an order preserving, additive mapping f of a finite set A into the reals, then for all $x,y,u,v \in A$, we have

$$xoy \succsim uov \quad \text{iff } f(x) + f(y) \geqslant f(u) + f(v).$$

Each inequality in the data structure defines a numerical inequality that is satisfied if the additive representation is valid. Clearly $\langle A, \succsim, o \rangle$ has such a representation in the reals only if the system of inequalities, defined by the ordering in the data structure, has at least one solution. An extensive literature exists on the solution, uniqueness, and algorithmic aspects of the problem of systems of linear inequalities.

2. Qualitative Probability

In probability theory the principal primitive notion is that of an 'event' usually interpreted to be a <u>subset</u> of the <u>universal set</u> or <u>sample space</u> X. To cope effectively with infinite sample spaces, it has proved necessary to restrict the system of events so as not to include all subsets of X. Specifically, we confine ourselves to a non-empty system $\mathcal{E}$ of subsets of X that satisfies the following conditions:

i) if $A \in \mathcal{E}$, then $\bar{A} \in \mathcal{E}$;

ii) if $A,B \in \mathcal{E}$, then $A \cup B$ $\mathcal{E}$.

Such a system is called an <u>algebra of subsets</u>. It follows from i) and ii) and non-emptiness, that $X = A \cup \bar{A} \in \mathcal{E}$ and so by i), $\emptyset = \bar{X} \in \mathcal{E}$; moreover, if $A, B \in \mathcal{E}$, then $A \cap B \in \mathcal{E}$ since $A \cap B = \overline{\bar{A} \cup \bar{B}}$. If the unions of countable collections of events are also events, $\mathcal{E}$ is called a σ-algebra.

A (finitely additive) <u>probability space</u> is defined to be a triple $\langle X, \mathcal{E}, P \rangle$, for which $\mathcal{E}$ is an algebra of subsets of X, P is a measure from $\mathcal{E}$ into Re, i.e., for all $A, B \in \mathcal{E}$,

i) $P(A) \geqslant 0$;

ii) if $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$;

and P is a <u>probability measure</u> in the sense that also

iii) $P(X) = 1$.

This definition of a probability space and the interpretation of probability as a measure is due to Kolmogorov (1933).

The question "what is probability?" has given rise to controversies among frequentists, objectivists, Bayesians, subjectivists, logicists, etcetera. I feel that the question is of no different character from any other measurement question, such as "what is mass?". Indeed, one can imagine equally heated debates over the answer to that question, although they have not actually occurred. Alternatively, perhaps the arguments about probability are misplaced and it, too, should be treated as another problem of fundamental measurement. The controversies are due to the fact that whenever relative frequencies cannot be used, the most common measuring instrument in probability measurement is the all too variable human being.

The formal measurement problem of finding necessary and sufficient conditions for the existence of an order-preserving mapping of a system $\langle X, \mathcal{E}, \succcurlyeq \rangle$ into a probability space $\langle X, \mathcal{E}, P \rangle$ requires the existence of a weak ordering, $\succcurlyeq$, of qualitative probability on $\mathcal{E}$. Some ways in which this weak ordering of events can be obtained give rise to terms such as "subjective" or "intuitive" probability. These terms may prove misleading because they suggest an inherent subjectivism which, in fact, probably only reflects the present state of the art. The ways to assess $\succcurlyeq$ may alter with the development of the science, just as it has with other measurements. At one time the only instrument for comparing the mass of different objects must have been the human being. Gradually, man was replaced by more satisfactory - more consistent, reliable, precise - instruments, such as the pan balance, the carefully subdivided ruler, etc. So far in probability measurement, no really adequate instruments have been devised except when events are highly repeatable or when certain types of arguments based on physical symmetry are possible.

Observe that there already is a fair amount of structure in the system $\langle X, \mathcal{E}, \succsim \rangle$. Besides $\succsim$ being a weak order, we have assumed that $\mathcal{E}$ is an algebra of subsets. We start our axiomatization with de Finetti's (1937) requirements for a <u>qualitative probability structure</u>:

i) $\succsim$ <u>is a weak order over</u> $\mathcal{E}$,

ii) $A \succsim 0$ <u>for all</u> $A \in \mathcal{E}$, <u>and</u> $X \succ \emptyset$,

iii) <u>for all</u> $A,B,C,D \in \mathcal{E}$, <u>if</u> $A \cap B = \emptyset$, $C \cap D = \emptyset$ <u>and</u> $A \sim C$, <u>then</u> $B \succsim D$ <u>iff</u> $A \cup B \succsim C \cup D$.

The conditions are clearly necessary for the existence of the required numerical probability measure, but they are not sufficient. This was proved by Kraft, Pratt and Seidenberg (1959), who constructed the following ingeneous counter-example.

Suppose that X is the five-element set $\{a,b,c,d,e\}$, and $\mathcal{E} = 2^X$ (= set of all subsets of X). We first note that if $\succsim$ is a qualitative probability for which there is a representation, then from

$$\{a\} \succ \{b,d\}$$
$$\{c,d\} \succ \{a,b\}$$
$$\{b,e\} \succ \{a,d\}$$

it follows that

$$\{c,e\} \succ \{a,b,d\}.$$

The proof is very easy. Replace the three inequalities by their numerical analogues in the representation, e.g., $\{a\} \succ \{b,d\}$ by $P(\{a\}) > P(\{b\}) + P(\{d\})$. Add these three inequalities and cancel the same terms from both sides of the resulting inequality.

This yields $P(\{c\}) + P(\{e\}) > P(\{a\}) + P(\{b\}) + P(\{d\})$, from which $\{c,e\} > \{a,b,d\}$ follows.

Now suppose that we have some measure for which these four inequalities hold and for which there is no set $A \in \mathcal{E}$ with $P(A)$ between $P(\{c,e\})$ and $P(\{a,b,d\})$. For example, with $0 < \varepsilon < 1/3$, it is easy to see that the following measure will do:

$$P(\{a\}) = (4 - \varepsilon)/(16 - 3\varepsilon),$$
$$P(\{b\}) = (1 - \varepsilon)/(16 - 3\varepsilon),$$
$$P(\{c\}) = 2/(16 - 3\varepsilon),$$
$$P(\{d\}) = (3 - \varepsilon)/(16 - 3\varepsilon),$$
$$P(\{e\}) = 6/(16 - 3\varepsilon).$$

Since the ordering $\succcurlyeq$ induced by P satisfies the axioms of qualitative probability, so do those of $\succcurlyeq^*$ which is obtained from $\succcurlyeq$ by keeping everything else the same except $\{a,b,d\} \succcurlyeq^* \{c,e\}$. Obviously, $\succcurlyeq^*$ does not have a numerical representation since it violates the above four inequalities.

This makes it clear that more is needed to prove a representation theorem. One of the things that we need is an Archimedean axiom (though it is not enough, since it is satisfied in any finite system such as the Kraft <u>et al.</u> example). To formulate this, we need the following definition:

A sequence of events $A_1, \ldots, A_i, \ldots \in \mathcal{E}$ is called a <u>standard sequence relative to</u> A if there exist $B_i, C_i \in \mathcal{E}$, $i=1, 2, \ldots$, such that

i) $A_1 = B_1$ and $B_1 \sim A$

ii) $B_i \cap C_i = \emptyset$

iii) $B_i \sim A_i$

iv) $C_i \sim A$

v) $A_{i+1} = B_i \cup C_i$.

This inductive definition does not make the system unbounded since, for each A_i, we still have $X \succcurlyeq A_i$. We state the Archimedean axiom as:

iv) <u>For each $A > 0$, any standard sequence relative to A is finite</u>.

We can now continue in one of two quite different ways. The first, due to Scott (1964), is to state necessary and sufficient conditions for the finite case by using the linear inequality technique mentioned in the previous section on extensive measurement. The other, followed by Koopman (1940a, 1940b, 1941), de Finetti (1937), Savage (1954), and Luce (1967), involves simpler sufficient conditions but includes a rather strong existence (solvability) axiom. The first three authors postulated that there are partitions of X into arbitrarily many equiprobable events. The latter used instead:

v) <u>for all</u> $A,B,C,D \in \mathcal{E}$, <u>if</u> $A \cap B = \emptyset$, $A > C$ <u>and</u> $B \geqslant D$, <u>then there exist</u> $C',D',E \in \mathcal{E}$, <u>such that</u>

   a) $E \sim A \cup B$,

   b) $C' \cap D' = \emptyset$,

   c) $E \supseteq C'$ <u>and</u> $E \supseteq D'$,

   d) $C' \sim C$ <u>and</u> $D' \sim D$.

This axiom postulates the existence somewhere else in the space of disjoint, probability-equivalent copies of the not necessary disjoint sets C and D. Moreover, these copies are included in a copy of the union of two other disjoint sets A and B that are more probable than C and D. Axioms i)-v) together are sufficient for the existence of a probability measure. One proof first introduces a restricted concatenation operation as follows: If $\tilde{A}$ denotes the equivalence class containing A, then let

$$\mathcal{B} = \{(\tilde{A},\tilde{B}) \mid A > \emptyset,\ B > \emptyset, \text{ and } \exists A' \in \tilde{A},\ B' \in \tilde{B} \ni A' \cap B' = \emptyset\}.$$

When both A and B are very probable, they cannot be concatenated because no pair of events indifferent to the (A,B) pair will be disjoint. We now define the concatenation operation

$$\circ : \mathcal{B} \longrightarrow \mathcal{E}/\sim$$

by

$$\tilde{A} \circ \tilde{B} = \widetilde{A' \cup B'}.$$

By the definition of $\mathcal{B}$, concatenation is restricted, essentially, to disjoint events.

Theorem 2.1: <u>If</u> $\langle X, \mathcal{E}, \geqslant \rangle$ <u>satisfies axiom</u> i)-v), <u>then</u> $\langle \mathcal{E}/\sim, \mathcal{B}, \geqslant, \circ \rangle$ <u>is an extensive system</u> (i.e. <u>a positive, solvable, Archimedean, weakly-ordered local semigroup</u>).

Surprisingly enough, the only tricky part of the proof is to show associativity. It follows immediately from Theorem 1.3 that a measure P exists, and by Axiom ii) we may choose its unit so that $P(X) = 1$. Thus P is unique. An extension of this theory to a weak ordering of conditional events, i.e. of the form $A \mid B \geq C \mid D$, can be found in Luce (1968). The big problem there is that we must construct both the multiplicative structure inherent in the conditional probability representation, i.e.,

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)} \geq \frac{p(C \cap D)}{p(D)} = p(C \mid D) \text{ iff } A \mid B \geq C \mid D, \quad (1)$$

and, at the same time, the usual additivity of probability, i.e.,

$$p(A \cup B) = p(A) + p(B), \text{ if } A \cap B = \emptyset. \quad (2)$$

Additivity is established by showing that the ordering induced by $A|X \geq B|X$ on $\mathcal{E}$ satisfies the above unconditional axioms. The conditional axioms are also shown to lead, via extensive measurement theory, to a representation satisfying eq. (1) which is unique up to a positive power. The main difficulty in the proof is to show that the probability of eq. (2) is the same as one of the family satisfying eq. (1). Techniques of functional equations are used to show this.

3. Positive Difference Structures

A possible task for measurement theoreticians in the behavioral sciences is to try to reduce the natural formulation of their problems to cases of extensive measurement. A useful trick, it turns out, is to reduce them to the special case of extensive measurement known as positive difference structures. These structures can best be exemplified by an axiomatization of length measured on a long (possibly infinite) ruler.

If we compare length with mass, one of the main differences not captured by extensive measurement is the fact that length is naturally isomorphic with intervals on the real line. Intervals can be characterized by their endpoints, and the concatenation of adjacent intervals is especially natural: ab o bc = ac. The concatenation of non-adjacent intervals, such as ab and cd, has no comparable direct definition and one of our problems is to formulate an indirect one. Each interval can be identified in two ways: as ab and as ba. There is, however, a natural interpretation of direction, which leads to calling one a positive interval and the other negative. We will attend only to a subset, which will be called A^*, of the positive intervals.

The primitives for our axiom system are an abstract set A, a set $A^* \subset A \times A$, which will be axiomatized in such a way as to be interpreted as a set of positive intervals, and $\succsim$, an ordering on A^*, i.e. a subset of $A \times A \times A \times A$. The axioms are:

i) $\langle A^*, \succsim \rangle$ *is a weakly ordered set;*

ii) *if* $ab, bc, a'b', b'c' \in A^*$ *and if* $ab \succsim a'b'$ *and* $bc \succsim b'c'$, *then*

a) $ac, a'c' \in A^*$, *and*

b) $ac \succsim a'c'$;

iii) *if* $ab, bc \in A^*$ *then* $ac \succ ab, bc$;

iv) *if* $ab, cd \in A^*$ *and* $ab \succ cd$, *then* $\exists\, c', d' \in A \ni ad' \sim c'b \sim cd$,

v) *for all* $ab, bc, ab', b'c \in A^*$, *if* $ab \sim b'c$, *then* $ab' \sim bc$.

Axioms ii)-v) have a very simple interpretation in terms of length, and can be illustrated by drawing a line with the relevant points on it. In such a structure, a sequence $a_1, \ldots, a_i, \ldots \in A$ with $(a_{i+1}, a_i) \in A^*$, for all i, is called a *standard sequence* iff there exists an $ab \in A^*$ such that $a_{i+1}a_i \sim ab$ for each i.

vi) *If* $\{a_i\}$ *is a standard sequence, then* $\{n \mid n \in I^+, cd \succsim a_n a_1$ *is finite* (Archimedean axiom).

We now define

$$B = \left\{ (\widetilde{ab}, \widetilde{cd}) \mid \exists\, a', b', c' \in A \ni a'b', b'c' \in A^* \wedge ab \sim a'b' \wedge cd \sim b'c' \right\}$$

And if $(\widetilde{ab}, \widetilde{cd}) \in B$, then $\widetilde{ab} \circ \widetilde{cd} = \widetilde{a'c'}$.

Theorem 3.1: *If* $\langle A, A^*, \succsim \rangle$ *satisfies axioms* i)-vi), *then* $\langle A^* | \sim, B, \succsim, \circ \rangle$ *is an extensive system, provided that there exist* $ab, cd \in A^*$ *such that* $ab \succ cd$.

Corrollary 3.2: There exists a real-valued, order preserving mapping ψ on $A^*|\sim$ which is unique up to multiplication by a positive real number.

Corrollary 3.3: If $a, b \in A$, if not $ac \sim bc$ for all $c \in A$, and if for all $a, b, \in A^*$ either $ab \in A^*$ or $ba \in A^*$, then there exists a function φ: $A \rightarrow Re$, such that

$$\psi(ab) = \varphi(a) - \varphi(b);$$

moreover, φ is unique up to a positive linear transformation.

4. Additive Conjoint Measurement

For most attributes of interest in the behavioral sciences no natural concatenation operation is available. This means that the direct use of extensive measurement is impossible. If we accept N.R. Campbell's dictum "fundamental measurement = extensive measurement", then fundamental measurement is impossible in the behavioral sciences. This conclusion was reached after careful deliberation by the members (among them Campbell) of a British committee who investigated the possibility of measurement in psychology. It has proved far too pessimistic and premature since, in recent years, a number of quite different, but equally fundamental, systems have been proposed, among them conjoint measurement, the topic of this section, and subjective expected utility, the topic of the 6th one. In conjoint measurement no concatenation operation is assumed, but another kind of structure having to do with the fact that most attributes can be manipulated by several independent variables, sometimes permits representations of the following type.

Let $\geqslant$ be an ordering of a Cartesian product $\prod_{i=1}^{n} A_i$, where each A_i is a set. Such a structure is called decomposable relative to (a real-valued function) F: $Re^n \longrightarrow Re$ if there exist functions φ_i: $A_i \rightarrow Re$, $i=1, \ldots, n$, such that for all $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in \prod A_i$:

$$(a_1, \ldots, a_n) \geqslant (b_1, \ldots, b_n) \text{ iff } F(\varphi_1(a_1), \ldots, \varphi_n(a_n)) \geqslant F(\varphi_1(b_1), \ldots, \varphi_n(b_n))$$

This definition expresses the fact that the contributions of the variables to the overall measure are independent of one another. This is a very general requirement, but it is by no means a trivial one, since it is not satisfied in all cases. For example, suppose that n=2 and let $\varphi_i: A_i \to Re$ and $\psi_i: A_i \to Re$ for i=1, 2, be given functions. For all $a,b \in A_1$, $p,q \in A_2$, define the ordering $\geqslant$ on $A_1 \times A_2$ by

$$(a,p) \geqslant (b,q) \text{ iff } \varphi_1(a) + \varphi_2(p) + \psi_1(a)\psi_2(p) \geqslant \varphi_1(b) + \varphi_2(q) + \psi_1(b)\psi_2(q)$$

This "additive structure with independent interaction" is not in general decomposable relative to any function. In spite of the natural interest in this model as, perhaps, the simplest form of interaction, we know nothing about its properties. No necessary conditions (except weak ordering) have been discovered.

The only two-component cases* so far investigated are the additive one, $F(x,y) = x + y$, and the multiplicative one, $F(x,y) = xy$. In general, the multiplicative model can not be reduced to an additive one by a logarithmic transformation because the scale values may be negative. In the three-component case, the functions $F(x,y,z) = x + y + z$, xyz, $(x + y)z$, and $x + yz$, are thoroughly investigated. In the sequel we will confine ourselves to the n-component additive case.

The most familiar example of an additive model is, of course, the one from economics that says that the cardinal utility of a commodity bundle is equal to the sum of the utilities of each of its components. As a matter of fact this model inspired much of the earlier work on additive conjoint measurement (cf., for example, Debreu, 1960). In psychology a two-dimensional example is obtained if we let subjects compare the loudness of pure tones, varying both their intensities and frequencies. In both examples it is possible to draw indifference curves to represent the equivalence classes in the data structure. The theory establishes a systematic way to associate numbers with the indifference curves that, in a sense, represent the amount of attribute exhibited by that curve.

An example where this has been done successfully (but independently of the theory, I must admit) can be found in studies of Campbell and Masterson (1968) on the aversiveness of shocks. On one side of a shuttle box they placed shock with resistance Z and voltage V_Z, and on the other, shock with resistance Z_o and voltage V_o. Throughout

* of decomposition

the experiment Z_o was held fixed, and for each (Z,V_Z)-pair they discovered the value of V_o such that 50% of the animals selected each side. They found that the empirical values satisfied a relationship of the form

$$V_Z = \alpha + \beta V_o + \gamma Z + \delta V_o Z,$$

which is equivalent to the additive form

$$\log(\delta V_Z + \gamma\beta - \alpha\delta) = \log(\gamma + \delta V_o) + \log(\beta + \delta Z).$$

Many other examples of trade-off between variables can be found in the literature.

In all axiomatizations of additive conjoint measurement, certain necessary <u>cancellation conditions</u> play an important part. One of these conditions expresses, quite directly, the fundamental independence of the variables: for all $a,b \in A_1$ and $p,q \in A_2$,

$$\begin{aligned} &(a,p) \geqslant (b,p) \text{ iff } (a,q) \geqslant (b,q) \\ \text{and}\quad &(a,p) \geqslant (a,q) \text{ iff } (b,p) \geqslant (b,q). \end{aligned}$$

Notice how this follows if an additive representation is true:

$$\begin{aligned} (a,p) \geqslant (b,p) &\text{ iff } \varphi_1(a) + \varphi_2(p) \geqslant \varphi_1(b) + \varphi_2(p) \\ &\text{ iff } \varphi_1(a) \geqslant \varphi_1(b) \\ &\text{ iff } \varphi_1(a) + \varphi_2(q) \geqslant \varphi_1(b) + \varphi_2(q) \\ &\text{ iff } (a,q) \geqslant (b,q). \end{aligned}$$

As a matter of fact this condition justifies the natural definition of an induced order $\geqslant_1$ on the first dimension:

$$a \geqslant_1 b \text{ iff } \forall x \in A_2 \text{ we have } (a,x) \geqslant (b,x).$$

Similarly, we may define $\geqslant_2$ on the second dimension. Another property that can be arrived at in the same way is called <u>double cancellation</u>: for all $a,b,f \in A_1$ and $p,q,x \in A_2$,

$$\text{if } (a,x) \geqslant (f,q) \text{ and } (f,p) \geqslant (b,x), \text{ then } (a,p) \geqslant (b,q).$$

Other cancellation properties can be obtained by considering three or more inequalities in which all save four elements can be "cancelled". Later, we explicitly give one of the three inherently different forms of triple cancellation. All these conditions are necessary, and all may be checked directly in any set of data. For fairly large data structures, this is a very time-consuming task; indeed, it is only practical if a computer is used.

We may, however, restrict the number of necessary conditions needed to get an additive representation by strenghtening the sufficient conditions that impose structure on the system. This is done in the following axiomatization of n-dimensional conjoint measurement, with $n \geqslant 3$. Let A denote the cartesian product $\prod_{i=1}^{n} A_i$, when $n \geqslant 3$.

i) $\langle A, \geqslant \rangle$ is a weakly ordered set.

ii) If $N = \{1, 2, \ldots, n\}$, then for all $M \subseteq N$ the ordering induced on $\prod_{i \in M} A_i$ for any fixed choice of elements in $\prod_{i \in N-M} A_i$ is independent of that choice.

Axiom ii) permits us to define $\geqslant_i$ on A_i in the obvious way, and it turns out to be the only cancellation axiom that we need when $n \geqslant 3$, provided that we impose a strong solvability condition. Luce and Tukey (1964) postulated the following solution (of equations) axiom: $\forall\, a,b \in A_1,\ p \in A_2,\ \exists\, x \in A_2 \ni (a,p) \sim (b,x)$. This has been justly criticized as being too strong; in many examples it is easily seen not to be satisfied (e.g., loudness judgments). Therefore, Luce (1966) modified it to the following restricted solvability condition: $\forall\, a,b \in A_1,\ p \in A_2$, if $\exists\, \bar{x}, \underline{x} \ni (b,\bar{x}) \geqslant (a,p) \geqslant (b,\underline{x})$, then $\exists\, x \ni (b,x) \sim (a,p)$. A simple generalization gives us

iii) For all $(a_1, \ldots, a_n) \in A$, $(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n) \in A_1 \times \ldots \times A_{i-1} \times A_{i+1} \times \ldots \times A_n$, if there exist a $\bar{b}_i$ and a $\underline{b}_i$ such that $(b_1, \ldots, \bar{b}_i, \ldots, b_n) \geqslant (a_1, \ldots, a_n) \geqslant (b_1, \ldots, \underline{b}_i, \ldots, b_n)$, then there exists a b_i such that $(b_1, \ldots, b_i, \ldots, b_n) \sim (a_1, \ldots, a_i, \ldots, a_n)$

Moreover, we need a nontrivialness axiom:

iv) For at least three components A_i there exist $a_i, b_i \in A_i$ such that $a_i >_i b_i$. Such components are called essential.

In order to state the necessary Archimedean axiom, we need the following definition: Let N be a succession of integers, positive and/or negative, finite or infinite: A sequence $\{a_i^\gamma \mid a_i^\gamma \in A_i \wedge \gamma \in N \wedge \exists p,q \in A_j,\ j \neq i,\ \ni (\ldots,a_i^\gamma,\ldots,p,\ldots) \sim (\ldots,a_i^{\gamma+1},\ldots,q,\ldots) \text{ for } \gamma, \gamma+1 \in N\}$ is called a standard sequence. The Archimedean axiom simply is

v) Every bounded standard sequence is finite.

Theorem 4.1; If axioms i)-v) hold, then there exist $\varphi_i : A_i \rightarrow$ Re, $i = 1, \ldots, n$, such that for all $(a_1,\ldots,a_n),(b_1,\ldots,b_n) \in A$,

$$(a_1,\ldots,a_n) \geq (b_1,\ldots,b_n) \text{ iff } \sum_{i=1}^{n} \varphi_i(a_i) \geq \sum_{i=1}^{n} \varphi_i(b_i).$$

Moreover, if φ_i' is another set of such functions, then $\exists\, \alpha > 0$, β_i, $i=1,\ldots,n$, such that $\varphi_i = \alpha \varphi_i' + \beta$.

Notice that we have not yet formulated a representation theorem for the two-component cases. This we must do, not only because it is of interest and importance in its own right, but also because the only known proof of the n-component case involves reducing the problem to the two-component case. There is no need to alter the weak ordering, solvability, and Archimedean axioms in the two-component case. The property of independence is, however, too weak and it is replaced by two cancellation properties, namely, double cancellation:

(ii) for all $a,b,f \in A_1$ and $p,q,x \in A_2$

$$(a,x) \geq (f,p) \wedge (f,p) \geq (b,x) \text{ imply } (a,p) \geq (b,q)$$

and by one of the three forms of triple cancellation:

(iii) for all $a,b,f,g \in A_1$ and $p,q,x,y \in A_2$

$$(a,x) \geq (b,y) \wedge (f,y) \geq (g,x) \wedge (g,p) \geq (f,g) \text{ imply } (a,p) \geq (b,q).$$

From these assumptions it is easily shown that independence holds and so a weak ordering is induced on each component. The final assumption is that both co-ordinates are essential. This is all that is needed.

Theorem 4.2; <u>If $\langle A_1 \times A_2, \geqslant \rangle$ satisfies the weak ordering, double and triple cancellation, restricted solvability, Archimedean and essentialness conditions, then the conclusion of theorem 4.1 holds with n=2.</u>

It is an open problem to show that double and triple cancellation are independent axioms, or to derive one from the other.

We now outline the nature of the proofs of Theorem 4.1 and 4.2. Let A_i and A_j be any two essential components in the n-component case. It is easy to see that the induced order $\geqslant_{ij}$ satisfies all of the assumptions of Theorem 4.2 except the two cancellation properties. These also follow. It is fairly difficult to prove them for restricted solvability, but easy for unrestricted. For example, suppose $a, b, f \in A_i$ and $p, q, x \in A_j$ and $(a,x) \geqslant_{ij} (f,q)$ and $(f,p) \geqslant_{ij} (b,x)$. Let A_k be any other essential component and let $u \in A_k$. By solvability, $\exists\, v \in A_k$ such that

$$(f,x,v) \sim_{ijk} (a,x,u) \geqslant_{ijk} (f,q,u),$$

and so $(x,v) \geqslant_{jk} (q,u)$. Since $(f,x,u) \sim_{ijk} (a,x,u)$, then by independence x may be replaced by p, and so

$$(a,p,u) \sim_{ijk} (f,p,v) \geqslant_{ijk} (b,x,v) \geqslant_{ijk} (b,q,u).$$

Thus, $(a,p) \geqslant_{ij} (b,q)$. The proof for triple cancellation is similar. By theorem 4.2, there exists an additive representation $\varphi_i + \varphi_j$ on $A_i \times A_j$.

There is, however, a problem. Suppose we picked $i, j \in N$ and found mappings φ_i: $A_i \rightarrow Re$ and φ_j: $A_j \rightarrow Re$. We can of course also choose another pair $i, k \in N$, with $k \neq j$. This gives us the mappings φ_i': $A_i \rightarrow Re$ and φ_k: $A_k \rightarrow Re$. It must be shown that $\varphi_i' = \alpha\,\varphi_i + \beta$, with $\alpha > 0$. Finally, it can also be shown that if we choose our functions φ_i carefully so that the units and zeroes are appropriately related, then this provides an additive representation. When we accept this, the problem is reduced to the two-component one.

The next step in the reduction process used to prove Theorem 4.1 is to reduce the two dimensional system of Theorem 4.2 to a special, symmetric case. In Figure 1, the cartesian product $A_1 \times A_2$ is portrayed as a rectangle, as it will be in the desired representation.

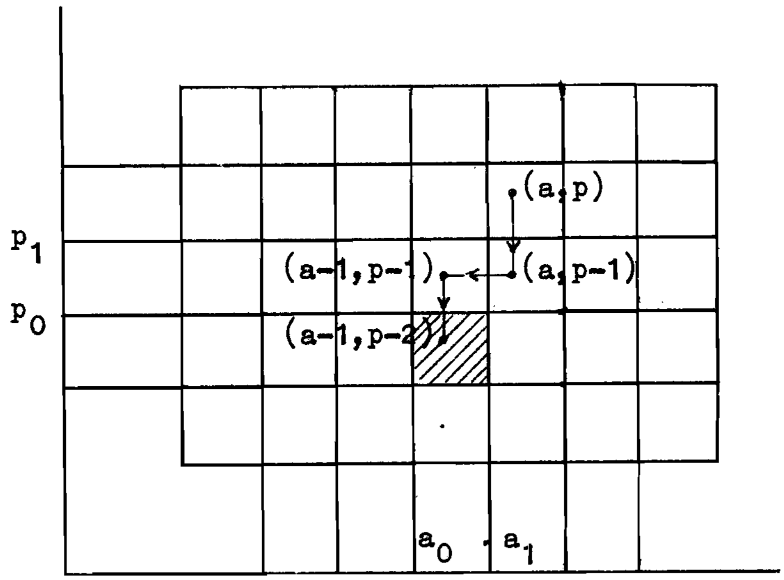

Figure 1

The points a_0, a_1, p_0, p_1 are chosen in such a way that $(a_0, p_1) \sim (a_1, p_0)$ and they determine the unit and the unit square consisting of all points (x, y) for which $a_1 \geq x \geq a_0$ and $p_1 \geq y \geq p_0$ (the shaded area in Figure 1).

Suppose that we now want to assign a number to the point (a,p) in Figure 1. We move unit steps in both directions until we arrive at a point in the unit square. This process is also illustrated in Figure 1. Then, of course, the sensible thing to try is the assignment:

$$\varphi_1(a) = \varphi_1(a-1) + 1$$

$$\varphi_2(p) = \varphi_2(p-2) + 2,$$

where the coordinates (a-1, p-2) are coordinates of a point in the unit square. The main part of the proof is to show that this inductive process can indeed be carried out. It relies heavily on the assumed triple cancellation and restricted versions of the other two triple cancellation conditions which can be proved from the axioms.

A system $\langle A_1 \times A_2, \succcurlyeq \rangle$ is called <u>symmetric</u> if for all $a, b \in A_1$, $\exists p, q \in A_2 \ni (a,p) \sim (b,q)$. Such systems can be mapped into a square, whereas the general case results in a rectangle. So we are done if we can get the representation in this case. Define the set

$A_1^* = \{ab \mid a,b \in A_1 \wedge a >_1 b\}$, and define $\geqslant_1^*$ on A_1^* by

if $ab, cd \in A_1^*$, then $ab \geqslant_1^* cd$ iff $\forall p,q \in A_2$ whenever

$$(a,p) \sim (b,q), \text{ then } (d,q) \geqslant (c,p)$$

Do the same thing for the second coordinate, which gives an A_2^* and a $\geqslant_2^*$. Now it can be proved that $\langle A_1, A_1^*, \geqslant_1^* \rangle$ and $\langle A_2, A_2^*, \geqslant_2^* \rangle$ satisfy the axioms for positive difference systems, so by corrolary 3.3 (previous section)

$$\exists \varphi_1 : A_1 \to Re, \ \varphi_2 : A_2 \to Re \ni$$

$$ab \geqslant_1^* cd \text{ iff } \varphi_1(a) - \varphi_1(b) \geqslant \varphi_1(c) - \varphi_1(d)$$

$$pq \geqslant_2^* uv \text{ iff } \varphi_2(p) - \varphi_2(q) \geqslant \varphi_2(u) - \varphi_2(v)$$

Moreover we pick an $a_1 > a_0$ and $p_1 > p_0$ such that $(a_0, p_1) \sim (a_1, p_0)$ and we set $\varphi_1(a_0) = \varphi_2(p_0) = 0$ and $\varphi_1(a_1) = \varphi_2(p_1) = 1$, and define $\psi_1(ab) = \varphi_1(a) - \varphi_1(b)$, $\psi_2(pq) = \varphi_2(p) - \varphi_2(q)$. The next thing to be established is that the two systems $\langle A_1, A_1^*, \geqslant_1^* \rangle$ and $\langle A_2, A_2^*, \geqslant_2^* \rangle$ have essentially the same structure.

Define: $\theta(\widetilde{ab}) = \widetilde{qp}$ iff $(a,p) \sim (b,q)$,

then it is shown that θ is an isomorphism, and that $\psi_1 = \psi_2(\theta)$. The final step in the proof of Theorem 4.2 is simple: It just remains to be proved that the ψ's are order preserving. Observe that

$$\begin{aligned}
\varphi_1(a) + \varphi_2(p) \geqslant \varphi_1(b) + \varphi_2(q) & \text{ iff } \varphi_1(a) - \varphi_1(b) \geqslant \varphi_2(q) - \varphi_2(p) \\
& \text{ iff } \psi_1(\widetilde{ab}) \geqslant \psi_2(\widetilde{qp}) \\
& \text{ iff } \psi_1(\widetilde{ab}) \geqslant \psi_2(\theta(\widetilde{cd})), \text{ where } (c,p) \sim (d,q) \\
& \text{ iff } \psi_1(\widetilde{ab}) \geqslant \psi_1(\widetilde{cd}) \\
& \text{ iff } \widetilde{ab} \geqslant_1^* \widetilde{cd} \\
& \text{ iff } (a,p) \geqslant (b,q).
\end{aligned}$$

5. Bisymmetry Systems

A theory due to Pfanzagl, which assumes a concatenation operation, can be reduced to additive conjoint measurement. His theory is similar to extensive measurement, but it is more general in that, among other things, it axiomatizes the formation of weighted means. Suppose that ρ is a fixed number in $[0,1]$ and for any real numbers a, b, we define $a \circ b = \rho a + (1-\rho)b$. We see that o has many properties different from extensive measurement. For example, $a \circ a \sim a$, o is not commutative, and o is not associative.

Pfanzagl begins with a structure $\langle A, \circ, \geqslant \rangle$ and he assumes:

i) $\langle A, \geqslant \rangle$ is a weakly ordered set,

ii) $a \geqslant b$ iff $a \circ c \geqslant b \circ c \wedge c \circ a \geqslant c \circ b$,

iii) A is connected in the order topology induced by $\geqslant$,

iv) $a \circ b$ is continuous in both a and b,

v) $(a \circ b) \circ (c \circ d) \sim (a \circ c) \circ (b \circ d)$.

This last axiom, the bisymmetry axiom, does not imply that the system is associative and/or commutative. Observe that this axiom is true for weighted means. Pfanzagl proved the following result:

Theorem 4.3: If $\langle A, \circ, \geqslant \rangle$ satisfies axioms i)-v), there exist real numbers $\rho, \sigma > 0, \lambda$ and a function $\varphi: A \rightarrow Re$ such that

i) $a \geqslant b$ iff $\varphi(a) \geqslant \varphi(b)$,

ii) φ is continuous,

iii) $\varphi(a \circ b) = \rho\,\varphi(a) + \sigma\,\varphi(b) + \lambda$,

iv) if φ' also satisfies i)-iii) then $\varphi' = \alpha\varphi + \beta$, with $\alpha > 0$.

Corrollary 4.3:

i) if $a \circ a \sim a$, then $\lambda = 0$, and $\rho + \sigma = 1$

ii) if the structure is commutative, then $\rho = \sigma = 1$

iii) if it is both commutative and associative, then $\rho = \sigma = 1$, and $\lambda = 0$

In the last case we have extensive measurement (set $\psi = \varphi + \lambda$). In the first case we have the weighted mean interpretation. The proof of the theorem can be carried out by reducing it to the twodimensional additive conjoint case.

Define

$$(a,p) \geqslant (b,q) \quad \text{iff} \quad a \circ p \geqslant b \circ q$$

The various axioms of theorem 4.2 must be proved. We establish double cancellation as an example: suppose $(a,x) \geqslant (f,q)$ and $(f,p) \geqslant (b,x)$, then by definition $aox \geqslant foq$ and $fop \geqslant box$. Now consider

$$\begin{aligned}
(aop)o(xox) &\sim (aox)o(pox) && \text{(by bisymmetry)}\\
&\geqslant (foq)o(pox) && \text{(by monotonicity)}\\
&\sim (fop)o(qox) && \text{(by bisymmetry)}\\
&\geqslant (box)o(qox) && \text{(by monotonicity)}\\
&\sim (boq)o(xox) && \text{(by bisymmetry)}
\end{aligned}$$

It follows by monotonicity and transitivity of $\geqslant$ that $(aop) \gtrsim (boq)$, hence $(a,p) \geqslant (b,q)$. The proof of triple cancellation is somewhat more complicated, but similar. The most difficult part is to derive the solvability and Archimedean conditions from the topological axioms (iii) and iv). This can be done.

6. Conditional Expected Utility Theory

Expected utility theories attempt to describe the behavior of a rational decision maker when confronted with choices among uncertain prospects. The principal primitive notions are "event" and "consequence". An "uncertain prospect" or "gamble" consists of a finite number of chance events, say $e_1,\ldots,e_n$, and a consequence associated with each event, denoted by $c_1,\ldots,c_n$. Expected utility theories construct two real-valued functions: a utility function u that maps the set of consequences into the reals and a probability measure P that is defined on the events. The expected utility EU is computed by taking expectations:

$$EU = \sum_{i=1}^{n} u(c_i)P(e_i),$$

and it orders gambles in the same way as do the preferences of the rational decision maker.

The first modern discussion of expected utility theory is in an appendix of the 1947 edition of Von Neumann and Morgenstern's classic book. They were concerned with simple gambles in which consequence <u>a</u> arises with probability <u>p</u> and consequence <u>b</u> with probability 1-p. The probabilities were assumed to be given in a numerical form. Such a gamble can be written as apb. Von Neumann and Morgenstern also introduced a compounding operation that makes it possible to construct more complicated gambles, such as (apb)qc, from simple ones. They axiomatized an ordering $\succcurlyeq$ over simple gambles and simple compounds of them. The axioms they introduced garanteed the existence of an order-preserving numerical expected utility function. The construction of this function depended on the numerical values of the objective probabilities, which obviously means that their procedure is not an example of fundamental measurement. Blackwell and Girshick, Samuelson, and others generalized this approach to n-component gambles (cf. Luce and Raiffa, 1957). A much deeper generalization was provided by Savage (1954), who completely dropped the objective probability assumption. He introduced axioms that are sufficient for the existence of both a subjective probability measure P and a utility u with the property that a gamble f is preferred to a gamble g if and only if the subjective expected utility of f is greater than the subjective expected utility of g. Savage's theory is very general, but as we shall see it has certain unnatural features.

Another line of development was initiated by the philosopher Ramsey (1931) whose ideas were worked out in detail by Suppes and his collaborators (cf. Davidson and Suppes, 1956). The main difference from Savage's approach is that Ramsey first constructed a utility function, from which a probability function is then developed; whereas, Savage began with axioms sufficient for the existence of a (subjective) probability measure, and he then defined utilities in terms of these probabilities (as Von Neumann and Morgenstern did in terms of objective probabilities). The Ramsey-Suppes-Davidson-Winet approach is still restricted to very simple, two-component gambles with independent events. Pfanzagl (1967) generalized this in such a way that compounding of non-independent events became possible. A still more general treatment of the problem is given by Luce and Krantz (1968). Not only does their axiom system cover general gambles as well, but a more general representation is obtained which also generalizes some ideas of Jeffrey. Schematically this historical discussion may be summarized as follows:

Von Neumann & Morgenstern

Blackwell & Girshick — Ramsey
Samuelson — Suppes et al.
Savage — Pfanzagl — Jeffrey

Luce & Krantz

In comparing Savage's statistical decision theory with the Luce-Krantz conditional theory, it is important to recognize that in most cases of practical interest people think in a conditional way. If you want to decide whether to go to Paris by plane or by car, for example, the reasoning is typically conditional. When considering the events that can arise when going by car, the fact that the plane may crash is simply not relevant. It is, of course, when considering the possible events associated with the flight. Moreover, the unconditional formulation of decision theory in the Savage-approach is highly inefficient. Consider the following simple example, in which there are two gambles: the first one is the throw of a die (D)

event	1	2	3	4	5	6
consequence	-3	-2	-1	1	2	3

The second one is the throw of a coin (C)

event	Head	Tail
consequence	10	-10

In the Savage formulation we must consider the complete cartesian product of states of nature and list all outcomes.

	1H	1T	2H	2T	3H	3T	4H	4T	5H	5T	6H	6T
D	-3	-3	-2	-2	-1	-1	+1	+1	+2	+2	+3	+3
C	+10	-10	+10	-10	+10	-10	+10	-10	+10	-10	+10	-10

In the conditional formulation, we list all possible consequences and for each decision list the events that cause each to arise:

	-10	-3	-2	-1	1	2	3	10
D	ϕ	1	2	3	4	5	6	ϕ
C	T	ϕ	ϕ	ϕ	ϕ	ϕ	ϕ	H

Of course the information in these two schemes is the same, but evidently the Savage notation is the more redundant, even in this simple example.

That the two approaches can be translated into each other in the finite case may not be clear at first sight. To show it, we first list the primitives of the statistical approach: a set $\mathcal{S}$ of states of nature, a set $\mathcal{T}$ of consequences, and mappings of the form f: $\mathcal{S} \to \mathcal{T}$. The representation involves two functions, a utility u: $\mathcal{T} \to Re$ and a probability Q: $\mathcal{S} \to [0,1]$, and subjective expected utility is defined as $\sum_i u[f(s_i)]Q(s_i)$. The summation is over all states of nature.

In the conditional approach, the primitives are a set X, an algebra $\mathcal{E}$ of subsets of X, a set of consequences $\mathcal{T}$, and mappings f_A: $A \to \mathcal{T}$ for $A \in \mathcal{E}$. The representation yields u: $\mathcal{T} \to Re$, and P: $\mathcal{E} \to [0,1]$, and the expected utility is defined as $\Sigma u(c_i)P[f_A^{-1}(c_i)|A]$, where the summation is over all consequences. The translation of the statistical theory into the conditional is trivial: define $X = \mathcal{S}$; $\mathcal{E} = 2^{\mathcal{S}}$; $f_X = f$; $P = Q$; and u = u. The translation from the conditional to the statistical theory is less obvious. Define $\mathcal{S} = \prod_{A_i \in \mathcal{E}} A_i$; $Q(s_i) = \prod_{\ell} [P(s_i^{\ell}) \mid P(A_{\ell})]$, where $s_i \in \mathcal{S}$, i.e.
$s_i = (s_i^1, \ldots, s_i^{\ell}, \ldots, s_i^{\mathcal{E}})$; $f(s_i) = f_A(s_i^{\ell})$ where ℓ is the index such that $A_{\ell} = A$; and u = u. One then shows that the latter expectation implies the former.

The primitives of the Luce-Krantz axiomatization are as follows. First, we have an algebra of subsets $\mathcal{E}$ of some abstract set X. Moreover, a subset $\eta \subseteq \mathcal{E}$ must be characterized by the axioms. Intuitively, the elements of η can be thought of as those events that are judged as having no probability of occurring. We condition only on events in $\mathcal{E} - \eta$ to avoid division by 0 in the representation. Again $\mathcal{T}$ denotes the set of consequences. The decisions are a set of functions $\mathcal{D} \subseteq \{f_A | f_A: A \to \mathcal{T}, A \in \mathcal{E} - \eta\}$. Finally we have a binary relation $\succsim$ on $\mathcal{D}$. Decisions can be compounded in the following way: if $A \cap B = \emptyset$, $A, B \in \mathcal{E} - \eta$, $f_A, g_B \in \mathcal{D}$ then $f_A \cup g_B(x) = \text{def} \begin{cases} f_A(x) & \text{if } x \in A \\ g_B(x) & \text{if } x \in B \end{cases}$. If $B \subseteq A$ and $B \in \mathcal{E} - \eta$, then $(f_A)_B$ denotes the restriction of f_A to the set B. The axioms of a conditional decision structure can now be stated as follows.

For all $A, B \in \mathcal{E} - \eta$, f_A, f'_A, $g_B \in \mathcal{D}$

i) a. $A \cap B = \emptyset \Rightarrow f_A \cup g_B \in \mathcal{D}$

b. $B \subseteq A \Rightarrow (f_A)_B \in \mathcal{D}$

ii) $\langle \mathcal{D}, \succsim \rangle$ is a weakly ordered set.

iii) $A \cap B = \emptyset \wedge f_A \sim g_B \Rightarrow f_A \cup g_B \sim f_A$.

iv) $A \cap B = \emptyset$, $f_A \succsim f'_A \Leftrightarrow f_A \cup g_B \succsim f'_A \cup g_B$.

This last axiom bears an obvious resemblance to the monotonicity condition in extensive measurement (cf. section 1, axiom vi). Moreover, the famous sure-thing principle is a special case of axiom iv. The sure-thing principle asserts that if for each possible outcome the consequence of gamble f is preferred or indifferent to the consequence of gamble g and for at least one outcome it is strictly preferred, then gamble f will be preferred to gamble g.

For our next axiom we need the following definition: a sequence $\left\{f_A^{(i)} \mid f_A^{(i)} \in \mathcal{D},\ i \in N\right\}$ is called a <u>standard sequence</u> if $\exists\, B \in \mathcal{E} - \eta \ni A \cap B = \emptyset \wedge \exists g_B^{(0)},\ g_B^{(1)} \in \mathcal{D} \ni \forall\, i, i+1 \in N,\ f_A^{(i)} \cup g_B^{(1)} \sim f_A^{(i+1)} \cup g_B^{(0)}$.

The Archimedean axiom is as usual:

v) <u>every bounded standard sequence is finite</u>.

Moreover, we use the notion of a standard sequence in:

vi) $\left\{f_A^{(i)} \mid N\right\}, \left\{h_A^{(i)} \mid N\right\}$ are s.s. $\wedge \exists\, k, k+1 \in N \ni f_A^{(k)} \sim h_A^{(k)} \wedge$
$f_A^{(k+1)} \sim h_A^{(k+1)} \Longrightarrow \forall\, i \in N,\ f_A^{(i)} \sim h_A^{(i)}$.

Our next axiom characterizes the set η.

vii) a. $R \in \eta \wedge S \subseteq R \Rightarrow S \in \eta$.
b. $R \in \eta \Leftrightarrow f_{A \cup R} \sim (f_{A \cup R})_A$.

The next axiom is a non-triviality assumption.

viii) a. $\mathcal{E} - \eta$ <u>has at least three pairwise disjoint elements</u>.
b. $\mathcal{D}/\sim$ <u>has at least two equivalence classes</u>.

The final axiom is similar to the solvability axioms that we have previously used in the other measurement models. Suppose that $A, B \in \mathcal{E} - \eta$, $h_A^{(1)}, h_A^{(2)}, g_B, f_{A \cup B} \in \mathcal{D}$,

ix) a. $\exists\, h_A \in \mathcal{D} \ni h_A \sim g_B$.
b. $A \cap B = \emptyset \wedge h_A^{(1)} \cup g_B \gtrsim f_{A \cup B} \gtrsim h_A^{(2)} \cup g_B \Rightarrow \exists\, h_A \in \mathcal{D} \ni$
$h_A \cup g_B \sim f_{A \cup B}$.

Clearly the first part of axiom ix) is a form of unrestricted solvability, and this is one part of the axiom system that we would really like to weaken. Surprisingly enough, the second part of axiom ix) is independent of the first part, although it is rather like the restricted solvability assumption we made in additive conjoint measurement.

Theorem 5.1: *If* $\langle X, \mathcal{E}, \eta, \mathcal{T}, \mathcal{D}, \geq \rangle$ *satisfies axioms i)-ix), then there exist functions* $u: \mathcal{D} \rightarrow Re$ *and* $P: \mathcal{E} \rightarrow [0,1]$ *such that, for all* $f_A, g_B \in \mathcal{D}$,

a. $\langle X, \mathcal{E}, P \rangle$ *is a finitely additive probability space.*

b. $R \in \eta \Leftrightarrow P(R) = 0.$

c. $f_A \geq g_B \Leftrightarrow u(f_A) \geq u(g_B).$

d. *if* $A \cap B = \emptyset$, *then* $u(f_A \cup g_B) = u(f_A)P(A|A \cup B) + u(g_B)P(B|A \cup B).$

e. P *is unique, and u is unique up to a positive linear transformation.*

There are some important differences from Savage's result. In the first place, the utility function is not defined on the set of consequences. As a matter of fact, we could even do without $\mathcal{T}$ since the axioms nowhere refer to $\mathcal{T}$. Part d. of theorem 5.1 does not say that u is an expectation in the statistical sense; it has a major property of an expectation, but only under very special conditions is it actually one. The infiniteness of Savage's system is explicitly built into the set of states. The infiniteness of the Luce-Krantz system is in the set of decisions, but not necessarily in $\mathcal{E}$; the algebra of subsets $\mathcal{E}$ is only specified to be non-empty. A serious weakness in the statistical approach is the essential role in the construction of u and P that is played by constant decisions, i.e. decisions that have the same consequence independent of the state of nature. Constant decisions may be realizable in simple situations, but they are very unrealistic in realistic settings. It can be shown that there is a realization of the Luce-Krantz system in which there are no constant decisions. To get a utility function over $\mathcal{T}$, let c_A denote the constant decision with $f_A(x) = c$ for all $x \in A$. If we add the following assumptions,

x) $c \in \mathcal{T} \Rightarrow \exists A(c) \in \mathcal{E} - \eta \ni c_{A(c)} \in \mathcal{D}$

xi) $c_A, c_B \in \mathcal{D} \Rightarrow c_A \sim c_B$

to axioms i)-ix), then it follows that there exists $v: \mathcal{D} \rightarrow Re$ such that for any gamble $f_A \in \mathcal{D}$, $u(f_A) = E[v(f_A) \mid A]$, where E denotes taking expectations. A gamble is a decision with a finite image and for $c \in \mathcal{T}$

$f_A^{-1}(c) \in \mathcal{E}$. Note that assumption x) is trivially true if, for example, the toss of a coin is included in $\mathcal{E}$; it is much weaker than assuming that all constant decisions are in $\mathcal{D}$. Another important case, not admissible under Savage's axioms, is where the utility of a decision has contributions from both the consequences and the conditioning event. For example, suppose that $w: \mathcal{E} \to Re$ is such that for $A \cap B = \phi$, $w(A \cup B) = w(A)P(A \mid A \cup B) + w(B)P(A \mid A \cup B)$, and $v: \mathcal{T} \to Re$, then a utility for gambles may be defined as $u(f_A) = E\left[v(f_A) | A\right] + w(A)$. This more realistic model, which can be interpreted as admitting utility-for-gambling, is consistent with axiom i)-x) but, of course, not with xi) except when $w \equiv 0$.

Savage's method of proof was to obtain subjective probabilities, from which he constructed the utility function along the lines of Von Neumann and Morgenstern's proof. In our approach P and u come out simultaneously. We briefly sketch the nature of the proof. Take an arbitrary non-null event A and partition it into non-null sets A_i, $i=1, \ldots, n$. Let $\mathcal{D}_{A_i}$ denote the subset of $\mathcal{D}$ whose elements are defined on A_i. Define $\succsim_1$ on $\prod_i \mathcal{D}_{A_i}$:

$$(f_{A_1}, f_{A_2}, \ldots, f_{A_n}) \succsim_1 (g_{A_1}, g_{A_2}, \ldots, g_{A_n}) \Leftrightarrow f_{A_1} \cup f_{A_2} \cup \ldots \cup f_{A_n} \succsim g_{A_1} \cup g_{A_2} \cup \ldots \cup g_{A_n}$$

It is possible to show that in $\left\langle \prod_i \mathcal{D}_{A_i}, \succsim_1 \right\rangle$ the axioms of n-dimensional conjoint measurement are satisfied (if $n=2$ we have to be careful, but starting with a more refined partitioning of A makes $n \geq 3$ again). Therefore $\exists\, \varphi_i: f_{A_i} \to Re$, and these functions are, of course, order-preserving and unique up to positive linear transformations. A different partitioning defines other functions, but they can be shown to be of the same family. We have to pick a particular one: if $f^{(1)} > f^{(0)}$, then for any A there exist $g_A^{(1)} \sim f^{(1)}$ and $g_A^{(0)} \sim f^{(0)}$. Zero and unit are established by picking u_A so that $u_A(g_A^{(0)}) = 0$ and $u_A(g_A^{(1)}) = 1$. For any $\mathcal{D}_A$, this defines a unique u_A. And because $\mathcal{D}$ is the union of the $\mathcal{D}_A$, let u on $\mathcal{D}$ be the union of all these functions. For $A \cap B = \phi$, we have
$u(f_A \cup g_B) = \varphi_A(f_A) + \varphi_B(g_B) = P(A|A \cup B)u(f_A) + \beta_{A,B} + P(B|A \cup B)u(g_B) + \beta_{B,A}$, where $\varphi_A(f_A) = P(A|A \cup B)u(f_A) + \beta_{A,B}$ is simply the linear transformation that relates $\varphi_A(f_A)$ and $u(f_A) = u_A(f_A)$. It remains to be proved, that the P-values are conditional probabilities, that $\beta_{B,A} + \beta_{A,B} = 0$, and that u is

order-preserving. Since $g_A^{(0)} \sim f^{(0)} \sim g_B^{(0)}$, we also have $g_A^{(0)} \cup g_B^{(0)} \sim f^{(0)}$, so $u(g_A^{(0)} \cup g_B^{(0)}) = 0 = P(A|A \cup B)u(g_A^{(0)}) + P(B|A \cup B)u(g_B^{(0)}) + \beta_{A,B} + \beta_{B,A} = 0 + 0 + \beta_{A,B} + \beta_{B,A}$. By using the $g^{(1)}$-elements in a similar way, we establish that the P's are conditional probabilities. The proof that u is order-preserving is rather difficult.

Although this is perhaps the most satisfactory existing axiomatization of subjective expected utility, some improvements are clearly possible. First axiom ix) is rather strong; we especially would like to weaken part a). Second, some data and examples suggest that the sure-thing principle is not a very good description of actual decision making (although it certainly seems rational to many people, though not to all). And finally, in the algebra $\mathcal{E}$, all events are treated as equally realizable, and it may be useful to partition $\mathcal{E}$ into those that are realizable and those that are only used for mathematical purposes. For example, the sub-events of a plane flight and of a car trip seem more natural conditioning events for decisions than does, say, the event {plane arrives 3 hrs. late, auto arrives on time}. The peculiar structure of $\mathcal{E}$, in which all events are treated alike, may even be related to the violations of the sure-thing principle.

<u>Concluding Note</u>

In these lectures we have discussed a number of different examples of fundamental measurement. They are summarized in the following diagram, in which the reductions used to prove the representation and uniqueness theorems are indicated by arrows.

		conditional expected utility theory
		↓
		n-dimensional additive conjoint measurement
		↓
bisymmetry structures	⟶	2-dimensional additive conjoint measurement
		↓
conditional probability	⟶	positive difference structures
		↓
unconditional probability	⟶	extensive structures

This is not the only possible hierarchy of axiom systems. Pfanzagl, for example, reduces all systems he investigates to bisymmetry structures.

Selected Literature

ad 1. Extensive Measurement

1. Representations of weakly ordered systems

Birkhoff, G. Lattice theory. Americ. Math. Soc. Colloquium publication XXV, 1948, 1967.

Cantor, G., Beiträge zur Begründung der transfiniten Mengenlehre, Math. Ann., 1895, 46, 481-512.

Debreu, G., Representation of a preference ordering by a numerical function. In R.M. Thrall, C.H. Coombs and R.L. Davis (eds.) Decision processes. New York: Wiley, 1954. Pp. 159-165.

2. Hölder's Theorem and extensive measurement, classical

Fuchs, L. Partially ordered algebraic systems. Reading, Mass.: Addison-Wesley, 1963.

Hölder, O. Die Axiome der Quantität und die Lehre von Mass. Ber.D.Säch., Geselsch.D.Wiss., Math.-Phy. Klasse, 1901. 53, 1-64.

Suppes, P. A set of independent axioms for extensive quantities. Portugaliae Math., 1951, 10, 163-172.

3. Hölder's Theorem and extensive measurement, generalized

Behrend, F.A., A contribution to the theory of magnitudes and the foundation of analysis. Math. Zeit., 1956, 63, 345-362.

Krantz, D.H., Luce, R.D., Suppes, P., and Tversky, A. Foundations of measurement, in preparation, 1968, (probable date of publication, 1970).

Luce, R.D. and Marley, A.A.J. Extensive measurement when concatenation is restricted and maximal elements may exist. In S. Morgenbesser, P. Suppes, and M.G. White (eds.) Essays in honor of Ernest Nagel. New York: St. Martins Press, 1968. In press.

Pfanzagl, J. Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens. Schrift.d.Stat.Inst.d.Univ.Wien, Neue Folge Nr. 1, 1959.

4. Extensive measurement, discursive

Campbell, N.R. Physics: the elements. Cambridge: Cambridge Univ. Press, 1920. Reprinted as Foundations of Science: the philosophy of theory and experiment. New York: Dover, 1957.

Campbell, N.R. An account of the principles of measurement and calculation. London: Longmans, Green, 1928.

Ellis, B. Basic concepts of measurement. Cambridge: Cambridge Univ. Press, 1966.

5. Extensive measurement with a non-connected order

Giles, R. Mathematical foundations of thermodynamics. New York: MacMillan, 1964.

Roberts, F. and Luce, R.D. Axiomatic thermodynamics and extensive measurement. Synthese, 1968, in press.

6. Measurement inequalities

Gale, D. The theory of linear economic models. New York: McGraw-Hill, 1960.

Scott, D. Measurement models and linear inequalities. J. Math. Psychol., 1964, 1, 233-243.

ad 2. Qualitative Probability

1. Probability axioms

Feller, W. An introduction to probability theory and its applications. New York: Wiley, Vol.I, 1950, second edition: 1957.

Kolmogorov, A.N. Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlin: Springer, 1933. Translated into English as Foundations of the theory of probability. New York: Chelsea, 1956.

2. Qualitative probability

de Finetti, B. La prévision: ses lois logiques, ses sources subjectives. Annales de l'Institut Henri Poincaré, 1937, 7, 1-68. Translated into English in H.E. Kyburg and H.E. Smokler (eds) Studies in subjective probability. New York: Wiley, 1964. Pp. 93-158.

Kraft, C.H., Pratt, J.W., and Seidenberg, A. Intuitive probability on finite sets. Annals Math.Stat. 1959, 30, 408-419.

Luce, R.D. Sufficient conditions for the existence of a finitely additive probability measure. Annals Math.Stat., 1967, 38, 780-786.

Savage, L.J. Foundations of Statistics. New York: Wiley, 1954.

Savage, L.J. The foundations of statistics reconsidered. Proc. 4th Berkeley Symp.Math. and Prob. Berkeley: Univer.Calif.Press, 1961. Reprinted in H.E. Kyburg,jr., and H.E. Smokler (eds) Studies in subjective probability. New York: Wiley, 1964. Pp. 171-188.

Villegas, C. On qualitative probability σ-algebras. Annals Math.Stat., 1964, 35, 1787-1796.

3. Qualitative conditional probability

Koopman, B.O. The axioms and algebra of intuitive probability. Annals of Math., 1940, 41, 269-292, (a).

Koopman, B.O. The bases of probability. Bull.Amer. Math.Soc., 1940, 46, 763-774. Reprinted in H.E. Kyburg,jr. and H.E. Smokler (eds) Studies in subjective probability. New York: Wiley, 1964, Pp. 159-172, (b).

Koopman, B.O. Intuitive probability and sequences. Annals of Math., 1941, 42, 169-187.

Luce, R.D. The numerical representation of qualitative conditional probability. Annals Math.Stat., 1968, 39, 481-491.

ad 4. Conjoint Measurement

1. Additive

Aczel, J., Pickert, G., and Rado, F. Nomogramme, Gewebe und Quasigruppen. Mathematica, 1960, 2, 5-24.

Blaschke, W., and Bol, G., Geometrie der Gewebe. Grundlehren math. Wiss., 1938, 9.

Campbell, B.A., and Masterson, F.A. Psychophysics of punishment. In B.A. Campbell and R.M. Church (eds) Punishment. New York: Appleton-Century-Crofts, 1968, in press.

Debreu, G. Topological methods in cardinal utility theory. In K.J. Arrow, S. Karlin, and P. Suppes (eds.) Mathematical Methods in the social sciences, 1959. Stanford: Stanford Univ. Press, 1960. Pp. 16-26.

Krantz, D.H. Conjoint Measurement: the Luce-Tukey axiomatization and some extensions. J. math. Psychol., 1964, 1, 248-277.

Luce, R.D. Two extensions of conjoint measurement. J. math. Psychol., 1966, 3, 348-370.

Luce, R.D., and Tukey, J. Simultaneous conjoint measurement: a new type of fundamental measurement. J. math. Psychol., 1964, 1, 1-27.

Scheffé, H. The analysis of variance. New York: Wiley, 1959, Pp. 95-96.

Tversky, A. Additivity, utility, and subjective probability. J. math. Psychol. 1967, 4, 175-201.

2. Polynomial

Cliff, N. Adverbs as multipliers. Psychol. Rev., 1959, 66, 27-44.

Krantz, D.H., and Tversky, A. The diagnosis and testing of simple three-variable decomposition models for ordinal data. In preparation.

Logan, F.A., Decision making by rats: delay vs. amount of reward. J.comp. Physiol. Psychol., 1965, 59, 1-12.

Roskies, R., A measurement axiomatization for an essentially multiplicative representation of two factors. J. math. Psychol., 1965, 2, 266-276.

Tversky, A., A general theory of polynomial conjoint measurement. J.math. Psychol., 1967, 4, 1-20.

ad 5. Bisymmetry Systems

Pfanzagl, J. Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens. Schrift.d.Stat.Inst.d.Univ.Wien, Neue Folge Nr. 1, 1959.

ad 6. Conditional Expected Utility

1. Surveys of Utility Theory

Fishburn, P.C. Decision and value theory. New York: Wiley, 1964

Fishburn, P.C. Utility Theory. Management Sci., 1968, 14, 335-378.

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

Luce, R.D. and Suppes, P. Preference, utility and subjective probability. In R.D. Luce, R.R. Bush and E. Galanter (eds) Handbook of mathematical Psychology. Vol. III. New York: Wiley, 1965. Pp. 248-410.

2. Unconditional Theories

Davidson, D. and Suppes, P. A finitistic axiomatization of subjective probability and utility. Econometrica, 1956, 24, 264-275.

Pfanzagl, J. A general theory of measurement - applications to utility. Naval Research Logistics Q., 1959, 6, 283-294.

Raiffa, H. and Schlaifer, R. Applied Statistical decision theory. Boston: Harvard Univ. Press, 1961.

Ramsey, F.P. Truth and probability. In F.P. Ramsey, The foundations of Mathematics and other logical essays. New York: Harcourt, Brace, 1931, Pp. 156-198. Also in H.E. Kyburg, Jr., H.E. Smokler (eds) Studies in subjective probability. New York: Wiley, 1964, Pp. 61-92.

Savage, L.J. The foundations of statistics. New York: Wiley, 1954.

Suppes, P. and Winet, Muriel. An axiomatization of utility based on the notion of utility differences. Management Sci., 1955, 1, 259-270.

Von Neumann, J. and Morgenstern, O. Theory of games and economic behavior. Princeton: Princeton Univers. Press, 1944, 1947, 1953.

3. Conditional Theories

Luce, R.D. and Krantz, D.H. Conditional expected utility theory, 1968. Econometrica, in press.

Pfanzagl, J. Subjective probability derived from the Morgenstern-Von Neumann utility concept. In M. Shubik (ed) Essays in mathematical Economics in honor of Oskar Morgenstern. Princeton: Princeton Univ. Press, 1967, Pp. 237-251.