BERICHTE

aus dem

PSYCHOLOGISCHEN INSTITUT

der UNIVERSITÄT BONN

Band 27, Heft 1 (2001)

Take The Best, Dawes' rule, and compensatory decision strategies: A method for classifying individual response patterns

Arndt Bröder

Abstract

Strategy descriptions like the “Take The Best”-heuristic (G. Gigerenzer et al., 1991), the weighted additive rule, and the equal weight decision rule are competing theories on information integration in probabilistic inference tasks. Behavioral decision research is confronted with the problem of drawing conclusions about unobservable decision strategies from behavioral data. Although there has been considerable progress due to methodical traditions like “Structural Modeling” and “Process Tracing”, these paradigms have certain limitations in testing specific hypotheses about individual strategies. Some of these problems are summarized briefly. A deductive method for classifying individual response patterns is introduced. Predictions about regression coefficients are deduced from competing substantial hypotheses about strategies for probabilistic inferences. These can be tested at the level of individual participants. The validity of this classification procedure is demonstrated with a Monte Carlo simulation. Some useful applications of the method are described, limitations of the method and potential generalizations are discussed.

Zusammenfassung

Strategien wie die “Take The Best”-Heuristik (G. Gigerenzer et al., 1991), das gewichtete additive Modell oder die “Equal Weight Decision Rule” stellen konkurrierende Theorien über die Informationsintegration in probabilistischen Inferenzaufgaben dar. Die so genannte “Behavioral Decision Research” oder “Deskriptive Entscheidungsforschung” steht vor dem Problem, aus Verhaltensdaten Rückschlüsse auf nicht beobachtbare kognitive Prozesse ziehen zu müssen. Obwohl die beiden dominierenden Forschungstraditionen “Structural Modeling” und “Process Tracing” viel zum Fortschritt der deskriptiven Entscheidungsforschung beigtragen haben, weisen sie einige Probleme auf, die hier kurz zusammengefasst werden. Eine Methode zur Klassifikation individueller Entscheidungsmuster wird vorgestellt, die auf Deduktionen aus den substanziellen Hypothesen beruht. Aus den Hypothesen über kognitive Strategien werden Vorhersagen für Regressionsgewichte abgeleitet. Die Hypothesen können so auf individueller Ebene geprüft werden. Eine Simulationsstudie demonstriert die Validität der Methode. Es werden einige Anwendungen geschildert und Probleme der Methode diskutiert.

Take The Best, Dawes' rule, and compensatory decision strategies: A method for classifying individual response patterns[1]

Like any other field in cognitive psychology, the special branch called "Behavioral Decision Research" (BDR; e.g. Maule & Svenson, 1993; Payne et al., 1992) is concerned with formulating theories about cognitive processes. In the case of BDR, process models are invented to describe human thinking in various situations in which judgments or decisions have to be made. Sometimes, the BDR branch is characterized as *descriptive* as opposed to *normative* models of decision making. However, the term "descriptive" might have a somewhat misleading connotation in this context because the aim is not merely to describe actual data, but to describe *cognitive processes* which are hypothetical constructs within cognitive theories. Of course, these theoretical claims have to be confronted with actual data in order to test their validity and predictive power. Linking unobservable theoretical constructs to potentially observable data patterns is the central problem of psychology in general, and it can be termed the "measurement problem".

In BDR, this measurement problem has been tackled by the invention of a variety of methodical developments that try to link data to theory. In general, these attempts can be divided into two broad classes that are often referred to as *Structural Modeling* (e.g. Brehmer, 1994; Dawes, 1979) and *Process Tracing* (e.g. Payne, 1976; 1982; Payne *et al*., 1988; 1992; 1993), respectively. Structural Modeling "typically focuses on the end result of a decision process and tries to relate the final decision to parameters characterizing the decision problem" (Svenson, 1983, p. 140). Various methods might be subsumed under this class, for example Anderson's (1981; 1982) Information Integration Theory or the Brunswikian tradition that resulted in the development of Social Judgment Theory (e.g. Brehmer, 1988). Process Tracing, on the other hand, "directly assesses what information was accessed to form a judgment and the order in which the information was accessed. This information is used to make inferences about what decision strategies have been used in arriving at a choice" (Ford et al., 1989, p. 76). Thus, both approaches aim at testing theories about decision strategies, but they rely on different types of data. The information derived from both types of data can be viewed as complementary, as Einhorn et al. (1979) have acknowledged. Both traditions have proved to be of enormous value in contributing to our

[1] I am grateful to Edgar Erdfelder, Jürgen Bredenkamp and Ingo Wegener for helpful comments on an earlier version of this manuscript.

understanding of human decision making in multi-attribute situations (see e.g. Brehmer, 1994; Slovic & Lichtenstein, 1971; Payne et al., 1993, for some reviews).

While the merits of these approaches are out of question, there have been several repeated methodical criticisms concerning both traditions which limit their general applicability (e.g. Westenberg & Koele, 1994). Especially in the context of detailed and precise theories about cognitive processes, the mere adaption of approved methods may not be adequate because they are often not suitably tailored to address the specific theory under investigation. In these cases, new methods have to be developed in order to achieve strict tests of existing theories. On the one hand, this conclusion seems obvious, but on the other hand, researchers often stick to traditional methods, arguing about their *general usefulness* instead of asking whether they are able to solve a *specific* question. (Of course, there are also numerous exceptions to this assertion, e.g. Aschenbrenner et al., 1984; Böckenholt & Kroeger, 1993; Busemeyer & Townsend, 1993; Huber, 1983; Wallsten & Barton, 1982; to name just a few).

This article deals with the so-called "Take The Best"-heuristic which was proposed by Gerd Gigerenzer and his colleagues as a boundedly rational cognitive model of multiple-cue probabilistic inferences (Gigerenzer et al., 1991; Gigerenzer & Goldstein, 1996; 1999). Despite being a precise model, empirical consequences are not easy to obtain from its formulation, that is, no acceptable decision criteria exist for deciding whether a person adopted the strategy or not. Such a criterion will be developed here by deriving precise expectations about the structure of regression weights, given a specific set of experimental stimuli. This particular application is intended to demonstrate the power of a deductive way of reasoning in BDR that can supplement the existing tool-kit of research procedures. Whereas the particular method developed here is confined to the case of the Take-The-Best-heuristic and the well-known "Equal Weight Linear Model" or "Dawes' rule", the general line of reasoning is not.

The paper is structured as follows: First, the Take-The-Best-heuristic will be described along with the methodical problem of deriving empirical predictions from it. Second, criticisms of the above-mentioned traditions of BDR will be summarized briefly. Most of these problems have been acknowledged before, but a synopsis and some new arguments might be helpful. This is not done in order to devalue these approaches in general, but to demonstrate their limitations in the case of specific theory tests. Third, the classification method for individual response patterns will be described. A Monte Carlo simulation will be reported to demonstrate its validity. In addition, some successful applications will be described briefly. In the discussion section, several limitations of the method as well as potential remedies are discussed.

The Take-The-Best-heuristic and problems of empirical tests

The Take-The-Best-heuristic (hereinafter: TTB) is a simple model of information integration in binary probabilistic inference tasks. Its original formulation appeared within the theory of *Probabilistic Mental Models* proposed by Gigerenzer et al. (1991) which was designed to explain the choice process *and* the calibration of confidence judgments in general knowledge tasks. TTB has now become a central building block of the research program on "Simple heuristics that make us smart" (Gigerenzer et al., 1999). Gigerenzer and his colleagues have shown that simple noncompensatory heuristics that violate traditional assumptions about rational decision making may perform equally well as more "rational" compensatory heuristics like the weighted additive model, for instance (Gigerenzer & Goldstein, 1996; 1999; see also Johnson & Payne, 1985; Thorngate, 1980). The success of these heuristics in combination with their simplicity render them plausible psychological models of decision making.

TTB is intended to describe the cognitive processes in dealing with binary probabilistic inference tasks. In such a task, an inference must be made with respect to an unknown value of a target variable. For instance, the general knowledge item "Which city has more inhabitants: Chicago or Detroit?" might be asked. Most people will not be able to retrieve the information about city populations with certainty, so they have to rely on *probability cues* that are believed to be correlated to the target variable (e.g. the existence of an international airport). The predictive power of such a cue with respect to the target variable is called its *ecological validity*, defined as the conditional probability of drawing the correct inference when one object possesses the critical feature while the other does not, and the former object is chosen (see Björkman, 1994; Gigerenzer et al., 1991). Drawing on Brunswik's (1956) assumption of well-adapted cognitive systems, Gigerenzer et al. (1991) assume that ecological cue validities are known to observers who had repeated experience with the environment. Obviously, given a set of cues, the correct decision could be made with maximum probability if a complex Bayesian computation were applied. This, however, is not plausible as a cognitive model, because the processing demands increase exponentially with the number of cues (Oaksford & Chater, 1993). According to the TTB assumption, people just look up the most valid cue. If this cue discriminates, they choose the object with the critical attribute, and thinking stops. In case of equal cue values for both options, the next most valid cue is examined and so on. It is obvious that this heuristic is noncompensatory because the most valid discriminating cue will determine the choice without potential revision of that decision due to less valid cues. In addition, TTB is a special case of the lexicographic decision rule (e.g. Fishburn, 1974).

The assumption of such a simple strategy used by actual respondents is a bold one. Other strategies like a weighted additive model or a simple equal weight model for combining cue values may be candidates of comparable plausibility. Therefore, an *empirical* evaluation of the TTB hypothesis is indispensable in order to assess its adequacy as a theory of probabilistic inferences. However, strict empirical tests of the TTB hypothesis require bridging the gap between theory and data which is quite problematic in this case. The problems are discussed in detail in Bröder (2000a). Here they will be only briefly summarized.

The first problem deals with the undefined universality and precision of the TTB hypothesis. Gigerenzer and Todd (1999) explicitly state that TTB is one of a larger set of strategies people are equipped with. This implies the possibility of different strategies of different (same) persons in the same (different) situations. Given this possibility of interindividual (and intersituational) differences, a method for testing the TTB hypothesis at the *individual level* is indispensable. Group statistics would be misleading in such a case. On the other hand, the deterministic formulation of TTB is unrealistic. Even if the strategy is simple, there may be occasional processing errors or response errors that might influence the data. Rejecting the hypothesis of a TTB strategy because of some random responses that do not fit the predicted pattern would be extremely unfair. Hence, the inclusion of an explicit random response error model is necessary.

The second problem can be termed the “problem of separation”. Consider a set of items consisting of two alternatives each. These objects are characterized by binary cue values. One can now compare the response vectors that are predicted by various decision strategies, such as TTB, Dawes’ rule or a weighted additive model. The striking fact is that one will find an enormous overlap of the predictions from these models (almost 92% identical predictions of all strategies in the city population environment, the “drosophila” environment analyzed by Gigerenzer & Goldstein, 1996; 1999; see Bröder, 2000c). This overlap of predictions is obviously a severe problem for theory testing because no single choice can be attributed to a specific strategy. If occasional response errors are considered, the problem gets even worse.

The following conclusions can be drawn from the above-mentioned methodical problems: The evaluation of the TTB hypothesis is only possible when a criterion for hypothesis testing at the individual level is at hand because individual differences in strategy use have to be expected (see Brehmer, 1994; Brehmer & Brehmer, 1988; Slovic & Lichtenstein, 1971). This method must be able to solve the separation problem in order to test hypotheses about specific strategies. Before introducing such a method, some problems of traditional methodical approaches will be briefly summarized. A detailed discussion can be found in Bröder (2000a).

Structural Modeling and Process Tracing

In this section, the traditional approaches of BDR will be briefly examined for their potential to solve the above-mentioned problems concerning strict tests of the TTB hypothesis at the individual level.

The dominating approach in the Structural Modeling area is the so-called "Social Judgment Theory" which is based on Brunswik's (1956) lens model. The preferred method consists of mapping a judgment vector onto a matrix of cue values via multiple regression procedures. A good model fit in terms of the multiple correlation is seen as indicative for a compensatory strategy (e.g. Brehmer, 1994; Einhorn et al., 1979). The regression coefficients, on the other hand, yield information about the utilization of cues, that is, their relative importance in forming the judgments (see Stewart, 1988, for a discussion). However, a perfect fit of linear regression equations can be achieved by noncompensatory strategies as well if these can be formalized as linear models. Thus, while equating the terms "linear" and "compensatory" may be appropriate in most cases, this need not always be true. As will be shown below, TTB (a noncompensatory rule) is equivalent in performance to a linear integration model. The fit of a linear regression model will therefore not be able to differentiate between different cognitive models in our case.

According to our intuition, the magnitudes of the regression coefficients should reflect cue importance and therefore are expected to reproduce the rank order of cue validities. However, as a simulation of Hoffrage et al. (1997) has shown, this is *not* necessarily the case. Hoffrage et al. generated a sample of TTB decisions from their city populations environment. In the subsequent regression analysis of the response vectors, the regression weights did not conform to the intuition. The most valid of nine cues obtained rank five in a rank order of regression weights. Two problems hamper the interpretation of the regression weights here: First, the use of a binary dependent variable might distort the "true" state of affairs, and second, the regression weights are not only a function of cue validity, but also depend on the discrimination rates in the stimulus sample (see Gigerenzer & Goldstein, 1996) and the intercorrelations of the cues. This must caution us from interpreting regression weights as measures of cue importance within this paradigm. To summarize: The successful regression methodology of Social Judgment Theory cannot readily be adapted to the empirical problem of testing the TTB hypothesis without modification.

Following the pioneering work of John Payne and his coworkers (Payne, 1976; 1982; Payne et al., 1988; 1992; 1993), the Process Tracing methodology from problem solving research has been applied to multi-attribute decision research with tremendous success. Sequences of information acquisition are monitored by think-aloud protocols (e.g. Einhorn et al., 1979; Harte et al., 1994), eye movements (e.g. Russo & Dosher, 1983), or the information board technique in which hidden attribute information has to be actively uncovered by the participant by turning cards or opening information boxes on the computer screen (Payne et al., 1988). Reviews of different techniques of Process Tracing are provided in Abelson and Levi (1985), Svenson (1979), and van Raaij (1983). Several measures for various aspects of the search are created, representing the depth, the selectivity, and the sequence of information search which indicate the use of more noncompensatory strategies when information search shows less depth, greater selectivity, and more attribute-wise processing than alternative-wise processing. However, even if these measures allow for a classification of strategies as more or less noncompensatory, it is difficult to relate them to hypotheses about *specific* strategies (e.g. the weighted additive model or Dawes' rule). For instance, Harte et al. (1994) derived predictions about processing sequences from different strategies. Unfortunately, none of the actual sequences derived by think-aloud protocols perfectly fitted the predicted sequences. Unless explicit error models are formulated, the classification of such sequences remains problematic (see the discussion of this point in Harte et al, 1994, p.113, and in Einhorn et al., 1979; p.481). Another problem - although more theoretical - is the lack of a strict implication connecting information integration and information search. An individual could show a "compensatory" information search pattern and nevertheless adopt a noncompensatory information *integration* strategy. The correspondence assumption of search and integration may be reasonable in most cases (see Abelson & Levi, 1985; p. 256 f.), but a test of e.g. the TTB hypothesis might be attenuated by (unknown) violations of this assumption. Thus, Process Tracing data alone cannot guarantee a strict test.

Another problem arises from the fact that the results of the two methodical traditions seem to imply different conclusions concerning human judgment and decision making. Structural Modeling studies tend to corroborate the assumption of the preponderance of compensatory strategies (e.g. Brehmer, 1994), whereas Process Tracing studies tend to imply the opposite conclusion. This apparent contradiction will not be discussed here (see Bröder, 2000a, for details). Some studies comparing the two approaches tend to favor the assumption that both methods might focus on different phases of the decision process (Billings & Marcus, 1983; Maule, 1994). However, for this reason, it is recommended here to combine the methods whenever possible in order to cross-validate the results, or at least in order to enrich the data base.

The considerations above lead to the conclusion that the traditional paradigms of BDR do not provide tools for the test of the TTB hypothesis, at least when they are applied without modification.Therefore, another method must be developed which is specifically tailored to handle this problem. In the next section, a method for testing the TTB hypothesis at the individual level is introduced. Statistical hypotheses are deduced from the substantive hypothesis, allowing for testing the latter by means of statistical tests. The validity of the method is tested by means of a Monte Carlo simulation. The subsequent sections describe successful applications and discuss potential extensions of the method that might overcome some limitations.

A deductive, idiographic procedure for classifying decision strategies

Following Einhorn's (1970) suggestion to use an "idiographic approach" in BDR, a method for evaluating hypotheses for individual respondents is necessary. That is, hypotheses about the use of strategies have to be tested at an individual level in order to classify decision patterns. This approach will be adopted here.

One obvious problem is the formulation of the rules: most decision strategies are formulated in a deterministic manner. Taken seriously, this deterministic formulation would render theory-testing extremely simple because most rules imply a certain vector of choices. Whenever behavior does not match this vector exactly, the hypothesis that this strategy was used could be considered wrong. Many psychologists would probably agree that this test would be unfair because the notion of completely error-free responding is not very realistic in psychology. Most theories assume -explicitly or implicitly- some error component in responses. When comparing response vectors with predicted vectors, we are not interested in *unsystematic* deviations due to processing errors, but we want to detect *systematic* deviations caused by the application of a different strategy. The methodical implication is obvious: For any decision model, a random error model reflecting unsystematic response errors has to be specified. As a consequence, any test of the hypothesis that a specific rule was administered will test the conjunction of rule plus error model.

The strategies tested: "Take The Best" and "Dawes' rule"

The TTB heuristic is formulated for inductive inferences concerning binary choices based on a set of dichotomous cues. According to TTB, people form an internal cue hierarchy with respect to the predictive power (called "validity") of each cue. The "best" cue is then examined whether it discriminates between two objects. If it does, one object is chosen without reference to other cues. If the best cue does not discriminate, the next cue is examined and so on. TTB is a special case of the noncompensatory lexicographic rule (Fishburn, 1974). Another simple and effective decision rule can be called "Dawes' rule" (Gigerenzer, Czerlinski, & Martignon, 1999; Dawes, 1979) which is an equal weight linear model (hereinafter: EWL). People are thought to count positive cue values for every object regardless of cue importance. The object with the higher number of positive cues will be chosen. It has been repeatedly demonstrated that this strategy is fairly accurate as compared to linear models with "optimal" weights (Dawes, 1979; Dawes & Corrigan, 1974; Wainer, 1976). However, EWL is less "frugal" than TTB because all cue information has to be searched for in memory in order to reach a conclusion.

As has been outlined above, even if these strategies are simple, we cannot expect people to use them completely error-free. We will have to specify an error model which defines unsystematic deviations from the rule. In the demonstration presented here, the simplest error model will be considered: It will be assumed that if a certain strategy is used, a participant will make an error with probability α. That is, in an expected proportion of $\alpha * 100$ percent of her choices, the participant will chose the object *not* favored by the strategy. This error probability is assumed to be uniformly distributed over all possible comparisons of objects. Of course, the principle of deductive tests of models is not confined to this particular error model, others might be considered more appropriate in some situations. In the next section, predictable consequences will be deduced from TTB and EWL in conjunction with the error model.

Deducing implications of the models

The logic of deriving predictions is as follows: the observable data pattern we get from every participant is a vector of choices. In principle, if nonsystematic response errors are allowed in the decision process, every possible choice vector can be generated by every decision rule, so we cannot derive deterministic implications from assumed strategies for the observable data patterns. Rather, we must find a statistical criterion to decide on the hypothesis that a particular strategy was used. That is, the goal is to derive implications for a statistical null hypothesis from the hypothesis that certain heuristics were employed. Therefore, the formal properties of the decision rules must be examined. The rules mentioned above (TTB and EWL) can be conceptualized as linear integration models (with a specific weight structure) on a latent decision variable Θ that is mapped onto a binary choice variable. Whereas "linearity" is obvious in the case of EWL, it is counterintuitive in the case of the noncompensatory TTB-heuristic. However, Martignon and Hoffrage (1999) have shown, that TTB is equivalent (in performance) to a weighted additive model with a noncompensatory weight structure (see also Gigerenzer, Czerlinski, & Martignon, 1999).

As was demonstrated by Hoffrage et al. (1997), a regression analysis of this binary variable does not necessarily yield the expected weight structure. However, if certain conditions are met, then a derivation of expected regression weights is straightforward.

For simplicity, the derivations for TTB will be demonstrated in accordance with the experimental situation in which $n = 4$ binary cues (coded "0" and "1" for the absence or presence of the cue feature, respectively) were used. With n binary cues there are $N = 2^n$ possible cue-patterns, yielding 16 patterns in the case of four cues. A complete set of paired comparisons between all 16 patterns consists of $N^2/2 - N/2 = 120$ pairs.

If the lexicographic TTB rule is consistently applied, the 16 cue-patterns consisting of "0" and "1" will be ordered lexicographically. The pattern (1,1,1,1) will have the highest rank in this hierarchy because it dominates all other patterns. The next patterns will be (1,1,1,0), (1,1,0,1), (1,1,0,0) and so on until pattern (0,0,0,0) which is the lowest in rank because it is dominated by all other patterns. If we denote the rank of each pattern j in this hierarchy with R_j (ranging from 0 to 15), then Equation (1) holds.

$$R_j = 8x_{j1} + 4x_{j2} + 2x_{j3} + 1x_{j4} \qquad (1)$$

In Equation (1), the x_{ji} denote the values of cue i for pattern j. As can be seen, the rank in the lexicographic hierarchy can be perfectly expressed as a linear function of the cue values (see Martignon & Hoffrage, 1999). The coefficients (i.e. the weights) are *noncompensatory*. That is, the weight of a more valid cue can never be exceeded by the sum of weights of less valid cues. It is easy to see that Equation (1) must be true when the 16 patterns are hierarchically ordered in a table. Consequently, for each comparison of any two patterns j and k, Equation (2) will result:

$$\Delta R_{j,k} = (R_j - R_k) = 8 * (x_{j1} - x_{k1}) + 4 * (x_{j2} - x_{k2}) + 2 * (x_{j3} - x_{k3}) + 1 * (x_{j4} - x_{k4}) \quad (2)$$

The differences in cue-values of patterns j and k will be denoted as $c_{ijk} = (x_{ji} - x_{ki})$ resulting in $c_{ijk} \in \{-1, 0, 1\}$. The TTB rule defines a nonlinear transformation of this rank difference variable, yielding the choice variable *TTB* according to the decision rule given in Equation (3).

$$\left[TTB_{j,k} = (-1) \Leftrightarrow \Delta R_{j,k} < 0\right] \wedge \left[TTB_{j,k} = 0 \Leftrightarrow \Delta R_{j,k} = 0\right] \wedge \left[TTB_{j,k} = 1 \Leftrightarrow \Delta R_{j,k} > 0\right] \quad (3)$$

This means that whenever object j is superior to object k in the lexicographic hierarchy, object j is chosen (coded "1"), when the opposite is true, object k is chosen (coded "-1"). Note, that this formal representation of the TTB-heuristic is not the same as the cognitive model. The cognitive model of sequential cue-wise testing does not involve the subjective representation of a lexicographic hierarchy. Nevertheless, the strategy is *formally equivalent* to this representation. If two patterns are the same (which is not possible within the paired comparisons considered here) the TTB rule will not result in a choice (coded "0"). Note that this coding scheme of *TTB* suggested in Equation (3) is arbitrary. However, combining Equations (2) and (3), we can write the variable *TTB* as a function of the cue-value-differences c_{ijk}. Consider a complete paired comparison of all 16 cue-patterns: Every element of the *TTB*-vector is completely determined by the c_{ijk}. There is another arbitrary aspect in Equation (3): We are free to choose which cue-pattern is denoted as "j" or "k", respectively, in any comparison, resulting in different values of the c_{ijk} and *TTB*, dependent on that choice. This does not change the formal structure of Equations (2) and (3), but this will lead to different structures of cue-intercorrelations and correlations of the cues with *TTB*. Obviously, across all 120 paired comparisons, there are 2^{120} possible coding schemes that will lead to different correlational structures. We can now choose one of these schemes that has the following properties: (1) all cue-means are zero, (2) all cue variances are equal, and (3) all cue-intercorrelations are zero. If these conditions are met, then it can be shown that the correlations of *TTB* with the c_i will show the structure in Equation (4).

$$r_{c1,TTB}=2*r_{c2,TTB}=4*r_{c3,TTB}=8*r_{c4,TTB} \tag{4}$$

Furthermore, if the TTB-rule is consistently applied, but "contaminated" with a constant, but unknown, error probability α, then the *expected* values of these correlations are given by Equation (5).

$$E(r_{c1,TTB})=2*E(r_{c2,TTB})=4*E(r_{c3,TTB})=8*E(r_{c4,TTB}) \tag{5}$$

The proof of these assertions can be found in the Appendix. There one can also find a description of the coding method that yields the necessary properties to derive this prediction.

Together with the conditions mentioned above, the derived structure of correlations has direct implications for the regression weights if a multiple regression of the variable *TTB* on the independent variables c_i is performed: Whenever the predictors in a regression equation are uncorrelated (condition 3), the standardized regression coefficients β_i are equal to the bivariate correlations of the predictors and the dependent variable (Cohen & Cohen, 1983, p. 101), and hence, they must show the same noncompensatory structure. The unstandardized coefficients B_i depend on the standardized coefficients and the predictor variances (see Cohen & Cohen, 1983; p. 100). As these variances are equal (condition (2)), it follows for the B_i:

$$E(B_1)=2*E(B_2)=4*E(B_3)=8*E(B_4) \tag{6}$$

We can apply the same logic of reasoning to the EWL-rule, and will find that if a choice vector was generated by this strategy (including response errors with probability α), then the expected structure of regression weights follows Equation (7).

$$E(B_1)=E(B_2)=E(B_3)=E(B_4) \tag{7}$$

The two different choice strategies imply different structures of expected regression weights when an observed choice vector is analyzed in a multiple regression with the cue-differences as predictors. Therefore, it is possible to test the hypotheses that a person applied the TTB-heuristic or the EWL-heuristic at the individual level by testing the null hypotheses about the regression weights. This approach does not rely on any surface indicators of "noncompensatory" behavior, and the regression analysis is not seen as a "model" of the decision process, but merely as a statistical tool to test hypotheses about correlations, that were derived form the substantial models. In this respect, the procedure can be classified as "deductive".

Hypothesis tests

Testing hypotheses about regression weights is straightforward and can be achieved by comparing the model fit of an unrestricted regression model to the fit of an appropriately restricted model. The unrestricted model is given in Equation (8).

$$\hat{Y} = \hat{B}_1 c_1 + \hat{B}_2 c_2 + \hat{B}_3 c_3 + \hat{B}_4 c_4 + \hat{B}_0 \tag{8}$$

The null hypotheses implied by TTB and EWL are given in Equations (9) and (10), respectively.

$$\hat{Y} = \hat{B}_1 c_1 + \tfrac{1}{2} \hat{B}_1 c_2 + \tfrac{1}{4} \hat{B}_1 c_3 + \tfrac{1}{8} \hat{B}_1 c_4 + \hat{B}_0 \tag{9}$$

$$\hat{Y} = \hat{B}_1 c_1 + \hat{B}_1 c_2 + \hat{B}_1 c_3 + \hat{B}_1 c_4 + \hat{B}_0 \tag{10}$$

A statistical test of these null hypotehses is straightforward in linear regression: If the restricted models in (9) and (10) explain significantly less variance than the unrestricted model in (8), then the corresponding null hypothesis is rejected.

Classification procedure

In the preceding section we have developed two null hypotheses reflecting the expected patterns of regression weights resulting from the choice rules TTB and EWL, respectively, when these are combined with a simple error model. For every person completing all 120 paired comparisons, we can test these hypotheses by applying the above-mentioned multiple regression models. On the basis of these tests, the participant's behavior will be classified according to a simple decision rule. If we can reject the TTB-hypothesis while the EWL-hypothesis is not rejected, the participant will be classified as "EWL" because her decision behavior is compatible with the EWL rule, but not with TTB. Consequently, if the reversed pattern emerges, the participant will be classified as "TTB". If *both* hypotheses are rejected, we can assume that the participant used another weighting scheme which is presumably compensatory but does not follow an equal weight rule ("COMP"). Therefore, a simple statistical test can tell us whether the observed data pattern is compatible with the assumption of certain strategies plus response errors.

An undesirable situation emerges when *neither* hypothesis is rejected. This would mean that the data were compatible with two completely different weighting schemes. However, these cases will only result when R^2 is low in either case and thus, the linear model is not a good model anyway. These cases will be called "unclassified". Although the possibility of "unclassified" patterns is not satisfying theoretically, they do not appear to play a significant role in empirical applications of the procedure. For instance, none of the 160 empirical choice vectors reported in Bröder (2000b, Experiments 2,3, and 4) remained "unclassified".

Classification criteria, power and the appropriateness of the F-Test

To test a restricted regression model with *k* predictors against an unrestricted model with m predictors the following *F*-statistic may be used (Cohen & Cohen, 1983):

$$F = \frac{R_u^2 - R_r^2}{1 - R_u^2} * \frac{N - 1 - m}{m - k} \tag{11}$$

R_u^2 and R_r^2 are the squared multiple correlation coefficients of the unrestricted and restricted models, respectively. Under the null hypothesis of equal *R*'s and other conditions (see below) this statistic is centrally *F*-distributed with $df_{\text{num}}=m-k$ and $df_{\text{den}}=N-1-m$. In the experiment described below, every participant completed all N = 120 paired comparisons, resulting in the appropriate *F*-statistic with df_{num}= 4-1 = 3 and df_{den}= 120-4-1=115 for the tests of the hypotheses. We will refer to the *F*-values as "F_{TTB}" and "F_{EWL}". What are our chances to detect deviations from the TTB model or EWL if we apply this *F*-test? A power analysis reveals that a "medium effect size" (f^2 = .15) according to the conventions from

Cohen (1988) can be detected with the error probabilities $\alpha = \beta = .05$ with this sample size.[2] The critical F is 2.68. According to these values, our error probabilities for misclassifying do not exceed .05 if the F-test is valid in this situation. Note that our effect size defines the alternative hypothesis: we do not consider deviations smaller than f^2=.15 to be "substantial" deviations from the null hypothesis.

The need for validation studies

Before the classification procedure outlined above is applied to substantial research problems, its validity has to be examined empirically. Two potential problems questioning the validity have to be addressed. First, a statistical prerequisite for administering the F-test is violated in this application of multiple regression (see below). That is, the theoretical misclassification probabilities might deviate from the actual ones. Second, the method might be insensitive to detect deviations from the null hypotheses and therefore overestimate the proportion of TTB users and EWL users in a sample. To address the first problem, a Monte Carlo simulation was conducted.

Nominal and actual error probabilities: a Monte Carlo simulation

The central F-distribution of the before-mentioned statistic can only be assumed when the assumption of independently and identically normally distributed residuals is fulfilled, which is obviously violated in our case of a dichotomous criterion. Whether this violation leads to a distortion of type-I error probability in testing the hypotheses must be examined in a robustness study. For this reason, large samples of response vectors were generated according to the TTB-rule and the EWL-rule, respectively. These response vectors were "contaminated" by various response error probabilities (5%, 10%, 20%, 30%) according to the simple error model. Both decision rules were combined with all four error rates, resulting in eight cells of the design. 10,000 vectors were generated for each cell. Additionally, 10,000 random vectors were generated. These can be considered a random sample of all possible 2^{120} response patterns. For all data patterns, three multiple regression analyses were performed to compute the values F_{TTB} and F_{EWL} according to Equation (11) on which the classification as "TTB", "EWL", "COMP" or "unclassified" was based

[2] The analysis was performed with the program G-Power for Macintosh Power PC (Buchner, Faul, and Erdfelder, 1996). The exact value of (1-β) is .9526.

like outlined above. The critical *F*-value was F_{crit} = 2.68 which stands for a nominal probability of a type-I error of α=0.05. The classification results are shown in Table 1.

Table 1: Results of the Monte-Carlo-Simulations. Data sets were generated by using the TTB rule or the EWL rule with error rates of 5%, 10%, 20%, and 30%, respectively. 10,000 data sets were generated in each of the eight resulting conditions and analyzed by multiple regression as described in the text.

		Percentage classified as...			
generating rule	error rate (%)	TTB	EWL	COMP	unclassified
random	--	1.47	1.15	3.49	**93,89**
TTB	5	**99.94**	--	0.06	--
	10	**98.99**	--	0.96	0,05
	20	**85.02**	0.48	2.33	12,17
	30	**40.24**	0.78	3.31	55,67
EWL	5	--	**99.66**	0.34	--
	10	--	**98.76**	1.21	0,03
	20	0.25	**87.71**	2.37	9,67
	30	0.80	**44.39**	2.83	51,98

Note: The theoretical critical F(3,115) for α=.05 is F_{crit}=2.68.

The first row of Table 1 shows the classification of the randomly generated response vectors that remained “unclassified” in 93.89% of the cases. This observation has two implications: First, if a person responds completely unsystematic, then her behavior will be erroneously classified as a systematic strategy in only about 6% of the cases. Second, as these randomly generated vectors are a random sample of all possible vectors, the result shows, that the classification as “TTB”, “EWL” or “COMP” is not trivial.

The classification procedure yields extremely high hit rates (>98%) for response vectors generated by EWL or TTB when the error rates are low or moderate (5% or 10%). However, this hit rate decreases to about 40% with an error rate of 30%. But note, that the decreasing hit rate does not result in a misclassification rate that exceeds the nominal α of 0.05. Merely, an increase of unsystematic errors in responding leads to more unclassified patterns, but not to a systematic bias in favor of another strategy. A response error rate of

30% is extremely high and can be seen as a limiting case to speak of “systematic” behavior at all.

Altogether, the simulation study shows that despite of the violation of its assumptions, the *F*-test is robust in the sense that it is not biased against the null hypotheses. The error probabilities of falsely rejecting the null hypotheses are less than the nominal value of 0.05. Therefore, the *F*-test tends to be conservative, rather than anticonservative, with respect to the null hypotheses.

In addition to this simulation study, Bröder (2000c, Experiment 4) has reported a validation experiment which aimed at showing that the classification method behaves rational when applied to actual behavioral data. Three experimental conditions were realized, in which participants were expected to respond “TTB-like”, “EWL-like” or according to a compensatory weighting scheme. This was achieved by manipulating the cue validities in the environment in a way that the corresponding strategy was optimal in each of the given situations. 90% of the participants (27 of 30) were classified as expected.

The simulation study as well as experiment show that the proposed classification procedure is rational in the sense that the classification of response vectors is not particularly error prone despite of some violations of assumptions necessary to justify the *F*-test. Most of the observed response patterns in the experiment were classified in a manner consistent with the strategies expected to be adopted by “rational” participants in the corresponding experimental conditions. These are two empirical arguments that support the validity of the deductive regression-based classification procedure outlined above.

Some applications of the procedure

The validation studies may have convinced some readers that the classification procedure is valid to some degree. However, this does not answer the question “What is it good for?”. To be a methodical progress, the procedure must be shown to offer the possibility of substantial insights in the decision research domain or at least in the “simple heuristics” perspective fostered by Gigerenzer et al. (1999). BDR is characterized by a growing interest in situational and task factors that might trigger the use of different strategies under varying conditions. Drawing on the contingency model by Beach and Mitchell (1978), Payne et al. (1993) emphasize the hypothesis that decision makers are “adaptive”, using different strategies contingent on the demands of a specific decision situation. For example, time pressure (see Svenson & Maule, 1993) or the dispersion of cue validities (Johnson & Payne, 1985; Payne et al., 1988) have been shown to have predictable effects on information search behavior, suggesting that decision makers choose appropriate strategies from a “tool-kit” of strategies they can use. This assumption is explicitly adopted by Gigerenzer and his colleagues (Gigerenzer & Todd, 1999). With a valid classification method of individual response strategies, an examination of the impact of task factors on strategy selection is possible. Furthermore, the identification of *individual* strategies in principle allows for linking these strategy preferences to other variables, such as cultural differences or personality traits. That is, a situational or task factor might lead to a strategy switch for most people, whereas some might retain a preferred strategy under varying conditions. The individual identification of strategies could be used to assess the relative importance of strategy preferences or personality traits versus situational factors in a truly interactive manner.

Although the above-mentioned interactive perspective has not been realized yet, a brief example may illustrate the value of the method. Bröder (2000b, Experiments 3 and 4) varied the costs for acquiring information about the cue values in a hypothetical stock market game. The participants had to choose one of two presented shares in each of the 120 decisions. The shares were described by positive or negative values on four cues with different validities. However, to get access to the cue information, participants had to acquire the cue information by clicking the appropriate buttons with the mouse. In one condition, this was the only “investment” to get access to that information (Experiment 4, condition 3). In two other conditions, participants had to “pay” either a low amount of virtual money for the cue-information (Experiment 3, condition 1) or a rather high amount of virtual money (Experiment 3, condition 2 and Experiment 4, condition 4). This money was subtracted from a virtual “private account”. Participants’ response vectors were classified with the regression procedure outlined above. While 65% percent of the participants were classified as TTB users in the high cost condition, only 40% were classified as TTB users in the low cost condition. In the no cost condition, the proportion of TTB users was only 15%. The influence of information costs on strategy use is highly significant ($\chi^2_{(2)} = 12.23$, $p< 0.01$) and turned out to be the most potent factor among others motivating participants to use TTB in this situation.

These applications imply several conclusions. First, they show that the method is capable of examining factors that lead to strategy switches contingent on the demands of the task, thus supporting the claim of “adaptive decision making”. Second, the results show that even in the presence of powerful situational factors like “information costs”, individual differences exist in the preferences for certain strategies. That is, not all respondents show the same strategy under identical conditions. These strategy preferences might be related to personality variables, such as processing capacity, motivation, rigidity, or whatever. Obviously, a valid classification of individual strategies opens the field for relating these observed preferences to such variables. This might lead to a clearer picture of the factors determining strategy selection. Third, the results show that TTB seems to be a strategy often used in probabilistic inference tasks, at least under certain conditions.

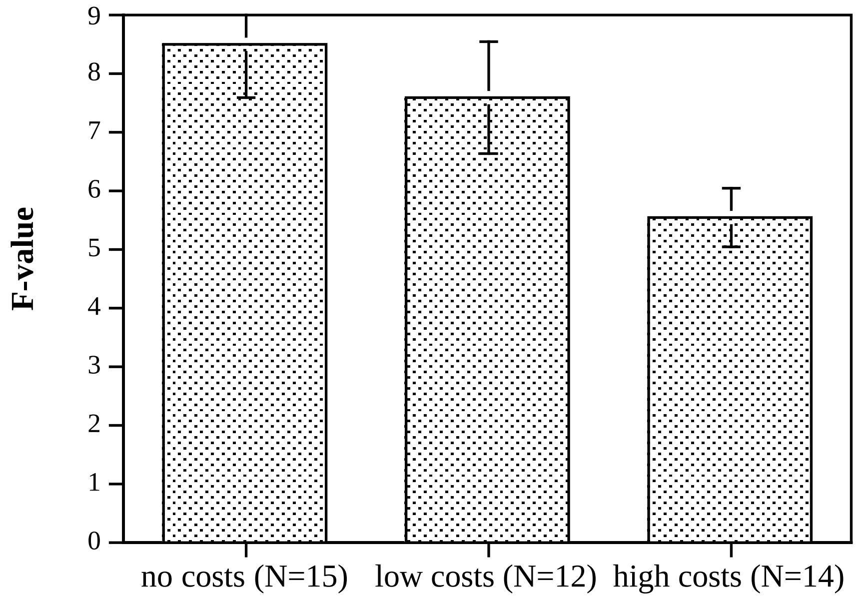

Figure 1: Means (and SEs) of the F_{TTB}-values as measures of “compensatoriness” for participants with a compensatory strategy in experimental conditions with no vs. low. vs. high costs of information acquisition (Bröder, 2000b).

Some researchers might feel uncomfortable with the apparently coarse-grained classification of strategies as TTB, EWL, or COMP (see e.g. Koele & Westenberg, 1995). The latter category is something like a rest category which contains all “compensatory” strategies that do not fit the TTB rule or an EWL rule, respectively. However, these patterns might reflect different degrees of deviation from the noncompensatory strategy and therefore different degrees of weighting less valid cues. The classification procedure provides such a quantitative measure of “compensatoriness”, namely the F_{TTB}-statistic on which the classification is based. Obviously, this F-value reflects the degree of deviation from the ideal noncompensatory strategy and can serve as a supplementary measure to analyse the within-strategy variation of cue-weighting.[3] For example, Fig. 1 depicts the mean F_{TTB}-values for all respondents who were classified as “COMP” in the above-mentioned experimental conditions (no vs. low. vs. high information costs). As one can easily see, the degree of compensatory cue weighting decreases with increasing information costs. An ANOVA of the F-values reveals that this difference is significant ($F(2,38)=3.65$, $p<0.05$). That is, even if there are participants who retain a compensatory strategy under conditions with information costs, their deviation from a noncompensatory strategy is less pronounced the higher these costs are. Therefore, in addition to the strict deductive hypothesis test about decision strategies, a more fine-grained analysis of the variation *within strategies* is possible.

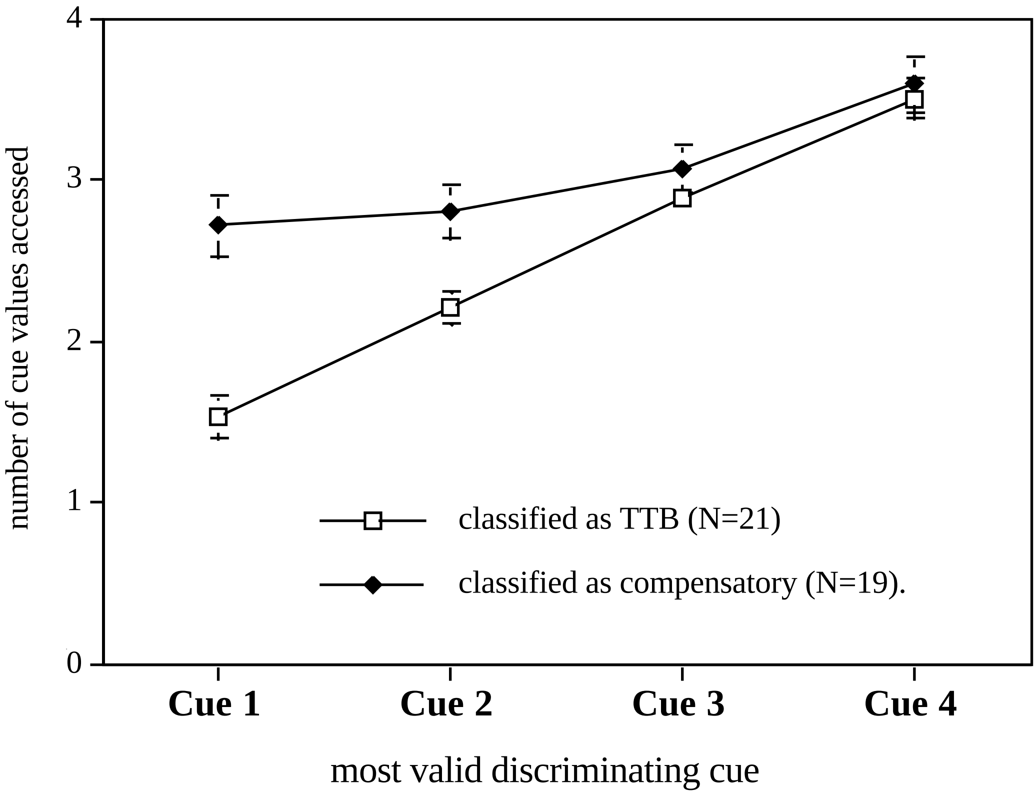

Figure 2: Mean number of cues acquired in Experiment 3 of Bröder (2000b) (Note: error bars represent SE of estimate)

[3] Comparing F-values as measures of deviation from the null hypothesis is only useful as long as the degrees of freedom are the same. Otherwise, one would prefer to use anoter effect size measure, e.g. Cohen’s (1988) f^2.

A last point should be illustrated here. With the possibility of a (valid) strategy classification of individual respondents, a cross-validation with process tracing measures can be achieved in principle. In Fig.2, the mean number of cues that were accessed by the respondents prior to their decision is plotted for those respondents who were classified as "TTB" and those who were classified as "compensatory", respectively, in Experiment 3 of Bröder (2000b). On the abscissa of Fig. 2, different decision situations are presented in which the most valid cue differentiated between the options (Cue1), the second most valid cue differentiated (Cue 2) and so on. As can be seen, the information acquisition of TTB users is extremely restrictive as compared to that of respondents classified as compensatory decision makers. When the mean number of cues acquired is compared for both strategy classifications, the difference is highly significant ($t(38) = 4.36$, $p < .001$). If, on the other hand, both experimental *conditions* are compared (low costs vs. high costs), the difference does not turn out to be significant ($t(38) = 1.75$, $p = .09$). The result shows that a clear difference in acquisition behavior exists *between strategies*. However, when *experimental conditions* are compared, this difference is attenuated by the fact that both groups are composed of a mixture of different strategies.[4]

If the numbers of cues acquired with each item type are used as predictors in a discriminant analysis, 85% of participants (34 of 40) are classified to the same strategy as with the regression-based procedure, showing some substantial overlap between acquisition data and decision data. Thus, Process Tracing data can be cross-validated by the classification method, and this allows for testing the correspondence assumption between decision and acquisition measures in every application of the method.

General Discussion

In this paper, a method for classifying individual response vectors for binary decisions was introduced. In 1970, Hillel J. Einhorn stated that "[i]f one can build a model of individual behavior and at the same time provide a theoretical classification of different 'types', then one

[4] Some readers may wonder why the upper line in Figure 2 is not completely parallel to the abscissa at the level of four cues (Compensatory decision makers always buy all information). This is a misunderstanding. Compensatory decision making does *not* imply full information search in every single situation. For example, consider choice situations in which the first two cues favour one alternative. A rational - and compensatory - decision in this case is *not to buy* further information of less valid cues because they cannot override the impact of the first two cues.

can combine an idiographic and a nomothetic approach." (Einhorn, 1970, p. 229). Obviously, large individual differences exist in applying decision strategies. Hence, an identification of these strategies can add valuable information to the analysis of group statistics which is normally done in Process Tracing studies. This has been demonstrated by the fact that grouping participants according to experimental conditions did not reveal a reliable difference in information acquisition behavior, whereas grouping respondents according to the strategy they presumably employed showed a large difference.

To summarize, the method developed here seems capable to detect individual differences. Its validity was demonstrated by a simulation study and a validation experiment (Bröder, 2000c). The usefulness of the method has been illustrated by an application which showed that, although task factors have an impact on strategy selection, the strategy switch is not perfect. In principle, the individual identification of strategies allows for relating strategy preferences to other variables, such as cultural or individual (trait) differences. Hence, it can be used to investigate the determinants of strategy selection more completely. The main methodological advantage is that hypothesis tests were derived directly from substantial hypotheses without any need to rely on validity assumptions for ad-hoc measures of compensatory and noncompensatory behavior (see Bröder, 2000b).

However, the classification method suffers from some limitations that will be discussed now. One apparent disadvantage is the restricted applicability of the method for testing the TTB hypothesis and the EWL hypothesis for multiple-cue probabilistic inferences, respectively. The method is specifically designed to address this question. Hence, it is not a general purpose tool for analyzing data in BDR. This limited applicability might decrease its value in the view of some researchers. Although "off-the-shelf", general purpose research procedures may be desirable for pragmatic reasons, we feel that this criticism rather points to the strength of the method than to its weakness. Specific research questions (e.g. whether people really adopt the TTB strategy) can often only be answered by the introduction of research methods specifically tailored to address this particular theory or hypothesis. The above-mentioned "problem of separation" hampered progress concerning the investigation of TTB until now because it could not be solved by other methods. According to the author's view, this approach should generally be adopted in BDR. Although Structural Modeling and Process Tracing are valuable general data analytic tools, their limitations for addressing specific issues require supplemental methodical developments. The method presented here may be seen as *one example* of such an approach.

A more severe limitation of the method, as presented here, is the necessity of a full paired comparison of all possible cue patterns. This is critical for two reasons: First, the restriction to four-cue problems seems unavoidable because even with five binary cues, 496 decisions would be required from each participant which is beyond the scope of a normal laboratory experiment (unless torturing participants is the goal of the study). Second, true Brunswikians will criticize that this restriction does not allow for constructing representative sets of stimuli. To put this more generally, no variation of the stimulus characteristics (except cue validities and semantic embedding) is possible. For instance, the question whether the number of dominated alternatives in a series of decisions has an influence on strategy selection (e.g. Johnson & Payne, 1985; Payne et al., 1988) cannot be readily addressed because the proportion of dominated alternatives is dictated by the need for a full paired comparison. Whereas these restrictions confine the range of potential applications, they do not call into question the *general logic* of the approach adopted here. The full paired comparison was introduced for technical reasons and did not appear to be a critical problem in the experiments reported above. With such a set of stimuli, a zero correlation of the appropriately coded cues can be achieved which allows for an easy derivation of to-be-expected regression weights (see Appendix). However, this does not preclude the possibility of deriving predictions for expected regression weights (or other statistics) when this technical assumption is not met. Presumably, these derivations will become more awkward, but they should be possible in principle for any subset of paired comparisons that is drawn from the class of all paired comparisons of cue patterns. Monte Carlo simulations to assess the robustness of the F-test are recommended for these applications. In most instances, the use of the full set of comparisons will not be a problem, allowing for an unmodified application of the method as presented here. Also, an extension to cases with more than two decision alternatives is a future desideratum.

The above-mentioned problems set the stage for further developments of the method, possibly resulting in a generalization and the use of test statistics that are more satisfying theoretically. The version introduced here is a starting point for developing methods that link theoretical claims about decision strategies to unambiguous empirical predictions. At the moment, the potential merits of the procedure outweigh its disadvantages.

Gigerenzer et al. (1999) have started a huge research program in which they examine a variety of short-cut heuristics with respect to their (surprisingly good) efficiency in “real world” environments. TTB is one prominent candidate within the tool-kit. However, the question of the descriptive adequacy of these heuristics as cognitive models of actual human behavior has not been addressed until now (but see Rieskamp & Hoffrage, 1999, for an exception). Some time ago, Gigerenzer (1981) pointed out that successful solutions of the measurement problem are indispensable for any progress of psychology as an empirical science. This is also true, of course, for his own research agenda on "simple heuristics that make us smart". Hopefully, the method developed here is a fruitful contribution to solving the problem in the case of simple decision heuristics.

References

Abelson, R. P. & Levi, A. (1985). Decision making and decision theory. In G. Lindzey and E. Aronson (Eds.), *Handbook of social psychology. Vol I. Theory and method (3rd edition*), (pp. 231-309). New York: Random House.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of information integration theory.* New York: Academic Press.

Aschenbrenner, K. M., Albert, D. & Schmalhofer, F. (1984). Stochastic choice heuristics. *Acta Psychologica, 56*, 153-166.

Beach, L. R. & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*, 439-449.

Billings, R. S. & Marcus, S. A. (1983). Measures of compensatory and noncompensatory models of decision behavior: process tracing versus policy capturing. *Organizational Behavior and Human Performance, 31*, 331-352.

Björkman, M. (1994). Internal cue theory: Calibration and resolution of confidence in general knowledge. *Organizational Behavior and Human Decision Processes, 58,* 386-405.

Böckenholt, U. & Kroeger, K. (1993). The effect of time pressure in multiattribute binary choice tasks. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 195-214). New York: Plenum Press.

Brehmer, B. (1988). The development of Social Judgment Theory. In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view*, (Vol. 54, pp. 13-40). Amsterdam: North-Holland.

Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica, 87,* 137-154.

Brehmer, A. & Brehmer, B. (1988). What have we learned about human judgment from thirty years of policy capturing? In B. Brehmer & C. R. B. Joyce , Eds.), *Human judgment: The SJT view* (pp. 75-114). Amsterdam: North-Holland.

Bröder, A. (2000a). A methodological comment on behavioral decision research. *Psychologische Beiträge,42,* 645-662.

Bröder, A. (2000b). Assessing the empirical validity of the “Take The Best”-heuristic as a model of probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26,* 1332-1346.

Bröder, A. (2000c). *”Take The Best - Ignore The Rest”. Wann entscheiden Menschen begrenzt rational?* [“Take The Best - Ignore The Rest”. When do people decide boundedly rational?]. Lengerich, Germany: Pabst Science Publishers.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments (2nd ed.)*. Berkeley: University of California Press.

Buchner, A., Faul, F., & Erdfelder, E. (1996). *G•Power: A priori, post-hoc, and compromise power analyses for the Macintosh* (Version 2.1.1) [Computer program]. Trier, Germany: University of Trier.

Busemeyer, J. R. & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*, 432-459.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.

Cohen, J. & Cohen, P. (1983). *Applied multiple regression / correlation analysis for the behavioral sciences (2nd ed.)*. Hillsdale: Erlbaum.

Czerlinski, J., Gigerenzer, G. & Goldstein, D. G. (1999). Accuracy and frugality in a tour of environments. In G. Gigerenzer, P. M. Todd and the ABC Research group (Eds.), *Simple heuristics that make us smart* (pp. 59-72). New York: Oxford University Press.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571-582.

Dawes, R. M. & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95-106.

Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin, 73*, 221-230.

Einhorn, H. J., Kleinmuntz, D. N. & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86*, 465-485.

Fishburn, P. (1974). Lexicographic order, utilities and decision rules: a survey. *Management Science, 20*, 1442-1471.

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M. & Doherty, M. L. (1989). Process tracing methods: contributions, problems, and neglected research problems. *Organizational Behavior and Human Decision Processes, 43*, 75-117.

Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie.[Measurement and modeling in psychology]* München: Reinhardt.

Gigerenzer, G. & Goldstein, D. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review, 103*, 650-669.

Gigerenzer, G. & Goldstein, D. G. (1999). Betting on one good reason: The Take The Best heuristic. In G. Gigerenzer, P. M. Todd & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75-95). New York: Oxford University Press.

Gigerenzer, G., Czerlinski, J. & Martignon, L. (1999). How good are fast and frugal heuristics? In J. Shanteau, B. A. Mellers, & D. A. Schum (Eds.), *Decision science and technology: Reflections on the contributions of Ward Edwards* (pp. 81-103). Norwell, MA: Kluwer Academic Publishers.

Gigerenzer, G., Hoffrage, U. & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98,* 506-528.

Gigerenzer,G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Harte, J. M., Westenberg, M. R. M. & van Someren, M. (1994). Process models of decision making. *Acta Psychologica, 87,* 95-120.

Hoffrage, U., Martignon, L. & Hertwig, R.(1997). *Does "Judgment policy capturing" really capture the policies?* Poster, presented at the 16th conference on Subjective Probability, Utility, and Decision Making (SPUDM); University of Leeds, August 18-21, 1997.

Huber, O. (1983). Dominance among some cognitive strategies for multidimensional decisions. In L. Sjöberg, T. Tyszka & J. Wise (Eds.), *Human decision making* (pp. 228-242). Bodafors, S: Doxa.

Johnson, E. J. & Payne, J. W.(1985). Effort and accuracy in choice. *Management Science, 31*, 395-414.

Koele, P. & Westenberg, M. R. M. (1995). A compensation index for multiattribute decision strategies. *Psychonomic Bulletin and Review, 2*, 398-402.

Martignon, L. & Hoffrage, U. (1999). Where and why is Take The Best fast, frugal, and fit? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd and the ABC Research group (Eds.), *Simple heuristics that make us smart* (pp. 119-140). New York: Oxford University Press.

Maule, A. J. (1994). A componential investigation of the relation between structural modelling and cognitive accounts of human judgement. *Acta Psychologica, 87,* 199-216.

Maule, A. J. & Svenson, O. (1993). Theoretical and empirical approaches to behavioral decision making and their relation to time constraints. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 3-25). New York, NY, US: Plenum Press.

Oaksford, M. R. & Chater, N. (1993). Reasoning theories and bounded rationality. In K. I. Manktelow & D. E. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 31-60). London: Routledge.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16,* 366-387.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534-552.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43,* 87-131.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Rieskamp, J. & Hoffrage, U. (1999). When do people use simple heuristics and how do we know this? In G. Gigerenzer, P. M. Todd and the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 141-167), New York: Oxford University Press.

Russo, J. E. & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 676-696.

Slovic, P. & Lichtenstein, S. (1971). Comparison of Bayesian and Regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance, 6,* 649-744.

Stewart, T. R. (1988). Judgment analysis: Procedures. In B. Brehmer and C. R. B. Joyce (Eds.), *Human judgment: The SJT view* (pp. 41-74). Amsterdam: North-Holland.

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance, 23,* 86-112.

Svenson, O. (1983). Decision rules and information processing in decision making. In L. Sjöberg, T. Tyszka and J. Wise (Eds.), *Human decision making* (pp. 131-162). Bodafors, S: Doxa.

Svenson, O. & Maule, A.J. (Eds.) (1993), *Time pressure and stress in human judgment and decision making*. New York: Plenum Press.

Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science, 25*, 219-225.

van Raaij, W. F. (1983). Techniques for process tracing in decision making. In L. Sjöberg, T. Tyszka and J. Wise (Eds.), *Human decision making* (pp. 179-196). Bodafors, S: Doxa.

Wainer, H. (1976). Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin, 83,* 213-217.

Wallsten, T. S. & Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 361-384.

Westenberg, M. R. M. and Koele, P. (1994). Multi-attribute evaluation processes: Methodological and conceptual issues. *Acta Psychologica, 87,* 65-84.

Appendix: Proof for noncompensatory correlations

Definitions:

Let $\underline{c}_{im} \in \{-1, 0, 1\}$ be the value of the cue-difference of cue $\underline{i}$ in a paired comparison m of cue patterns. c_1, c_2, c_3, and c_4 refer to the cues in the order of their importance (validity). Let N_i be the number of paired comparisons in which c_i discriminates, while none of the more important cues discriminates, i.e.:

$N_1 \equiv$ number of comparisons with $c_1 \neq 0$,

$N_2 \equiv$ number of comparisons with $c_1 = 0$ and $c_2 \neq 0$,

$N_3 \equiv$ number of comparisons with $c_1 = 0$ and $c_2 = 0$ and $c_3 \neq 0$, and

$N_4 \equiv$ number of comparisons with $c_1 = 0$ and $c_2 = 0$ and $c_3 = 0$ and $c_4 \neq 0$.

Necessary conditions:

The TTB rule defines a variable **T** dependent on the c_i and coded according to Equations (2) and (3) in the text. Suppose, a coding scheme can be found that fulfills the following conditions:

(1.) all bivariate cue correlations are zero (r_{cick}=0 for all $i \neq k$),

(2.) all cue variances are equal (S_{ci}=S_{ck} for all i,k),

(3.) all cue means are zero ($\bar{c}_i = 0$ for all i), and

(4.) the mean of **T** is zero ($\bar{T} = 0$).

Theorem:

If a person completes a full paired comparison of all 16 cue patterns employing the TTB rule with an unknown but constant error probability α, and her choices are coded according to Equations (2) and (3), and if, furthermore, the conditions (1.) to (4.) hold, then the expected correlations of the cues and T will follow the pattern $\mathrm{E}(\mathrm{r}_{\mathrm{Tc1}}) = 2\mathrm{E}(\mathrm{r}_{\mathrm{Tc2}}) = 4\mathrm{E}(\mathrm{r}_{\mathrm{Tc3}}) = 8\mathrm{E}(\mathrm{r}_{\mathrm{Tc4}})$.

Proof:

The correlation of a cue i with the variable T across N =120 paired comparisons m is defined as follows:

$$r_{Tci} = \frac{\sum_{j=1}^{N} (c_{im} - \bar{c}_i)(T_m - \bar{T})}{N * S_{ci} * S_T}.$$

As N, S_{ci} and S_T are constant for all i (condition 2), the following considerations can be confined to the cross-product in the numerator. For each cue i, two cases have to be considered:

Case 1: Cue i discriminates, while none of the more important cues discriminates

Case 2: All other cases

The expected cross-product can be determined for each case:

Case 1:

$$E\left[\sum_{Case\ 1} (c_{im} - \bar{c}_i)(T_m - \bar{T})\right] = N_i * 1 * (1-\alpha) + N_i * (-1) * \alpha = (1-2\alpha) * N_i$$

That is, in each case in which cue i is the most important discriminating cue (frequency of N_i), the expression $(c_{im} - \bar{c}_i)(T_m - \bar{T})$ will assume the value "1" when TTB is applied without error, whereas it assumes the value "-1" when an error is made with probability α.

Case 2:

If cue i does not discriminate, another cue will determine the choice. As all cues have zero correlations (condition 1), it follows that.

$$E\left[\sum_{Case\ 2} (c_{im} - \bar{c}_i)(T_m - \bar{T})\right] = 0$$

Combining Case1 and Case 2, the expected value of the cross-product is $(1-2\alpha)*N_i$ for each cue. So, the correlations depend only on the N_i. However, by simple application of combinatorial formulae it can be shown that for N=120 paired comparisons of the 16 possible cue-patterns, the N_i are N_1=64, N_2=32, N_3=16, and N_4=8, and therefore we expect the noncompensatory pattern of correlations $E(r_{Tc1}) = 2E(r_{Tc2}) = 4E(r_{Tc3}) = 8E(r_{Tc4})$.

A coding method to fulfill the conditions

The relations derived above only hold if the before-mentioned conditions (1) through (4) are met. So the task is to find an assignment method of the cue patterns to "j" and "k" (see Equations 2 and 3) which satisfies the conditions across the 120 paired comparisons. A pragmatic method is found as follows: If one assigns numbers j,k to the cue patterns in the order of their lexicographic hierarchy ($j,k \in \{1,2,...,16\}$), every paired comparison can be numbered $m \in \{1,2,....,120\}$ running through the indices j and $k>j$. Conditions 1 to 4 are met when patterns are assigned to “j” or “k” in an alternating fashion running through the comparisons m. The vector **T** predicted by TTB (without error) then contains the alternating elements “1” and “-1”. It is easy to see that at least condition 4 (mean of **T** is zero) must be satisfied by this procedure because **T** will contain as many “1” as “-1”. Furthermore this method yields pairwise zero correlations between cues (condition 4) and equal cue variances (*SD*=0.73). However, applying this method, cue 4 shows a slight deviation from the pattern ($\bar{c}_4 = 0.067$ and $SD = 0.732$), but this deviation is numerically irrelevant as it affects only the fourth decimal place of the correlation r_{Tc4}.

Berichte aus dem Psychologischen Institut der Universität Bonn

Die „Berichte aus dem Psychologischen Institut der Universität Bonn“ (ISSN 0931-024X) gibt es seit 1975. Ab Jahrgang 5 (1979) beginnt die Heftzählung in jedem Jahr bei Heft 1. Von 1984 bis 1994 wurden die „Bon-ner Berichte“ von Dr. Edgar Erdfelder und Dr. Joachim Funke herausgegeben. Seit 1995 fungieren Dr. Edgar Erdfelder und Dr. Michael Kavsˇek als Herausgeber. Eine Übersicht über die zuletzt publizierten Hefte gibt nachfolgende Liste.

Band 19 (1993)

Heft 1: Beckmann, J.F., Funke, J. & Guthke, J. (1993). Zum Zusammenhang von kom-plexem Problemlösen, Intelligenztest- und Lerntest-Indikatoren.

Heft 2: Meiser, T. (1993). Loglinear Rasch models: Theory and applications in longitudinal rese-arch.

Heft 3: Faul, F. & Erdfelder, E. (1993). Vergleich verschiedener experimenteller Verfahren zum Test additiver Informations-Integrations-Modelle.

Heft 4: Hinnersmann, H. (1993). Evaluation eines Trainingansatzes zum Schlußfolgern und Urteilen.

Heft 5: Faul, F. (1993). NCM: Numerisches Conjoint Measurement. Eine Programm-doku-menta-tion.

Heft 6: Meiser, T., Fichtel, M., Rompe, P. & Rudinger, G. (1993). Mixture distribution Rasch models for the analysis of longitudinal discrete data.

Band 20 (1994)

Heft 1: Erdfelder, E., Faul, F. & Buchner, A. (1994). GPOWER: A general power analysis program.

Band 21 (1995)

Heft 1: Ruoß, M. (1995). Schmerz und Kognition. Kognitive Faktoren bei der Entstehung, Aufrechterhaltung und psychologischen Behandlung chronischer Schmerzen.

Heft 2: Scholz, O.B. & Adler, F. (1995). Atopische Dermatitis und ihre Beziehungen zu der-matologischen, allergologischen, ökologischen und psychosozialen Einflußfaktoren. Selbstberichte davon Betroffener.

Heft 3: Erdfelder, E. & Buchner, A. (1996). Process dissociation measurement models: Good versus better.

Band 22 (1996)

Heft 1: Funke, J., Dilger, S. & Grube-Unglaub, S. (1996). Manual zum "Skript-Monitoring-Test" (SMT).

Heft 2: Krüger, T. (1996). Eine Normierung der Ergänzung deutscher Wortanfänge zu Substantiven mit fünf oder sechs Buchstaben.

Heft 3: Kavsˇ ek, M. (1996). Visuelle Wahrnehmung in der frühesten Kindheit: Reflexionen zum Stand der Forschung.

Band 23 (1997)

Heft 1: Musch, J. & Lieberei, W. (1997). Eine auswertungsobjektive Postkorbübung für Assessment Center.

Band 24 (1998) (Kein Heft erschienen)

Band 25 (1999)

Heft 1: Bröder, A. (1999). PMM theory, overconfidence, and representative sampling of items: A review of data.

Eine Übersicht über alle bislang publizierten Hefte bietet die WWW-Homepage der "Bonner Be-richte" (http://www.psychologie.uni-bonn.de/allgm/publikat/bericht.htm). Über diese WWW-Homepage können darüber hinaus teilweise die Zusammenfassungen, teilweise auch die gesamten Berichte eingesehen und ggf. kopiert werden.