Measures for Quantitative Process-Tracing Methods

Rene Riedl
Johannes Kepler University Linz

Eduard Brandstatter
Johannes Kepler University Linz

Follow this and additional works at: http://aisel.aisnet.org/icis2006

Recommended Citation

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2006 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
MEASURES FOR QUANTITATIVE PROCESS-TRACING METHODS

Quantitative Research Methods

Riedl, René
Dept. of Business Informatics
Johannes Kepler University Linz
rene.riedl@jku.at

Brandstätter, Eduard
Dept. of Social and Economic Psychology
Johannes Kepler University Linz
eduard.brandstaetter@jku.at

Abstract

To know the decision strategy and the underlying information acquisition behavior of a person is essential for the design of computer programs that support managerial and consumer decision making. Both quantitative methods (e.g. computer-assisted process-tracing systems) and qualitative methods (e.g. verbal protocols) are used to investigate information acquisition behavior of managers and consumers, respectively, and to detect their decision strategies. This article presents measures that allow for the detection of important decision strategies—that is, one is able to distinguish decision strategies from each other. The presented measures are the methodological basis of widely used quantitative process-tracing methods such as computerized process-tracing (CPT) tools or eye-tracking systems. The article concludes that the measures and their application within CPT tools or eye-tracking systems constitute the basic foundation of an emerging research field called clickstream analysis, which is a method to investigate human decision processes by analyzing computer log files.

Keywords: Decision process, decision strategy, information acquisition behavior, computer-assisted process tracing

Introduction

Humans daily make several decisions. Some of them are habitual in nature (e.g. should one take the car or bus to get to work?), whereas others require a greater amount of deliberation (e.g. buying an apartment). From a research perspective, the issue here is not which means of transport or apartment is chosen per se. Rather, it is the decision-making process that is of interest, and in particular, three facets of that process: What information does a decision maker use? How is the information used? When is the information used?

As emphasized by Todd and Benbasat (1987), to know human decision processes is essential for the design of decision support systems. Kim and Talalayevsky (2005), for example, examined the impact of intermediaries on consumer strategies and concluded: “Intermediaries provide in-depth decision support systems for consumers to search and sort out vast amount of information available on Internet” (p. 1626). Hence, considering the importance of in-depth analysis of decision processes, i.e., “opening the black box”, it is essential to advance the development of techniques (i.e., process-tracing methods) that enable people’s information acquisition behavior to be revealed.

Black box models dominated research in psychology from the 1910s to the late 1960s. This was the heyday of behaviorism—a research paradigm that neglected internal processes but concentrated on simple stimulus–response patterns (Skinner 1974; Watson 1913). But the limitations of behaviorism, that is, this neglect of internal processes, ultimately led to a new paradigm that investigated them. From the 1960s cognitive psychology relied on rigor experiments and other quantitative research methods to uncover the processes that take place between the onset of a stimulus and a response or choice (Brandstätter et al. 2006; Newell 1966; Nisbett and Wilson 1977; Payne 1976a; Payne 1976b; Russo 1978; Russo and Rosen 1975; Simon 1979).
Research in the field of decision making has long been focused on measuring the effects of the variation of independent variables (e.g. availability of decision aids) on the dependent variable (e.g. decision quality). Yet as Todd and Benbasat (1987) postulated almost 20 years ago, “there is a need for methods that open the black box” (p. 494). Process-tracing methods may serve this purpose.

There are several process-tracing methods, including verbal protocols (Ericsson and Simon 1980; Newell and Simon 1972), information display boards (Finch et al. 1987; Payne 1976a; Stokmans 1992; Verplanken et al. 1992), tracing of eye movements (Lohse and Johnson 1996; Russo and Dosher 1983), phased narrowing (Jasper and Levin 2001; Levin and Jasper 1995), and computerized process tracing (CPT) such as the MOUSELAB system (Payne et al. 1993). Each of these methods, usually used in the laboratory, has its strengths and weaknesses (for an overview see Todd and Benbasat 1987). In a typical process-tracing study participants are presented with the decision problem of choosing an option from a set of options, which are characterized by utility scores or attribute values that are presented in the cells of a matrix. At the beginning of the experiment all cells are covered. To arrive at the final decision, a participant has to uncover cells of the matrix. After the choice has been made, the researcher can analyze the information acquisition behavior, allowing inferences on the cognitive processes of a participant.

Process-tracing methods are not only of interest to psychologists (who developed them), nor are they exclusively used by them. Rather, various scientific fields such as marketing (Brucks et al. 2000), corporate finance (Swain and Haka 1996), and political science (Chin and Taylor-Robinson 2005) used these methods. In contrast to other management disciplines, the IS community has been contributing significantly to the advancement of process-tracing methods. Weissinger-Baylon et al. (1980), for example, demonstrated that a process-tracing method like visual mental imagery protocols provide step-by-step traces of executive decision making. Cooper (1983) proposed that the lack of a general theory of managerial information requirements is an underlying cause of the many MIS failures.

Viewing decision making as a production process (decisions being produced in much the same manner as normal goods or services), he employed microeconomic concepts to provide guidance for the development of MIS, thereby taking a first step toward developing a theory of managerial information requirements. In 1987, Todd and Benbasat presented a variety of process-tracing methods with their relative strength and weaknesses. In 1993, at the International Conference on Information Systems (ICIS), a panel discussion was conducted on the topic of modeling expert decision making (Chung et al. 1993). The session aimed at exploring interdisciplinary but controversial perspectives of modeling decision-making behavior. The session, importantly, focused on the “process view” of decision making—that is, the question of how a decision is reached. Furthermore, a recent study used process-tracing data to investigate the contextualized access to task domain knowledge enabled by hypertext-style links (Mao and Benbasat 1998).

In this article, we present process-tracing measures that can detect decision strategies in laboratory settings. Understanding when people use different decision strategies will help in the design of better decision tools. First, we describe 14 key decision strategies, which are explained on the basis of five characteristics. Based on these five characteristics, we introduce our measures, along with several examples. Afterward, we point out possible limitations of our approach and we present practical applications of our work. Finally, we discuss the relevance of our work for quantitative research methods in the IS discipline.

**Decision Strategies**

In accordance with Payne et al. (1992), we define a decision strategy “as a set of operations used to transform an initial stage of knowledge into a final goal state of knowledge where the decision maker feels the decision problem is solved” (p. 108). We use the term *decision rule* synonymously. We consider both algorithmic rules (e.g. the multi-attribute utility rule) and heuristics (rules solving a decision problem by intelligent guesswork rather than by following some pre-established formula) to be decision strategies.

First, decision strategies can be distinguished by whether they allow for compensating for a bad value on one attribute with a good value on another attribute. Such so-called compensatory strategies therefore require explicit trade-offs among attributes. Second, strategies can be distinguished by the degree to which the amount of processing is consistent or selective across options or attributes. That is, is the same amount of information examined for each option or attribute, or does the amount vary? Third, some decision strategies do not process all information, whereas others do. Hence, strategies can be distinguished by the amount of information processed. People’s *bounded rationality* is an important reason why they do not process all the information available (Gigerenzer and Selten 2001; Simon 1982). Fourth, information processing is either option based or attribute based. In option-based processing,
multiple attributes of a single option are considered before information about the next option is processed. In attribute-based processing, the values of several options on a single attribute are processed before information about a further attribute is processed. Fifth, decision strategies differ with regard to the degree of quantitative and qualitative reasoning used. In general, strategies that involve summing, subtracting, and/or multiplying values as well as counting are considered to be quantitative. In contrast, strategies that simply compare values are regarded as qualitative; note that this terminology is rooted in psychology (e.g. Payne et al. 1993). In computer science comparisons such as “greater than” etc. are considered to be quantitative.

In the following we describe 14 decision strategies: eight strategies that are in a typical adult’s “cognitive toolbox,” (Payne et al. 1993, p. 32), three “major strategies identified thus far by behavioral researchers” (Hastie and Dawes 2001, pp. 232-234), and three strategies defined in Jungermann et al. (2005).

1. **MAU (multi-attribute utility model)** chooses the option with the highest aggregated value that is defined as the sum of the weighted attribute values. In the literature (see, for example, Payne et al. 1993), MAU is often characterized as WADD (weighted additive rule). MAU (i.e., WADD) strategies are usually viewed as normative rules (Keeney and Raiffa 1976) because they systematically process all information.

2. **EQW (equal weights rule)** only differs from MAU in the way that it simplifies decision making by ignoring information about the relative importance of each attribute.

3. **ADD (additive difference rule)** compares two options at a time, attribute-by-attribute. Then, the differences across the attributes are summed to provide a single overall difference score across all attributes for that pair of options. The winner is then compared to the next option, and so on. The best option has lost no comparison.

4. **SAT (satisficing heuristic)** considers options sequentially, in the order in which they occur in the choice set. The value of each attribute for a particular option is considered to see if it meets a predetermined cutoff (aspiration) level for that attribute. If any attribute fails to meet the level, the option is rejected and the next option is considered. SAT is the oldest decision strategy described in the literature (Simon 1955).

5. **EBA (elimination-by-aspects strategy)** eliminates options that do not meet a minimum cutoff value for the most important attribute. This elimination process is repeated for the second most important attribute. Processing continues until a single option remains (Tversky 1972).

6. **LEX (lexicographic rule)** simply means selecting the option with the best value on the most important attribute. If there are several options with the same value, the next most important attribute is taken and the option with the best value is chosen. The process stops when only one option is left.

7. **FRQ (frequency of good and/or bad features rule)** starts with the development of cutoffs for specifying good and bad features. Then a decision maker counts the number of good and/or bad features of an option. A decision maker chooses the option with the most good features, the fewest bad features, or both (Alba and Marmorstein 1987).

8. **MCD (majority of confirming dimensions rule)** involves processing pairs of options (like ADD). The values for each of the two options are compared on each attribute. The option with the majority of winning attribute values is retained and is then compared with the next option. The process of pair-wise comparison stops if all options have been evaluated and the final winning option has been identified (Russo and Dosher 1983).

9. **DOM (dominance rule)** chooses the option that is at least as good as every other alternative on all important attributes. Some authors, for example, Hastie and Dawes (2001), state a broader definition; that is, they also define a strategy that aims at finding an option that is worse than any other option on all attributes and then throwing it out of the choice set as DOM.

10. **DIS (disjunctive rule)** first sets cutoff points on the important attributes and then looks for the first option that is at least as good as the cutoff value on any attribute.

11. **REC (recognition heuristic)** chooses the option with the highest value on the attribute “name recognition” REC is mainly used in choices where people are very poorly informed about the options considered (Goldstein and Gigerenzer 2002).

12. **MAJ (majority rule)** chooses the option with the highest number of dominant attribute values.
13. **LIM (least important minimum heuristic)** first determines the worst value of each option and then chooses the option with the least important worst value (“take the lesser evil”).

14. **LVA (least variance heuristic)** chooses the option with the lowest variance across the attribute values. LVA only makes sense for decision situations in which no dominant option exists.

In Table 1 the decision strategies are compared on the basis of the five mentioned characteristics. With the comparison of the 14 strategies we extend the compilation of decision strategies listed above.

### Table 1. Characteristics of Decision Strategies

<table>
<thead>
<tr>
<th>Compensatory (C) vs. Noncompensatory (N)</th>
<th>MAU</th>
<th>EQW</th>
<th>ADD</th>
<th>SAT</th>
<th>EBA</th>
<th>LEX</th>
<th>FRQ</th>
<th>MCD</th>
<th>DOM</th>
<th>DIS</th>
<th>REC</th>
<th>MAJ</th>
<th>LIM</th>
<th>LVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent (C) vs. Selective (S)</td>
<td>C</td>
<td>C</td>
<td>S*</td>
<td>S</td>
<td>S</td>
<td>C</td>
<td>S*</td>
<td>C</td>
<td>S</td>
<td>S</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Information ignored? Yes (Y) vs. No (N)</td>
<td>N</td>
<td>N*</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N*</td>
<td>N*</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Option-based (O) vs. Attribute-based (A)</td>
<td>O</td>
<td>O</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>O</td>
<td>A*</td>
<td>O</td>
<td>A*</td>
<td>A*</td>
<td>A</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Quantitative (QN) vs. Qualitative (QL) reasoning</td>
<td>QN</td>
<td>QN</td>
<td>QN</td>
<td>QL</td>
<td>QL</td>
<td>QN</td>
<td>QN</td>
<td>QN</td>
<td>QL</td>
<td>QL</td>
<td>QN</td>
<td>QN</td>
<td>QN</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** White background color indicates the source Payne et al. (1993, p. 32.). Dark grey background color indicates the source Hastie and Dawes (2001, pp. 232-234). Light grey background color indicates our contribution. * Our classification differs from the source.

All decision strategies in Table 1, except MAU, are boundedly rational (Payne et al. 1993). Bounded rationality describes how real people make decisions when time is limited, information costly, and the decision problem unclear. Many psychologists believe that people commit reasoning fallacies due to limited cognitive capacities (Gigerenzer and Selten 2001; Simon 1982). Many economists, in contrast, believe that people behave as if they optimize under constraints, such as information costs. These two views diverge sharply: one promotes irrationality, the other rationality. The father of bounded rationality, Simon, in contrast, believed that boundedly rational strategies are anchored in both the mind and the environment. In his view choice strategies exploit the capacities of the human mind (such as recognition memory), which allows for quick decisions. Choice strategies also exploit statistical structures (such as the signal–noise ratio), which allows people to ignore information. Unlike the “rational” MAU strategy, which concentrates on utility maximization, boundedly rational strategies often consider four goals (Payne et al. 1993): (1) maximizing utility, (2) saving cognitive effort, (3) minimizing negative emotions, and (4) justifying one’s decision to other people. EBA, for example, maximizes utility in certain environments (Payne et al. 1993), demands little cognitive effort, avoids negative emotions by eliminating the worst options (because they fall short of the aspiration levels), and its simplicity can be justified to other people. Using the four goals, one could build a meta-theory that classifies each of the 14 decision strategies in Table 1.

### Measures for Quantitative Process Tracing

Participants taking part in a laboratory experiment are presented with the decision problem of choosing one of several options. The options can be characterized by either utility scores or attribute values, which are presented in the cells of a matrix. We call such a matrix a choice matrix. At the beginning of the experiment all cells are covered. To arrive at the final decision, a participant has to uncover particular cells of the matrix. After the choice has been made, the researcher can analyze the information-acquisition behavior. Hence, it is possible to deduce the decision strategy pursued by the participant and to infer the cognitive processes of the participant. In the following, we present our measures for quantitative process tracing.
Measure 1: Ratio of Option-wise Transitions to Attribute-wise and Mixed Transitions

The first measure concerns the ratio of option-wise transitions to attribute-wise and mixed transitions. A transition is defined as option-wise if a participant uncovers two cells within an option, and attribute-wise if a participant uncovers two cells within an attribute. Mixed transitions are both option-wise and attribute-wise (Figure 1).

Consider a choice matrix containing \( o \) options and \( a \) attributes. MAU, EQW, FRQ, LIM, and LVA require the participant to look at all cells within one option. Within one option there are \( a \) attributes. Thus a participant using one of these five decision strategies makes \((a-1)\) option-wise transitions within one option. Multiplied by the number of options \( o \), any of the five strategies predicts that the number of option-wise transitions \((OT)\) for a choice matrix is

\[
OT = (a-1) \times o, \quad (1)
\]

where \( a \) and \( o \) are the number of attributes and options of the choice matrix, respectively.

After a participant has uncovered all cells within an option, any of these five strategies predicts that the participant will move to the next option. This transition from one option to another option can be either an attribute-wise or a mixed transition. Because there are \( o \) options, any of the five strategies predicts that the number of attribute-wise and mixed transitions \((AT+MT)\) for a choice matrix is

\[
AT + MT = o-1, \quad (2)
\]

where \( 0 \leq AT \leq (o-1) \) and \( 0 \leq MT \leq (o-1) \); note that at least one of the two factors has to be positive, otherwise no decision problem would exist.

In the following we give examples of how to calculate \( OT \) as well as \( AT + MT \). We use both symmetrical and asymmetrical choice matrices, because the validity of a measure can be influenced by the proportion of the number of options and attributes in a matrix. For example, Boeckenholt and Hynan (1994) found that Payne’s (1976b) Search Index (SI)—an index that shows whether information selection behavior is option-wise or attribute-wise—only leads to valid results if the number of options and attributes in a matrix is equal. Boeckenholt and Hynan (1994) therefore proposed an alternative index, the so-called Strategy Measure (SM). But Payne and Bettman (1994) showed that if symmetrical matrices are used to describe a choice problem, then the SM also has some weaknesses. A pragmatic solution to the problem therefore is to use the SI with symmetrical matrices and the SM with asymmetrical matrices. (Note that the SI and the SM have the same direction: a positive value stands for option-wise and a negative value for attribute-wise processing. The SM is not restricted to the range of \(-1\) to \(+1\) as is the SI, because the range of the SM depends on the matrix size and the number of transitions.)

Suppose, for example, a choice matrix has four options and three attributes (which is an asymmetrical matrix). In this matrix \( OT = 8 \) and \( AT + MT = 3 \) (see the top row in Figure 2). With five options and five attributes (which constitutes a symmetrical matrix) \( OT = 20 \) and \( AT + MT = 4 \) (see the bottom row in Figure 2). Furthermore, the left column in Figure 2 shows examples where \( MT = 0 \); the middle column shows examples where \( AT = 0 \); and the right column demonstrates a blend of \( OT, AT, \) and \( MT \).

For any choice matrix of the size \( a \times o \) we can calculate the ratio \( OT / (AT+MT) \):

\[
OT / (AT+MT) = \frac{(a-1) \times o}{(o-1)}. \quad (3)
\]
The more a participant’s ratio resembles the ratio in Equation 3, the more likely this participant has used MAU, EQW, FRQ, LIM, or LVA. Equation 3 is context sensitive because it allows predicting different ratios for different choice matrix dimensions.

**Figure 2. Examples of Measure 1 Using MAU, EQW, FRQ, LIM, or LVA**

**Measure 2: Ratio of Time Spent on Options**

A second measure to identify whether a participant pursues a particular decision strategy is the ratio of time spent on options. MAU, EQW, ADD, FRQ, MCD, DOM, MAJ, LIM, and LVA require participants to look at all options and all attributes, which means that a participant has to open all cells in a choice matrix. If the cells contain utility scores (rather than attribute values), all the cells need the same processing time. If so, MAU, EQW, FRQ, DOM, REC, MAJ, LIM, and LVA predict that participants look at all options equally long and the ratio of time spent on the options thus equals 1:1: …:1, where the number of 1s represents the number of options of the choice problem.

How much time will participants using ADD (or MCD, which uses a similar procedure) spend on the options? Consider a choice problem with two options. A participant using ADD or MCD will compare all attribute pairs and thus look at each option equally long. This participant will look at two options with a ratio of 1:1 (Table 2). Next consider a choice problem with three rather than two options (Table 2). A participant using the ADD or MCD strategy will compare the attribute utilities of options 1 and 2. Assuming that option 1 outperforms option 2, the participant will then compare options 1 and 3. In total, the participant has looked at four options (1 and 2, 1 and 3) with a time ratio of 2:1:1. That is, this participant looked twice as long at option 1 as at options 2 or 3. A ratio of 2:1:1 or 1:2:1 or 1:1:2 thus indicates participants are using ADD or MCD, whereas ratios of, for instance, 2:2:2 or 3:1:1 do not. Table 2 lists all possible ratios for choice matrices with different numbers of options. Note that these ratios are independent of the number of attributes and we assume that a participant’s short-term memory has a limited capacity to keep the attribute utilities in mind. Considering Miller’s (1956) work—that human short-term memory capacity has a limit of 7±2 information digits—our assumption is realistic. Equation 4 shows how to calculate the sum of options looked at altogether (SO) when using ADD or MCD:

\[ SO = 2o – 2, \]  

where \( o \) represents the number of options in the choice matrix.

In the case of EBA and LEX it is possible to predict that the ratio of time spent on options is *not* 1:1: …:1, because the participant eliminates options during the decision process. Furthermore, in the case of SAT and DIS no prediction concerning the ratio of time spent on options is possible, because we do not know a participant’s cutoff levels and therefore we cannot predict when the search process stops.
Table 2. Example for Measure 2 Using ADD or MCD

<table>
<thead>
<tr>
<th>Number of Options</th>
<th>Sum of Options Looked at (with Repetitions)</th>
<th>Possible Ratios of Time Spent on Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>1:1</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2:1:1</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>2:2:1:1 or 3:1:1:1</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>4:1:1:1:1 or 3:2:1:1:1 or 2:2:2:1:1</td>
</tr>
<tr>
<td>o</td>
<td>2o-2</td>
<td></td>
</tr>
</tbody>
</table>

ADD … Additive Difference Rule  | MCD … Majority of Confirming Dimensions Rule  | o … Number of Options in the Matrix.

Measure 3: Correlation between Attribute Rank and Number of Cells Uncovered for Each Attribute

Attribute rank calculates the mean rank of all cells of a specific attribute, where the first cell uncovered gets the value 1, the second cell uncovered the value 2, and so on (Table 3). Generally, the lower the rank of an attribute the earlier that attribute has been looked at, whereas the attribute with the highest rank has been looked at most recently. Using the correlation between attribute rank and the number of cells uncovered for each attribute, EBA and LEX allow a specific prediction. Both heuristics start by looking at the most important attribute. If two or more options have the same attribute value on the most important attribute (LEX), or if two or more options surpass the cutoff value on the most important attribute (EBA), people just keep the winning options and scan the winning options on the second most important attribute. Again, if one attribute value is better than all others on the second most important attribute (LEX), or if one attribute value surpasses the cutoff value (EBA), a decision can be reached. If not, people scan the remaining winning options on the third most important attribute. This process is repeated until a decision can be reached. Both LEX and EBA assume that people look at the most important attribute first. The first attribute thus has the lowest attribute rank. The second most important attribute has the second lowest attribute rank and so on. Because fewer options are scanned on each attribute, both EBA and LEX predict a negative correlation ($r$) between attribute rank and the number of cells uncovered for each attribute.

Table 3 shows an example of pursuing the EBA strategy and the calculation of measure 3 can be easily comprehended. Consider a choice matrix with four options and four attributes. Consider a participant who started by looking at all four options (O1, O2, O3, and O4) on the first attribute (A1). Then, the participant looked at O1, O2, and O3 on the second attribute (A2); that is, O4 fell below the cutoff level on the first attribute. Afterward, the participant inspected O1 and O2 on the third attribute (A3); hence it is clear that O3 did not surpass the cutoff level on the second attribute. Then, the participant stopped the search process; let us assume that O1 was chosen. As described above, the correlation between attribute rank and the number of cells uncovered for each attribute has to be negative for a search pattern like the one in Table 3 (Pearson’s correlation coefficient is −0.995 in the example). For MAU, EQW, ADD, FRQ, MCD, DOM, DIS, MAJ, LIM, and LVA the correlation coefficient is expected to be zero, because the number of cells uncovered for each attribute is a constant factor in these decision strategies.

(Note that for SAT the correlation coefficient can be either negative or zero and in the case of REC no correlation coefficient can be calculated, because per definition REC only considers one attribute, namely, “name recognition”).
Quantitative Research Methods

Table 3. Example of Measure 3 Using EBA

<table>
<thead>
<tr>
<th>O1</th>
<th>O2</th>
<th>O3</th>
<th>O4</th>
<th>Sum</th>
<th>#CUA</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>A2</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>*</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>A3</td>
<td>8</td>
<td>9</td>
<td>*</td>
<td>*</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>A4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>—</td>
<td>0</td>
</tr>
</tbody>
</table>

A … Attribute  | O … Option  | * … Cell was not Uncovered  |
| ---           | ---         | ---                        |
— … No Calculation Possible  | #CUA … Number of Cells Uncovered for Each Attribute  | AR … Attribute Rank.  |

Measure 4: Final Decision

A further measure to identify a decision strategy refers to the final decision participants have made after they have searched in the cells that contain utility scores (and not objective attribute values). By stating a rank order of options, participants show their preferences among all options presented in the choice matrix. In addition to choosing an option, participants are asked to state the relative importance of every single attribute. Hence, we do have attribute weights. Knowing the attribute weights is important to be able to predict the final decision of some decision strategies (e.g. MAU).

For MAU, EQW, ADD, MCD, DIS, REC, MAJ, LIM, and LVA we can predict a person’s rank order of options (note that this is only possible when no dominant option is in the choice set). This allows us to compare a participant’s predicted rank order of options with his or her actual rank and to calculate the difference: The lower the difference, the more likely a particular decision strategy was used. Moreover, SAT can be linked to a participant’s final decision. SAT selects the first option that surpasses all cutoff values. If a person requires, for instance, that a satisfactory option must surpass five cutoff values, the person will dismiss an option as soon as any of the five attribute values falls short of the cutoff value of the particular attribute. Consequently, the chosen option must surpass all cutoff values. In sum, SAT predicts that people choose the option with the highest number of cells uncovered.

Furthermore, LEX can also be linked to the final decision participants have made. LEX predicts that people look at the most important attribute and select the option with the highest attribute value on this most important attribute. We define the most important attribute as the attribute participants have looked at first, that is, the attribute with the lowest attribute rank (see above). If the attribute with the lowest attribute rank contains a utility score that surpasses all other utility scores on that most important attribute, LEX predicts choosing the option with the highest utility score on the most important attribute. If there is not one but rather two or more utility scores with the highest value, LEX predicts that people select the option with the highest utility score on the attribute with the second lowest attribute rank, and so on. For EBA and FRQ a prediction is not possible, because we do not know a participant’s cutoff values.

Note that it is basically possible to identify DOM by including one dominant option in the choice matrix and then by comparing the predicted and actual final decision. In our experimental setting we do not include a dominant option because otherwise we would not be able to distinguish DOM from other strategies such as MAU, EWQ, LIM, or LVA.
Table 4. Measures for Quantitative Process Tracing

<table>
<thead>
<tr>
<th>Measure 1</th>
<th>Prediction of G7 ((A1+M2) possible?</th>
<th>YES</th>
<th>YES</th>
<th>NO</th>
<th>NO</th>
<th>NO</th>
<th>NO</th>
<th>YES</th>
<th>NO</th>
<th>NO</th>
<th>NO</th>
<th>YES</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure 2</td>
<td>Ratio of time spent on options...</td>
<td>= 1:1:...:1</td>
<td>= 1:1:...:1</td>
<td>See Table 2</td>
<td>No prediction possible</td>
<td>= 1:1:...:1</td>
<td>= 1:1:...:1</td>
<td>See Table 2</td>
<td>= 1:1:...:1</td>
<td>No prediction possible</td>
<td>= 1:1:...:1</td>
<td>= 1:1:...:1</td>
<td>= 1:1:...:1</td>
</tr>
<tr>
<td>Measure 3</td>
<td>Correlation (r) between A6 and #CUA</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r &gt; 0</td>
<td>r &gt; 0</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r = 0</td>
<td>r = 0</td>
</tr>
<tr>
<td>Measure 4</td>
<td>Prediction of final decision possible?</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

| Substitute 1 (for measure 1) | i(i,k,i) = i(k,n) | YES | YES | NO | NO | NO | NO | YES | NO | NO | NO | YES | YES |
| Substitute 2 (for measure 1) | Ratio of time spent on attributes... | = 1:1:...:1 | = 1:1:...:1 | = 1:1:...:1 | No prediction possible | = 1:1:...:1 | = 1:1:...:1 | = 1:1:...:1 | = 1:1:...:1 | No prediction possible | = 1:1:...:1 | = 1:1:...:1 |
| Substitute 3 (for measure 1) | Must all cells in the matrix be uncovered? | YES | YES | YES | NO | NO | NO | YES | YES | YES | YES | YES | YES |

MAU ... Multi Attribute Utility Model  | EQW ... Equal Weights Rule  | ADD ... Additive Difference Rule  | SAT ... Satisficing Heuristic  | EBA ... Elimination-by-Aspects Strategy  | LEX ... Lexicographic Rule  | FRQ ... Frequency of Good and/or Bad Features Rule  | MCD ... Majority of Confirming Dimensions Rule  | DOM ... Dominance Rule  | DIS ... Disjunctive Rule  | REC ... Recognition Heuristic  | MAJ ... Majority Rule  | LIM ... Least Important Minimum Heuristic  | LVA ... Least Variance Heuristic  | OT ... Optionwise Transition  | AT ... Attributewise Transition  | MT ... Mixed Transition  | r ... Pearson’s Correlation Coefficient  | AR ... Attribute Rank  | #CUA ... Number of Cells Uncovered for Each Attribute  | OR ... Option Rank  | n ... Multiple  | a ... Number of Attributes in the Matrix .
Substitutes for Measure 1

Instead of measure 1, three other measures can be used (because they are “100% substitutes”).

Substitute 1, difference between any two option ranks (ORs), calculates the mean rank of all cells of a specific option, where the first cell uncovered gets the value 1, the second cell uncovered the value 2, and so on. Generally, the lower the rank of an option, the earlier that option has been looked at, whereas the option with the highest rank has been looked at most recently. MAU, EQW, FRQ, LIM, and LVA predict that participants look at one option after the other because they process all attribute information available on every single option and then select the best option (what “best” means is defined by each strategy individually). Consider, for example, a decision problem with three attributes. The first option looked at has an option rank of 2 [i.e., (1 + 2 + 3)/3], the second option looked at a rank of 5 [i.e., (4 + 5 + 6)/3], the third a rank of 8 [i.e., (7 + 8 + 9)/3], and so on. Note that the differences between the option ranks, that is, 2, 5, and 8, are always three, which is the number of attributes in the choice problem. MAU, EQW, FRQ, LIM, and LVA thus predict that the difference between any two option ranks equals a multiple of the number of attributes of the choice problem. For example, in a choice problem with five attributes and three options the attribute ranks of 3, 8, and 13 indicate MAU, EQW, FRQ, LIM, or LVA. Equation 5 thus predicts

\[ |OR_i - OR_j| = n*a, \]  

where \( OR_i \) and \( OR_j \) are the option ranks of any two options, \( n \) is a multiple, and \( a \) represents the number of attributes in the choice matrix.

Substitute 2, ratio of time spent on attributes, follows the description of the ratio of time spent on options (measure 2). MAU, EQW, ADD, FRQ, MCD, DOM, MAJ, LIM, and LVA predict that participants look at all attributes equally long and the ratio of time spent on the attributes thus equals 1:1: … :1, where the number of 1s represents the number of attributes of the choice problem. In the case of EBA and LEX, participants do not look equally long at each attribute and therefore the ratio of time spent on the attributes does not equal 1:1: … :1. In the case of REC, only one attribute (“name recognition”) is of relevance and therefore the ratio of time spent on the attributes also does not equal 1:1: … :1. Finally, in the case of SAT and DIS no prediction concerning the ratio of time spent on attributes is possible, because we do not know a participant’s cutoff points and therefore we cannot predict when the search process stops.

Substitute 3, the number of cells that have to be opened at least once, can be used to distinguish several decision strategies from each other. In the case of MAU, EQW, ADD, FRQ, MCD, DOM, MAJ, LIM, and LVA (group 1) all cells in the matrix have to be uncovered at least once. In the case of SAT, EBA, LEX, DIS, and REC (group 2) it is not necessary to uncover all cells.

The flow diagram in the Appendix (note that we do not use the conventional notation due to shortage of space) shows a possible way to use our described measures for quantitative process tracing to find out which of the 14 decision strategies a participant used. (Only in three cases—SAT/DIS, ADD/MCD, and EBA/LEX—did our measures not allow for a precise discrimination.) As shown by the flow diagram, the measures are applied sequentially to analyze the information acquisition behavior and to identify the decision strategy of a participant. The flow diagram can be used as an algorithm to implement the measures in a software tool that can be used to conduct process-tracing laboratory studies. However, this is just one possible way to use our measures. Hence, applying our new measures (i) in combination with known measures—Boeckenholt and Hynan’s (1994) SM; Payne’s (1976) SI, among others—and (ii) in other sequences can result in a different finding regarding the discrimination of the decision strategies.

Limitations

In this article, we have presented measures that can be used for quantitative process-tracing methods. In four cases our measures cannot distinguish between decision-strategy pairs: (i) SAT/DIS, (ii) DOM/MAJ, (iii) ADD/MCD, and (iv) EBA/LEX; all other decision strategies, however, can be identified precisely (see Appendix). Both SAT and DIS include aspiration levels. Intuitively, a person using SAT is likely to uncover more cells than a person using DIS. MAJ can be considered as a special case of DOM. If no dominant option exists, MAJ selects the option with the highest number of dominant attribute values. Concerning ADD and MCD, precise distinction is impossible because the two strategies are nearly the same, that is, MCD is a special case of ADD (see the definitions above).
Our measures are limited by the assumption that people closely follow a strategy’s algorithm. In the case of MAU, for instance, we assume that people search option-wise rather than attribute-wise, because MAU chooses the option with the highest sum of the weighted attribute values (see definition above). Theoretically, however, a person may search attribute-wise, remember all attribute values, calculate the sum of weighted attribute values for each option, and choose the option with the highest sum. Limited capacity of short-term memory (Miller 1956), however, renders attribute-wise search unlikely.

Another limitation is that we have not yet validated our measures with real data. Although we have shown theoretically the correctness of our measures, future research efforts should focus on their application in laboratory studies. As mentioned above, the flow diagram in the Appendix may serve as an algorithm to implement our measures in a CPT tool.

Real decision makers often may not feel constrained to use a pure form of any one strategy, preferring instead a combination of strategies. One frequently observed strategy combination is an initial use of EBA to reduce the option set to two or three options followed by the use of MAU to select from among those remaining (Bettman et al. 1998). Hence, if two or more strategies are used sequentially in the decision process, future research could focus on dividing decisions into segments in order to find out which combination of strategies is used (for a more detailed discussion of this issue see, for example, Cook 1993).

Another limitation refers to the application of our measures when they are used in CPT tools or eye-tracking systems, which are usually used in laboratory settings. Hence, artificiality may negatively influence the external validity of the findings; that is, decision strategies in the laboratory may differ from strategies in the field, although the decision situation is the same (e.g. buying a car).

Increasingly more software vendors have started offering systems that facilitate the design of software by the users themselves (see Koutrika and Ioannidis 2004; Petit-Rozé and Strudgeon 2006). This personalization trend might challenge the practical value of our research from a software vendor’s perspective. The arguments for this assertion include that the design effort shifts from the software developer to the user. We believe, however, that a better understanding of people’s decision strategies can help in designing software that can be personalized.

From the Laboratory to the Field

In recent years, a new field of research has been developing. The use of clickstream data to analyze decision-making processes is gaining considerable momentum. Clickstream data can be defined as data that users generate as they move from page to page and click on items within a Web site, usually stored in log files. Hence, clickstream analysis seems to be of particular allure for IS researchers. If designers of Internet shops, for instance, know how their potential customers make decisions, then they can tailor systems that actively support customers’ decision-making processes (for recent research see Alves and Belo 2004; Bucklin and Sismeiro 2003; Chatterjee et al. 2003; Clark et al. 2006; Di Scala and La Rocca 2002; Montgomery et al. 2004).

A good example of a shopping Web site that already contains many features that support the execution of several decision strategies is ACTIVESHOPPER.COM (an Internet platform of Shelron Group Inc., New York). If someone wants to buy, for example, a printer, the system lists many alternatives by different manufacturers, one below the other (see Figure 3). Users can make a direct comparison of two or more printers using checkboxes (“Compare Products”). If, for example, a customer selects two printers and compares them on several attributes and afterward compares one of these two printers with another printer, then this user obviously follows ADD or MCD. Such a clickstream can be traced easily by log file analysis. Now suppose a customer uses the attribute filter (see “Filter the Results” in Figure 3). This customer is likely using EBA, because the program eliminates all options that fall short of the aspiration level. Again, the use of the attribute filter can be traced by log file analysis.

Customers can also acquire information by clicking on the images of the printers. Suppose a customer clicks on the image of the printer at the top of the list. A new window opens and several retailers are shown that sell this product. Further, within the new window the customer can click on a button labeled “View product specifications” Such a customer pursues an option-wise search strategy.

Note that the clickstream analysis could, for example, record the time customers have spent on various options. Hence, the measures presented in this article (like the ratio of time spent on options) can form the methodological basis for clickstream research. To sum up, analyzing clickstream data allows for a detection of decision strategies...
pursued by Internet shop customers. System designers can use these results to improve existing Web sites and conceive of new approaches to online sales.

Figure 3. Example of a Shopping Web site (Screenshot of ACTIVESHOPPER.COM Taken on 14.08.2006)

Relevance of our Work for Quantitative Research Methods in Information Systems

In social sciences, research methods are either quantitative or qualitative (Kerlinger and Lee 2000). Quantitative research—which emanates from a post-positivist tradition—aims at collecting numerical data in order (i) to analyze it statistically and (ii) to facilitate other scientists replicating the research study. If one accepts this definition, our research is quantitative by nature.

The discourse on quantitative methods has gained considerable momentum in the IS discipline in the recent past (see, for example, Carte and Russell 2003; Chin and Dibbern 2006; Lee and Baskerville 2003), although as early as the 1980s a less intensive discussion on quantitative methods was already taking place (e.g. Baroudi and Orlikowski 1989; Jarvenpaa et al. 1985). Chen and Hirschheim (2004) examined 1,893 articles published in eight major IS publication outlets between 1991 and 2001. In their review, the authors considered surveys, laboratory experiments, and field experiments to be quantitative research methods. The findings of the literature analysis clearly show that research published in the top IS publication outlets was quantitative and positivist (MIS Quarterly: 46% survey, 15% lab experiment, and 4% field experiment; Information Systems Research: 45% lab experiment, 38% survey, and 3% field experiment; ICIS: 39% survey, 22% lab experiment, 0% field experiment). Considering the importance of quantitative methods for scientific progress in the IS discipline, the methodological improvement of surveys and experiments is essential. Advancements are possible in the domains of both data collection and data analysis.

Often, the discourse on quantitative methods is focused on issues that primarily concern survey research and/or data analysis—the current call for papers of MIS Quarterly on Partial Least Squares Modeling (PLS) can serve as a good example. No doubt such issues are important, but in our opinion, the IS community should also be strengthening its
activities in the field of experimental research and data collection methods. Too often a laboratory experiment is considered to be a method with high reliability and internal validity—see, for example, Table 1 in Jenkins (1985). Unfortunately, researchers take these two criteria for granted, that is, they do not sufficiently take into account that the (i) measurement instrument, (ii) research design, and (iii) experimental task (Jarvenpaa et al. 1985) highly influence reliability and internal validity. In this article, we address one of these three important issues: We introduce a measurement instrument that contains measures for quantitative process tracing that allow for the identification of decision strategies. We demonstrate the reliability of our measurement model with realistic examples.

Acknowledgments

The authors would like to acknowledge the excellent direction and insights provided by four anonymous reviewers and the support of the associate editors and track chairs.

References


Quantitative Research Methods


Watson, J.B. “Psychology as the Behaviorist Views It,” *Psychological Review* (20), 1913, pp. 158-177.

How to Use the Measures for Quantitative Process Tracing (Flow Diagram)

MAU ... Multi Attribute Utility Model | EQW ... Equal Weights Rule | ADD ... Additive Difference Rule | SAT ... Satisficing Heuristic | EBA ... Elimination-by-Aspects Strategy | LEX ... Lexicographic Rule | FRQ ... Frequency of Good and/or Bad Features Rule | MCD ... Majority of Confirming Dimensions Rule | DOM ... Dominance Rule | DIS ... Disjunctive Rule | REC ... Recognition Heuristic | MAJ ... Majority Rule | LIM ... Least Important Minimum Heuristic | LVA ... Least Variance Heuristic | a ... Number of Attributes in the Matrix | o ... Number of Options in the Matrix | n ... Multiple.