Annals of Operations Research 19(1989)205 – 217

PREFERENCE INTENSITY MEASUREMENT

Peter H. FARQUHAR

Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, USA

and

L. Robin KELLER

Graduate School of Management, The University of California at Irvine, Irvine, California 92717, USA

Abstract

The concept of preference intensity has been criticized over the past sixty years for having no substantive meaning. Much of the controversy stems from the inadequacy of measurement procedures. In reviewing the shortcomings of existing procedures, we identify three objectives for developing a satisfactory procedure: (1) the capability of validating expressed preference differences by actual choices among naturally occurring options, (2) compatibility with the existing problem structure, and (3) no confounding of extraneous factors in the measurement of preference intensity. Several recently developed measurement procedures are criticized for failing one or more of these objectives. We then examine three different approaches for measuring preference intensity based on multiple perspectives. The *replication approach* emerges as a promising way of satisfying the three objectives above. This methodology applies to problems where an attribute can be replicated by "parallel components" that are independent, identical copies of the attribute. We illustrate the approach with two applications reported in the decision analysis literature. We also offer guidance on how to construct parallel components satisfying the requisite properties.

1. Introduction

The idea of *preference intensity* (or "strength of preference") was introduced by Pareto [37] and Frisch [18]; preference intensity involves the comparison of preference differences. For example, let w, x, y, and z denote levels of some attribute X. We might then compare one's degree of preference for x over w with one's degree of preference for z over y. This quaternary relation is denoted by $>^*$. Thus, $wx >^* yz$ implies the difference in preference in going from w to x is greater than the difference in preference in going from y to z. The concept of preference intensity has been criticized over the past sixty years for having no substantive meaning (e.g.

Fishburn [15], p. 81). Several procedures nevertheless have tried to operationalize the preference intensity concept. The purpose of this paper is to describe the limitations of well-known methods for measuring preference intensity and to examine new methods that purport to overcome these limitations under appropriate circumstances.

The motivation for investigating preference intensity measurement comes from recent theoretical contributions that rely on such measures in both riskless and risky decision making. Moreover, the classical notion of marginal utility in economics requires the comparison of preference differences. A critical issue now is the empirical examination of these theories to determine if satisfactory measurement procedures are available (Dyer and Sarin [7], p. 821). Indeed, Sarin ([42], p. 344) notes: "... Unless the conceptual problem of the meaning and measurement of strength of preference is satisfactorily resolved, this theory will only see limited commercial application".

Much of the controversy about strength of preference measurement hinges on two issues. Many decision analysts report that strength of preference judgments are easy to obtain and often simplify the utility analysis in a decision problem (e.g. von Winterfeldt and Edwards [50], pp. 208–211). On the contrary, others contend that the questions used in obtaining such judgments are meaningless. Consider the question: "For what value of i is your intensity of preference for \$$i$ over \$100 the same as your intensity of preference for \$100 over \$50?" Machina ([34], p. 169) states: "Personally, I would respond to this question by asking what it meant. Would I rather obtain \$100 after having hoped for \$$i$ or obtain \$50 after having hoped for \$100? Surely the former – \$100 is better than \$50 regardless of i. Would I prefer \$100 to an even chance of \$$i$ or \$50? ..." Moreover, in White's ([53], pp. 334–335) opinion, individuals simply cannot make judgments about preference differences.

One root of these objections is that preference difference judgments do not correspond to any real or hypothetical choice behavior. Although individuals obviously do provide answers to questions comparing preference differences, there appear to be few, if any, ways of validating such judgments with actual behavior. Another concern is that many procedures add an extraneous element (e.g. money or risk) to a decision problem to measure strength of preference. Finally, factors present in the decision problem might be inextricably bound together with preference intensity in some measurement situations.

Therefore, we introduce three criteria for evaluating procedures that measure strength of preference. First, the procedure should be *actionable*, that is, preference difference judgments must be capable of being revealed through some actions of the decision maker. Procedures relying on only introspective judgments are subject to biases and have other problems that may not be easily detected (e.g. Poulton [39], Lyons [33]). Second, the procedure should be *compatible* with the existing structure of the decision problem; no extraneous elements should be introduced to accommodate a particular measurement procedure. Third, the procedure should give a measure of strength of preference that is *unconfounded* with other factors present in the decision problem. Examples of procedures that violate this third criterion are described in sect. 3.

The paper proceeds as follows. Section 2 reviews several well-known measurement methods and discusses their shortcomings. Section 3 then examines a number of recent methods for measuring preference intensity and evaluates them on the three criteria above. Section 4 considers the circumstances under which particular procedures for measuring preference intensity can satisfy all of the criteria. The *replication approach* is then described and illustrated with two applications reported in the decision analysis literature. The paper concludes with a brief summary and some suggestions for further research.

2. Background

Under appropriate conditions on the relation $>^*$, one can obtain an interval-scaled measure v on X such that $wx >^* yz$ if and only if $v(x) - v(w) > v(z) - v(y)$. This v is called a *measurable value function.* Axioms on $>^*$ that yield this interval-scaled measure of preference intensity on X are described by Suppes and Winet [44], Fishburn [15], Krantz et al. [31], and others. Our focus here, however, is not on the axioms for preference intensity but rather on measurement procedures. Previous articles by Fishburn [14,16], Johnson and Huber [26], Kneppreth et al. [29,30], and Dyer and Sarin [9] review a number of procedures for assessing measurable value functions using preference difference judgments. Common procedures involve variations of (1) direct estimation, (2) willingness-to-pay, or (3) lotteries.

2.1. DIRECT ESTIMATION OF PREFERENCE INTENSITY

In a wide variety of applications (e.g. Edwards [10]), decision makers are asked to provide a direct rating of value $v(x)$ to reflect both ordinal preferences and strength of preference. Direct estimation methods include numerical ratings on anchored scales (Torgerson [45]), constant sum paired comparisons (Hauser and Shugan [23]), visual plots of positions on a numerical scale (Kneppreth et al. [29,30]), and several others (see Torgerson [45], Fishburn [14], Johnson and Huber [26], Farquhar [13]). A major advantage of these direct estimation methods is their ease of implementation. These methods typically instruct the decision maker to give ratings such that the differences correspond to the judged differences in preference intensity. Analogous methods require judgments about the ratios of differences (e.g. Galanter [19], Edwards [10], Hauser and Shugan [23]). For example, one might be asked to specify x such that the degree of preference for x over w is twice the degree of preference for z over y.

Other methods of direct estimation are based upon psychophysical measurement (Torgerson [45], Farquhar [13]). One commonly used method is the "bisection" technique. This method usually begins with the extreme elements of a continuous X, say a and b, and asks for the element x such that the preference differences ax and xb are equal. By bisecting each successive interval in this manner, one can approximate the measurable value function. Another method develops a "standard sequence" by constructing a series of steps equal in preference difference. Thus, one specifies $x_1, x_2, x_3, x_4, \ldots$ such that $x_1 x_2$ equals $x_2 x_3$, $x_2 x_3$ equals $x_3 x_4, \ldots$.

A major criticism of these direct estimation methods is that they are not actionable – questions eliciting preference differences are not formulated as choices between naturally occurring options. We return to this point again in later sections. Fishburn ([15], p. 82) notes: "this direct, introspective 'measurability' pill" is difficult to swallow unless one makes some accommodation in the theory. At the least, one has to address issues such as nontransitive indifference, response biases, and external validity in using introspective methods (Poulton [39], von Winterfeldt and Edwards [50]). The analogy with direct judgments of sensory magnitudes may not provide a satisfactory measurement of preference intensity.

2.2. THE WILLINGNESS-TO-PAY METHOD

Suppose the problem structure includes two attributes X and P; the second attribute is conveniently interpreted here as "price". Each option is thus described by the levels attained on these two attributes. Suppose a person is indifferent between the option of having x at price p and the option of having y at price q. This indifference is denoted by $(x, p) \sim (y, q)$. Krantz et al. ([31], p. 142) observe that if these two attributes satisfy an additive conjoint structure, then one can indirectly compare differences in preference on each attribute (see also Suppes and Winet [44], Dyer and Sarin [7], and Wakker [51]). For instance, one can construct an interval scale of preference intensity over X by "pricing out" equivalent differences on P: the person above appears "willing to pay" the difference $q - p$ to exchange x for y. This "willingness-to-pay" or "pricing-out" procedure is described in more detail in Raiffa [40], Dyer and Sarin [7], von Winterfeldt and Edwards [50], and Merkhofer [35].

The willingness-to-pay method is criticized on two grounds. First, there are many problems that involve decisions with just one attribute. The introduction of an extraneous second attribute may not be compatible with the original problem structure. For example, the measurement of human mortality or morbidity via a monetary attribute can raise strong objections among individuals who view money as inappropriate in this decision context or who simply refuse to make such tradeoffs. The second criticism is that the additivity condition implicitly assumed in the usual implementation of this procedure may not hold in a given problem. Since this condition is not always checked in practice, the opportunity exists for serious misapplication of the willingness-to-pay procedure. Raiffa [40], Keeney and Raiffa [27], French [17], Merkhofer [35], and many others, discuss further the advantages and drawbacks of willingness-to-pay methods.

2.3. THE LOTTERY METHOD

Suppose a person prefers having x for sure rather than the lottery that gives an even chance of receiving either w or z, where z is preferred to w. One interpretation of this statement is that the degree of preference for x over w is greater than the degree of preference for z over x, denoted $wx >^* xz$ (e.g. Harsanyi [22], von Winterfeldt and Edwards [50]). Using this interpretation, a series of lottery comparisons can be used to elicit a measurable value function.

When a person compares lotteries, judgments will generally be based both on attitude toward risk and on strength of preference for different attribute levels (e.g. Ellsberg [11], Fishburn [16], Dyer and Sarin [9]). If the measurable value function $v(x)$ and expected utility function $u(x)$ are identical, a strength of preference interpretation can be given to the responses. If the two functions are not identical, the only way to obtain strength of preference information from lottery comparisons is to know the specific relationship between $u(x)$ and $v(x)$ (Dyer and Sarin [9]). Keller [28] found that experimental subjects had different measurable value and utility functions. Barron et al. [2], on the other hand, found a linear relationship provided the best fit between utility and value functions in their experiment.

The lottery method therefore has two criticisms. Fishburn [15] notes that this method introduces artificial uncertainty into riskless problems. The lottery method is thus incompatible with many problem structures and can complicate the assessment task in other situations (e.g. Dyer et al. [6], Pliskin and Beck [38], and von Winterfeldt and Edwards [50]). If risk is already part of the problem structure, then the relationship between the measurable value function and the utility function must be known. For instance, Sarin [42] has established a "substitution of equal exchange condition" that is both necessary and sufficient for $v(x) = u(x)$. Such conditions should be tested in practice.

3. Recent methods for preference intensity measurement

The previous section criticizes three well-known approaches to measuring preference intensity. The direct estimation methods are not validated by choice behavior. The willingness-to-pay methods often require that a second attribute be additively independent of the one being measured. The lottery method complicates a problem by introducing risk and is appropriate only under very restrictive conditions.

Recently developed methods for measuring preference intensity try to overcome the limitations of earlier methods. The *repetition approach* creates identical, repeated situations to allow the simultaneous consideration of two or more levels of a given attribute. Preference intensity information is then inferred from the exchanges a person is willing to make across the separate situations. The *agent approach* imposes a higher authority on the problem structure to facilitate the comparison of preference differences. Both approaches are examined below.

3.1. REPETITION METHODS

Camacho [4,5] presents a novel approach for measuring strength of preference. This approach requires that the decision maker be able to imagine identical repetitions of a choice situation. For example, a person first imagines being presented with the choice set X (such as entrees at a particular restaurant) and choosing x (e.g. broiled swordfish). Then, the person imagines the same situation in which y (e.g. filet mignon) is now chosen. Denote these repeated choices by $[x, y]$. Camacho [4,5] establishes four axioms to measure strength of preference with these repeated choices; Vansnick [48,49] and Wakker [52] present similar sets of axioms for implementing this approach.

The *permutation axiom* states that the sequence in which imaginary choices occur is immaterial: $[x, y]$ is equivalent to $[y, x]$. The *independence axiom* states that the value derived from x does not depend on y, and conversely. The *repetition axiom* requires that if w is preferred to x, then n repetitions of w are preferred to n repetitions of x (e.g. with two repetitions, $[w, w]$ is preferred to $[x, x]$). The *rate of substitution axiom* is an Archimedean condition that assumes for every $x > w$ and $z > y$ in X there exist integers r and s such that $[x^r, y^s]$ balances $[w^r, z^s]$. The interpretation is that r gains of x over w exactly compensates for s losses of y over z.

The last axiom is a key part of the repetition approach. As an illustration, consider the problem of measuring your strength of preference for coupons worth different discounts on your next flight with a particular airline. Suppose you are indifferent between the following two sequences of imagined choices:

$$[x^r, y^s] = [\$300^3, \$0^2] = [\$300, \$300, \$300, \$0, \$0]$$

and

$$[w^r, z^s] = [\$150^3, \$500^2] = [\$150, \$150, \$150, \$500, \$500].$$

Then your preference intensity for \$300 over \$150 is presumably s/r, or 2/3, times your preference intensity for \$500 over \$0. A simpler procedure (analogous to the standard sequence method of direct estimation) is to ask for \$$y$ such that [\$300, \$$y$] balances [\$150, \$500].

Because the substantive meaning of a preference difference is not clear, the repetition approach in its current form is not testable with actual choices. Camacho ([4], p. 369) offers no operational definition and states: "We are not concerned with the empirical side of the problem of measurability of utility". Another issue is whether the permutation axiom is reasonable when simulating repeated choices in memory (e.g. see Nisbett and Ross [36], Tversky and Kahneman [47], Hogarth [24], and Isen et al. [25]).

3.2. AGENT METHODS

A related approach for measuring preference intensity uses a neutral agent who considers the perspectives of multiple individuals in a problem before making a decision (e.g. Dyer and Sarin [8], Bell [3]). For example, suppose the task is to compare the preference difference between exchanging w for x or exchanging y for z. We postulate two identical individuals: the first has w and the second has y. The agent (or "supra decision maker") must decide between either raising the first individual from w to x or raising the second from y to z. The key assumption is that the agent is fair and acts on only the difference in preference intensity (e.g. Harsanyi [21], Rawls [41], Grofman and Owen [20]). Pliskin and Beck [38] offer one application of this procedure; physicians measured the severity of illness by considering treatment allocations to improve the health of only a subset of patients.

The literature on social choice, however, suggests other factors might influence an agent's judgments in using this method. For instance, judgments based on a concern for "equity" might lead to choosing a pair of end positions for the two individuals (i.e. either w and z, or x and y) as similar as possible. Such judgments obviously lead to the confounding of preference intensity with a fairness factor.

On the other hand, one could introduce a "veil of ignorance", where the agent has an equal chance of occupying one of these two perspectives after the decision is made. Although this implementation presumably promotes fairness in the agent, it introduces risk into the decision problem as did the lottery method in sect. 2.3. Other variations of the agent approach exist, but they all seem to suffer from (1) the possible confounding of other factors with preference intensity in the agent's judgments, or (2) the introduction of risk over the multiple perspectives.

4. The replication method for measuring preference intensity

4.1. MEASUREMENT CRITERIA

Beginning with four sets of "commodities", w, x, y, and z, Alt [1] stated in 1936 the key question in measuring preference intensity: can one determine whether the increase of "utility" in exchanging w for x is greater than, equal to, or less than the increase in "utility" in exchanging y for z?

> . . . the main problem is still open, namely whether it is at all possible to make comparisons between the transitions of commodities by empirical observations. But I have hopes that within a short time this question will be answered in the affirmative. (Alt [1], p. 431).

The problem still remains to devise a satisfactory procedure for comparing preference differences. Based upon the discussion above, we have identified three main requirements for a measurement procedure. First, the procedure must be *actionable*, that is, capable of empirical test with actual choices. Introspective judgments about preference differences alone are generally not sufficient. Second, the procedure must be *compatible* with the existing problem structure. Procedures requiring either explicit or implicit judgments about risk are not compatible with situations that involve no uncertainty. Finally, the procedure must yield a measure of preference intensity *unconfounded* with any other factor that is not separately measurable by choices. In the agent method, for example, preference intensity was first confounded with equity and then confounded with risk.

A stumbling block for measurement is the substantive meaning of the exchange above. If the decision maker is to make a choice between exchanging w for x or exchanging y for z, then the individual must be in two distinct states, namely w and y, simultaneously. One accomplishes this objective either by combining the two states in a probability mixture, or by having some form of multiple perspectives in which the states can co-exist. The latter suggestion is the one pursued here.

Our focus is on problem structures where the decision can be naturally represented as a choice among parallel components that are independent, identical replications. This *replication approach* is illustrated below with two examples.

4.2. EXAMPLES OF THE REPLICATION APPROACH

Kulkarni et al. [32] used the idea of a spatial replication of choice situations to elicit several single-attribute measurable value functions in the evaluation of road improvement plans. The Kansas Department of Transportation commissioned Woodward-Clyde Consultants to help establish priorities for reconstruction of road and bridge segments throughout the state's highway system. Engineers considered road segments that were presumably identical on all important performance characteristics except "shoulder width". They were asked to indicate the level x such that they would be indifferent between increasing the shoulder width on one road segment from 0 to x feet, or increasing it on another road segment from x to 10 feet. The response was interpreted to mean that $v(x) - v(0) = v(10) - v(x)$.

The replication of road segments is actionable: judgments about road segments can be validated by actually choosing which improvement to make. However, a number of conditions must be met in the roads example for this replication procedure to be valid. First, the decision maker must envision two road segments that are not part of the same road (e.g. having road segments with different shoulder widths along the same road could be a safety hazard). In general, any two road segments in the judgment task must be independent. Second, the decision maker must believe the two segments are identical in all other aspects (e.g. if one route were more heavily traveled than another, then the corresponding segment might be more deserving of a wider shoulder). This assumption requires that the problem representation be completely specified. Third, the decision maker should recognize that the preference intensity measure on one attribute (e.g. shoulder width) is conditional on the fixed levels of other attributes (e.g. lane width) and could be affected by changes in these fixed levels. For practical reasons, one should determine whether the set of attributes describing road segments are "difference independent" (Dyer and Sarin [7], Wakker [51]).

Dyer et al. [6] provide another example of the replication approach. In evaluating the educational goals of alternative curricula, elementary school principals considered average performances on nationally standardized tests. For instance, a test on creativity was represented by two separate, but equivalent parts. A sample question was: "which increase would be worth more to you – part A going from 50 to 60 percentile or part B going from 70 to 85 percentile?"

The replication of sub-tests provides an actionable measurement procedure: the principal could be asked to choose between two curricula that would yield the exchange above. On the other hand, confounding factors might influence a principal's choice. For instance, one principal might seek to avoid a large variance in two parts of the same test, while another principal may want to raise whichever test score is lower. Likewise, one could argue above that "uniformity of road segments" or "minimum highway system performance" might be factors in the engineer's decision.

The lesson here is simple: there is no free lunch. Although the replication approach allows the direct comparison of preference differences, it also introduces the possibility of interaction among the multiple perspectives. The interaction can occur in the form of a meta-rule, equity, or other factors that are necessarily confounded with preference intensity. If the precise form of the interaction were known, it might be possible to infer a preference intensity measure from comparisons of preference differences, but there is no guarantee. Therefore, we must be sure that the replications are not only *identical* in all aspects, but are also *independent* in the sense of an additive conjoint structure. The roads example probably satisfies these conditions, while the educational example might fail the independence condition.

4.3. APPLYING THE REPLICATION APPROACH

Two characteristics of the decision structure are essential for developing a satisfactory preference intensity measurement procedure with the replication approach. First, if one expects to have a procedure capable of empirical testing with actual choices, the structure must allow multiple perspectives on the attribute being measured. This structure can occur in three ways: multiple attributes, temporal repetitions, or parallel components. If preference intensity is not to be confounded, the second essential characteristic of the problem structure is independence of the multiple perspectives.

We now consider each of the three cases. Suppose first that the decision problem has multiple attributes. If the attributes are additively independent, preference intensity measurement is straightforward (see sect. 2.2). However, if the attributes are not independent, some transformation of the attributes might possibly remove the interdependencies (see Farquhar [12]).

For example, Keeney and Raiffa ([27], p. 53) discuss a problem where attributes A and B measure crime in two separate areas of a city. These two attributes are not independent because conditional preferences over A depend on the fixed level of B (and conversely). By restructuring the problem so that $C = A + B$ and $D = |A - B|$, the transformed attributes C and D are likely to be independent. In the context of measuring strength of preference, this reformulation can isolate the effect of D so it will not confound the measurement of $A + B$.

Suppose the problem structure does not possess multiple, independent attributes. The temporal repetition approach suggested by Camacho [4] creates identical, repeated choice situations to compare preference differences across situations. In its current form, the approach is not actionable. Moreover, temporal order effects might confound preference differences: imagining a choice in the first situation likely alters perceptions and judgments in the second situation. Thus, the permutation axiom is in doubt. For these reasons, the repetition approach is probably not viable.

In the absence of multiple independent attributes, the only alternative is to try to create a choice situation with parallel components. These components must be independent, identical replicas of the attribute being measured. The task of creating identical components within a given attribute may not always be solvable, though we have given a few examples (i.e. patients, road segments, and sub-tests). This task might be helped with methodologies from non-standard logics (e.g. Turner [46]).

The next task is to ensure that the components are independent. As noted above, there are many ways in which interactions can occur. The challenge is to test the components for independence and, if necessary, seek transformations that identify or eliminate interactions (Farquhar [12]). In some cases, interactions such as equity, variety, or balance are quickly spotted, but nevertheless can be stubbornly confounded with strength of preference.

4.4. SUMMARY AND CONCLUSIONS

Drawing on applications by Dyer et al. [6], Kulkarni et al. [32], and others, we propose a replication approach for measuring preference intensity. This approach has the advantage of giving substantive meaning to expressed preference differences, because judgments can be revealed through choices between naturally occurring options. One disadvantage of the approach is that its applicability is limited to problem structures where a given attribute can be represented by independent, identical replicas. These so-called "parallel components" of the attribute must ordinarily be constructed by the decision analyst to have the requisite properties. If independence is not checked, other factors can be inadvertently confounded with the preference intensity measure.

Other procedures for measuring preference intensity were criticized on the basis of three criteria. In general, one seeks a measurement procedure that is *actionable* in the sense of revealed choices, *compatible* with the problem structure, and *unconfounded* with extraneous factors. Many procedures fail on one or more of these criteria.

We suggest further research consider the practical limitations of the replication approach in constructing parallel components of an attribute. For instance, it would be helpful to identify problem domains where the independence condition is likely to hold. In other problems where the condition fails initially, it would be helpful to know what transformations might yield independence. Finally, it would be useful to characterize those problem domains resistant to preference intensity measurements in that no form of multiple perspective approach apparently works.

Acknowledgements

We gratefully acknowledge support for this research from the Office of Naval Research under Contracts #N00014-80-C-0897, #N00014-84-K-0558, and #N00014-87-K-0201. We appreciate Michael H. Birnbaum's thoughtful comments on a very early draft of this paper. We presented a subsequent version of the paper at the TIMS/ORSA Meeting in San Francisco in May 1984. We also appreciate helpful discussions on this research topic with James S. Dyer and Rakesh K. Sarin.

References

[1] F. Alt, On the measurability of utility, Zeitschrift für Nationalokonomie 7(1936)161. Translated and reprinted in: *Preferences, Utility and Demand,* ed. J.S. Chipman, L. Hurwicz, M.K. Richter and H.F. Sonnenschein (Harcourt Brace Jovanovich, New York, 1971) pp. 424–431.

[2] F.H. Barron, D. von Winterfeldt and G.W. Fischer, Empirical and theoretical relationships between value and utility functions, Acta Psychologica 56(1984)233.

[3] D.E. Bell, Assessing single attribute measurable value functions, Working Paper, Harvard Business School, Boston, MA (1982).

[4] A. Camacho, Approaches to cardinal utility, Theory and Decision 12(1980)359.

[5] A. Camacho, Cardinal utility and decision making under uncertainty, in: *Foundations of Utility and Risk Theory with Applications,* ed. B.P. Stigum and F. Wenstop (Reidel, Dordrecht, 1983) pp. 347–370.

[6] J.S. Dyer, W. Farrell and P. Bradley, Utility functions for test performance, Manag. Sci. 20 (1973)507.

[7] J.S. Dyer and R.K. Sarin, Measurable multiattribute value functions, Oper. Res. 27(1979)810.

[8] J.S. Dyer and R.K. Sarin, Group preference aggregation rules based on strength of preference, Manag. Sci. 25(1979)822.

[9] J.S. Dyer and R.K. Sarin, Relative risk aversion, Manag. Sci. 28(1982)875.

[10] W. Edwards, Use of multiattribute utility measurement for social decision making, in: *Conflicting Objectives in Decisions,* ed. D.E. Bell, R.L. Keeney and H. Raiffa (Wiley, New York, 1977) pp. 247–275.

[11] D. Ellsberg, Classic and current notions of measurable utility, Economics Journal 64 (1954)528.

[12] P.H. Farquhar, Interdependent criteria in utility analysis, in: *Multiple Criteria Problem Solving*, ed. S. Zionts (Springer-Verlag, New York, 1978) pp. 131–180.
[13] P.H. Farquhar, Utility assessment methods, Manag. Sci. 30(1984)283.
[14] P.C. Fishburn, Methods of estimating additive utilities, Manag. Sci. 13(1967)435.
[15] P.C. Fishburn, *Utility Theory for Decision Making* (Wiley, New York, 1970).
[16] P.C. Fishburn, Cardinal utility: An interpretive essay, Rivista Internazionale di Scienze Economiche e Commerciali 23(1976)1102.
[17] S. French, *Decision Theory: An Introduction to the Mathematics of Rationality* (Ellis Horwood, Chichester, England, 1986).
[18] R. Frisch, Sur une problème d'économie pure, Norsk Mathematish Forenings Skrifter 1 (1926)1.
[19] E. Galanter, The direct measurement of utility and subjective probability, Amer. J. Psychology 75(1962)208.
[20] B. Grofman and G. Owen, eds., *Information Pooling and Group Decision Making* (JAI Press, Greenwich, Connecticut, 1986).
[21] J.C. Harsanyi, Cardinal utility in welfare economics and in the theory of risk taking, Journal of Political Economy 61(1953)434.
[22] J.C. Harsanyi, Bayesian decision theory and utilitarian ethics, Amer. Econ. Rev. 68(1978)223.
[23] J.R. Hauser and S.M. Shugan, Intensity measures of consumer preference, Oper. Res. 28(1980)278.
[24] R.M. Hogarth, ed., *Question Framing and Response Consistency* (Jossey-Bass, San Francisco, 1982).
[25] A.M. Isen, B. Means, R. Patrick and G. Nowicki, Some factors influencing decision-making strategy and risk taking, in: *Affect and Cognition, The Seventeenth Annual Carnegie Symposium on Cognition*, ed. M.S. Clark and S.T. Fiske (Lawrence Erlbaum, New Jersey, 1982) pp. 243 - 261.
[26] E.M. Johnson and G. Huber, The technology of utility assessment, IEEE Trans. on Systems, Man, and Cybernetics, SMC-7(1977)311.
[27] R.L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* (Wiley, New York, 1976).
[28] L.R. Keller, An empirical investigation of relative risk aversion, IEEE Trans. on Systems, Man, and Cybernetics 15(1985)475.
[29] N.P. Kneppreth, D.H. Gustafson, R.P. Leifer and E.M. Johnson, Techniques for the assessment of worth, Technical Paper 254, U.S. Army Research Institute for the Behavioral Sciences, Arlington, Virginia. Also in: *Health Status Indexes*, ed. R. Berg (Hospital Research and Education Trust, Chicago, Illinois, 1974).
[30] N.P. Kneppreth, W. Hoessel, D.H. Gustafson and E.M. Johnson, A strategy for selecting a worth assessment technique, Technical Paper 280, U.S. Army Research Institute for the Behavioral Sciences, Arlington, Virginia, 1978).
[31] D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky, *Foundations of Measurement*, Vol. 1 (Academic Press, New York, 1971).
[32] R. Kulkarni, S. Rothstein, A. Sicherman, F. Finn and V. Ozernoy, Development of a construction priority system, Final Report prepared for the Kansas Department of Transportation, Woodward-Clyde Consultants, San Francisco, CA (1981).
[33] W. Lyons, *The Disappearance of Introspection* (MIT Press, Cambridge, MA, 1986).
[34] M.J. Machina, "Rational" decision making versus "rational" decision modelling?, Journal of Mathematical Psychology 24(1981)163.
[35] M.E. Merkhofer, *Decision Science and Social Risk Management* (Reidel, Dordrecht, 1987).
[36] R. Nisbett and L. Ross, *Human Inference: Strategies and Shortcomings of Social Judgment* (Prentice-Hall, Englewood Cliffs, New Jersey, 1980).

[37] V. Pareto, *Manuel d'Economie Politique*, 2nd ed. (Giard, Paris, 1927).
[38] J.S. Pliskin and C.H. Beck, A health index for patient selection: A value function approach with application to chronic renal failure patients, Manag. Sci. 22(1976)1009.
[39] E.C. Poulton, Models of biases in judging sensory magnitude, Psychological Bulletin 86 (1979)777.
[40] H. Raiffa, *Decision Analysis – Introductory Lectures on Choices under Uncertainty* (Addison-Wesley, Reading, MA, 1968).
[41] J. Rawls, *A Theory of Justice* (Belknap Press, Cambridge, MA, 1971).
[42] R.K. Sarin, Strength of preference and risky choice, Oper. Res. 30(1982)982.
[43] R.K. Sarin, Measurable value function theory: Survey and open problems, in: *Essays and Surveys on Multiple Criteria Decision Making*, ed. P. Hansen (Springer-Verlag, New York, 1983).
[44] P. Suppes and M. Winet, An axiomatization of utility based on the notion of utility differences, Manag. Sci. 1(1955)259.
[45] W.S. Torgerson, *Theory and Methods of Scaling* (Wiley, New York, 1958).
[46] R. Turner, *Logics for Artificial Intelligence* (Ellis Horwood, Chichester, England, 1984).
[47] A. Tversky and D. Kahneman, The framing of decisions and the psychology of choice, Science 211(1981)453.
[48] J.-C. Vansnick, Strength of preference – theoretical and practical aspects, in: *Operational Research '84*, ed. J.P. Brans (North-Holland, Amsterdam, 1984) pp. 367–381.
[49] J.-C. Vansnick, Intensity of preference, in: *Toward Interactive and Intelligent Decision Support Systems*, Vol. 2, ed. Y. Sawaragi, K. Inoue and H. Nakayama (Springer-Verlag, New York, 1987) pp. 220–229.
[50] D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavioral Research* (Cambridge University Press, Cambridge, England, 1986).
[51] P. Wakker, Cardinal coordinate independence for expected utility, Journal of Mathematical Psychology 28(1984)110.
[52] P. Wakker, The repetitions approach to characterize cardinal utility, Theory and Decision 17(1986)33.
[53] D.J. White, *Operational Research* (Wiley, New York, 1985).