Heuristics in Multi-attribute Decision Making: Effects of Representation Format

Georg Jahn (georg.jahn@phil.tu-chemnitz.de)
Chemnitz University of Technology, Department of Psychology
Wilhelm-Raabe-Str. 43, D-09107 Chemnitz, Germany

Frank Renkewitz (frank.renkewitz@uni-erfurt.de)
University of Erfurt, Department of Psychology
Nordhäuser Str. 63, D-99089 Erfurt

Sonja Kunze (sonja.kunze@s2001.tu-chemnitz.de)
Chemnitz University of Technology, Department of Psychology
Wilhelm-Raabe-Str. 43, D-09107 Chemnitz, Germany

Abstract

When probabilistic inferences have to be drawn based on cue values stored in long term memory, participants appear to use a number of different heuristics. These heuristics vary in the amount of available cue information that is retrieved and integrated, and in the degree to which cue validities are considered. One factor that is assumed to affect strategy selection is the representational format of cue information: Holistic, image-based representations should induce the parallel retrieval of cue values and thus foster the use of compensatory strategies that take into account multiple cue dimensions. In contrast, verbal representations should favor non-compensatory strategies that draw inferences based on a single cue dimension. We tested this hypothesis in an experiment and explored how the learning procedure modulates the use of cue validities. The results suggest that the acquisition of cue knowledge from holistic images increases the use of a compensatory prototype matching strategy that neglects cue validities.

Keywords: Multi-attribute decision making, Probabilistic inference, Bounded rationality, Take-the-best heuristic .

Memory-based Multi-attribute Decisions

In most decisions the alternatives to choose from differ with regard to several attributes. Such multi-attribute decisions are linked to a fundamental issue in the study of human rationality: How do individuals integrate information?

A single attribute may be viewed as a value on a cue dimension. Then alternatives can be described by the values that they carry on cue dimensions. Usually, cue dimensions are not equally important. If cue values signal the likelihood of positive or negative outcomes (in probabilistic inferences), cue importance is referred to as cue validity.

In deciding between two alternatives, cue values and cue validity can be factored in to varying degrees (e.g., Gigerenzer, Todd, & the ABC Research Group, 1999). Individuals could compare cue values of alternatives on single dimensions, they could integrate the results of such comparisons from several dimensions, and they could weigh cue dimensions according to their validity.

Consider, for example, the task studied by Bröder and Schiffer (2003b, Experiment 4): Participants learned about ten suspects in a murder and then had to decide for several pairs of suspects who in a pair was more likely to be guilty based on the match with the description that witnesses gave of the murderer. The description of suspects and the description of the murderer contained four cues (accompanying dog, type of the jacket, type of trousers, and color of the shirt) that were the more valid the more witnesses were said to agree on the respective cue value.

The results of this experiment suggested that participants differed with regard to how completely they considered cue dimensions and cue validity. Furthermore, the format in which cue values were learned affected how they were used in decisions. Participants who were presented with suspects' images in the learning phase considered more cue dimensions than participants who learned lists of suspects' attributes. This effect of presentation format supported Bröder and Schiffer's prediction that integrated images should induce a holistic memory representation of alternatives. A holistic memory representation should then support retrieving and comparing several cue values in parallel in memory-based decisions.

More or less complete information integration corresponds to presumed strategies and heuristics in human information processing. Using all cue dimensions and cue validity, decisions could be based on all available cue values that are weighted by their validity. This strategy would choose the alternative with the highest weighted sum. Following Gigerenzer and his colleagues (1999), we will refer to this strategy as Franklin's rule (FR) after Benjamin Franklin who proposed a similar procedure.

A second strategy that also considers all cue dimensions but disregards cue validity has been named Dawes' rule (DR) after Robyn Dawes (Dawes, 1979). DR weighs each cue dimension equally and decides based on summed cue information. For example, if one suspect would match the description of the murderer in two attributes and the other suspect would match in three attributes, DR would select the latter.

A third strategy – prominent for its efficiency as a "fast and frugal" heuristic in some real-world domains as shown by Gigerenzer and his colleagues – is known as "Take the Best" (TTB). TTB uses cue validities, but decides on the basis of a single cue. Cue dimensions are considered in the order of validity starting with the most valid cue. As soon as a cue dimension is found on which the alternatives differ, the alternative with the superior cue value is selected and all dimensions of less validity are ignored. Consequently, TTB

is non-compensatory. Positive values on less valid dimensions cannot compensate for a negative value on a higher differentiating dimension.

The consistent application of a decision strategy (TTB, DR, or FR) produces a characteristic pattern of decisions (for a set of suitable decision items). Thus, the pattern of an individual's decisions can be classified by comparing it with the patterns that would be produced by TTB, DR, or FR. Such a strategy classification can be used to test the format hypothesis (Bröder & Schiffer, 2003b).

The format hypothesis predicts a prototype-matching strategy if internal representations are holistic and image-based (cf., Juslin & Persson, 2002). Prototype-matching would integrate cues in parallel and would appear as compensatory decision making (DR or FR). In contrast, with verbal encoding of cue values, the format hypothesis predicts sequential information retrieval, for which a sequential strategy such as TTB is better suited.

In Bröder and Schiffer's experiment (2003b, Experiment 4), participants' decision patterns were more often consistent with the sequential, non-compensatory strategy TTB when the cue values were presented verbally, and more often consistent with DR and FR when the cue values were presented as integrated images (i.e., pictures that each showed a dog and a human figure wearing the respective clothes). This result supported the format hypothesis.

The experiment that we report is similar to Bröder and Schiffer's experiment. Our objective was to test the format hypothesis with different materials and a change in the learning procedure that allowed us to disentangle to what extent sequential retrieval of information that favors TTB is induced by the presentation format or by the learning procedure.

Experiment

Our participants learned about travel offers (see Table 1) and then performed a sequence of binary choices (see Table 2 for examples). For each binary choice only the names of two offering companies were presented. The cue values of the respective offers had to be retrieved from memory. Participants were instructed to decide which of two travel offers a travel agency should include in its catalogue based on results of marketing research about optimal attributes (cue values) and attribute importance (cue validity).

Presentation format was varied as in Bröder and Schiffer (2003b). One group of participants learned lists of attributes (verbal condition), a second group learned the offers as integrated images (image condition, see Figure 1 for an example). A third group (symbols condition) learned lists of attributes, however, the attributes were listed as pictures of single attributes instead of verbally. Such lists of symbols produced no format effect compared to verbal lists in Bröder and Schiffer's Experiment 3. Thus, we expected a format effect between the image condition and the verbal and symbols conditions.

According to the format hypothesis, having learned attributes as lists should favor the sequential strategy TTB.

We presume that in addition a learning procedure that induces sequential retrieval in the order of cue validity favors strategies that use validity information (TTB and FR). Bröder and Schiffer's learning procedure potentially induced retrieval of cue values in descending order of validity, because their participants learned cue values by sequentially guessing them in descending order of validity. If this procedure induced later retrieval in descending order of validity, their learning procedure might have favored TTB and FR (cf., Bröder & Gaissmaier, in press).

Our participants learned cue values by guessing them in randomly changing order. Thus, our learning procedure neither induced sequential retrieval in descending order of cue validity, nor in any other particular order. Therefore, we expected that our participants' retrieval sequence would conform to the order of validity to a lesser degree. Consequently, decisions conforming to cue validity would not be boosted by our learning procedure. Any TTB and FR classifications would reflect more clearly the deliberate consideration of cue validity.

In addition to strategy classifications, we also measured decision times. If participants classified as TTB users would consider cues sequentially in descending order of validity and would stop as soon as they find a discriminating cue, decision times should be longer if the first discriminating cue is lower in validity (Bergert & Nosofsky, 2007; Bröder & Gaissmaier, in press; Glöckner & Hodges, 2006).

Method

Participants. Seventy-five students at Chemnitz University of Technology participated in the experiment (53 women, 22 men; mean age 23.1). They were either paid the equivalent of $5 or received partial course credit. Three participants had the chance to win another $10 contingent on their performance in the learning phase. 7 participants were replaced, because they did not meet the learning criterion after 90 minutes, 4 participants quit the learning phase earlier and were also replaced.

Design. Presentation format (image, symbols, and verbal) was varied between subjects. Participants were assigned randomly to three groups enforcing equal numbers in each condition.

Materials. We constructed 10 travel offers for family trips (see Table 1). Each consisted of the name of the organizing company and a combination of values on four cue dimensions: main attraction (animal park, botanic garden, or mosque), transportation (flight, bus, train, or car), accommodation (resort, hotel, or vacation home), and sport facilities (tennis, pool, or fitness). In the learning phase of the experiment, participants memorized these 10 offers without knowing the optimal cue values. In the decision phase, we presented pairs of these offers for binary choice. Participants were provided with information about optimal cue values and cue validities framed as results from marketing research and were asked to indicate which of two offers should be included in the catalogue of a travel agency.

Table 1: The decision alternatives that were presented in three different formats (image, symbols, and verbal) in the learning phase of the experiment. In the decision phase, only the names of alternatives were presented in paired choice items. The optimal combination of cue values was presented after the learning phase and determined the cue pattern of alternatives.

		Cue 1: Main Attraction	Cue 2: Transport	Cue 3: Accommodation	Cue 4: Sport	Cue pattern
	Optimal cue values	animal park	flight	resort	tennis	++++
Alternative	Name					
1	Eberhardt	animal park	bus	resort	tennis	+-++
2	Sander	botanic garden	flight	resort	tennis	-+++
3	Mayer	animal park	flight	hotel	fitness	++--
4	Hansen	animal park	car	hotel	tennis	+--+
5	Pohl	botanic garden	train	resort	tennis	--++
6	Wieland	animal park	train	resort	pool	+-+-
7	Berentz	mosque	flight	hotel	tennis	-+-+
8	Jäger	botanic garden	flight	resort	pool	-++-
9	Schönherr	animal park	bus	vacation home	fitness	+---
10	Zimmermann	mosque	flight	vacation home	fitness	-+--

Table 2: Item types and predicted choices between alternatives A and B. Take the Best (TTB) choices are determined by the first differentiating cue, Dawes' Rule (DR) chooses the alternative with more positive cue values, Franklin's Rule (FR) weighs cues according to their validity.

	Item type 1		Item type 2		Item type 3	
Alternative	A	B	A	B	A	B
Cue 1	+	-	-	-	+	-
Cue 2	-	+	+	-	+	+
Cue 3	+	+	-	+	-	-
Cue 4	+	-	-	+	-	+
TTB choice	A		A		A	
DR choice	A		B		A or B (50%)	
FR choice	A		B		A	
Guessing	A or B (50%)		A or B (50%)		A or B (50%)	

Table 3: Percentages of participants classified as using a certain heuristic (Take the Best, Dawes' Rule, Franklin's Rule or Guessing) based on the pattern of their choices among alternatives in items of type 1, 2 and 3 (Bayesian strategy classification following Bröder & Schiffer, 2003a).

	Strategy classification					
	TTB	DR	FR	Guessing	unclassified	N
Condition						
Image	24.0	72.0	4.0	0.0	0.0	25
Symbols	36.0	36.0	24.0	4.0	0.0	25
Verbal	36.0	28.0	24.0	4.0	8.0	25
Total	32.0	45.3	17.3	2.7	2.7	75

As shown in Table 2, three item types can be constructed from the 10 offers. For items of type 1, all three strategies (TTB, DR, and FR) predict the same choice. Only item types 2 and 3 are informative. Among all possible 45 binary choice items, 19 are of item type 1, 9 are of item type 2, and 17 are of item type 3. We presented all 45 possible pairs of offers and repeated the 9 items of type 2. In the repeated items, the order of offers was reversed. Thus the decision phase consisted of 54 binary choices after two practice trials.

In the image condition, travel offers were presented as integrated images in the learning phase. On a landscape background, smaller pictures were arranged that indicated single cue values (see Figure 1). In the symbols condition, the same pictures were used as symbols in lists of cue values. In the verbal condition, lists of verbal labels for cue values were presented.

Figure 1: Example of a travel offer as it was learned in the image condition.

Procedure. In the learning phase, participants went through three levels. First, they learned single travel offers. They were shown one offer. Then the cue values disappeared and they had to construct the offer cue by cue. For each cue, cue values were presented for selection. Participants received acoustic feedback and the correct value was added to the offer. Learning of single offers proceeded until the first reconstruction without errors. Cues were tested in random order and cue values for selection were shown in random order.

After 5 single offers had been learned, these 5 were tested and practiced as in single offer learning until a criterion of 95% correct cue value selections was achieved. Then, the cue values of these offers were presented without the name of the organizing company and participants had to select the correct name of the organizing company. This naming practice continued until all five offers were named correctly. Then, the same learning and testing procedure was completed for the remaining 5 offers.

Finally, all 10 offers were tested with cue reconstruction and naming until 90% correct selections were achieved. The order of offers was random in learning and testing. The assignment of offers to sets of five was balanced.

After the learning phase, participants were asked to imagine they had to decide which of two presented offers a travel agency should select for its catalogue. Information about cue validity was framed as the result of marketing research. Participants were told that the main attraction was rated the most important feature of family travel offers and that most customers would prefer the attraction animal park. The second important feature was transportation with a preference for flights, third was accommodation with a preference for resorts, and fourth was sport facilities with a preference for tennis.

In each trial of the decision phase, the names of two organizing companies were presented next to each other in the bottom half of the screen together with a display of the optimal cue values in the top half of the screen. Participants responded with two keys on a standard keyboard.

After the decision phase, participants completed a memory test that was similar to testing in the learning phase. In addition, in the image condition, the memory of participants for the offers was tested as a manipulation check. They were presented with two images and were asked to select the original image. The distractors contained the same small pictures representing cue values just at different, but still plausible locations in the image.

Results

On average, the participants completed the learning phase after 55.3 min ($SD = 44.4$). Learning took more time in the image condition ($M = 64.1$ min) than in the symbols and the verbal conditions (52.9 and 49.4 min, respectively). However, due to large interindividual differences this effect of presentation format was not statistically significant, $F(2, 71) = .72$, $p = .49$. The subsequent decision phase took on average 13 min in all three conditions.

In the memory test after the decision phase, participants' mean performance was 86.5% ($SD = 1.5$), which suggests that they remembered the alternatives during the decision phase quite well. Performance in the memory test did not differ significantly between format conditions, $F(2, 72) = 0.60$, $p = .55$. Participants in the image condition could discriminate between original and distractor images after the decision phase. On average, they correctly identified 8.1 of 10 images ($SD = 1.0$) in the manipulation check. Thus, the image condition successfully induced holistic memory representations.

Strategy classification. The format hypothesis predicts a higher frequency of compensatory strategies in the image condition than in the verbal and symbols conditions. To test this prediction, we classified participants' decision patterns with the same method as Bröder and Schiffer (2003b). This Bayesian strategy classification method is comprehensively presented and discussed in Bröder & Schiffer (2003a). We provide a brief description in the Appendix. The frequencies with which participants were

classified as using TTB, DR, FR, or Guessing are shown in Table 3.

The symbols condition and the verbal condition induced strategies with similar frequencies, as expected. However, the format hypothesis was not clearly supported. TTB classifications were only slightly less frequent in the image condition (24%) than in the symbols and verbal conditions (36%). Nonetheless, presentation format affected the frequency of strategies.

In the image condition, DR clearly was the predominant strategy, whereas in the symbols and the verbal conditions, DR was only slightly more frequent than FR. Thus, the frequencies of TTB, DR, and FR significantly differed between the three conditions, $\chi^2(4, N = 71) = 10.25$, $p = .04$. The symbols and the verbal conditions did not differ significantly from each other, $\chi^2(2, N = 46) = 0.16$, $p = .92$, whereas the image condition differed significantly from the symbols as well as from the verbal condition, $\chi^2(2, N = 49) = 7.15$, $p = .03$, and $\chi^2(2, N = 47) = 8.86$, $p = .01$, respectively.

Decision times of TTB users. To test whether the decision times of participants who were classified as TTB users increased if cues high in validity did not discriminate and lower cues had to be considered, we split the set of 54 items into four sets, depending on the position of the best discriminating cue in descending order of validity. For each participant we computed the median response time in each of the four sets. For example, if in an decision item the most valid cue had the same value in both alternatives and the second valid cue had differing values, this item would belong to the set of "Cue 2" items.

Table 4: Mean decision times (and standard errors) in seconds as a function of presentation format and of the first discriminating cue in descending order of validity for participants who were classified as TTB users (N = 24)

| Condition | First discriminating cue | | | |
	Cue 1	Cue 2	Cue 3	Cue 4
image	15.82	18.33	14.51	13.79
(N = 6)	(5.27)	(6.10)	(4.69)	(4.86)
symbols	5.06	7.20	5.60	5.26
(N = 9)	(0.47)	(1.01)	(1.04)	(0.71)
verbal	4.49	7.81	5.60	6.54
(N = 9)	(0.61)	(1.85)	(1.48)	(2.38)

Table 4 shows the means of these individual response time medians across the 24 participants who were classified as TTB users. Clearly, the median decision times did not increase across the four item sets, however, decision times were higher in the image condition. This effect of presentation format was confirmed in a 3 (condition) x 4 (item set) mixed ANOVA, $F(2, 21) = 4.535$, $p = .023$. In all conditions, decision times were increased for items, in which the second cue was the first discriminating cue. This resulted in a significant effect of item set, $F(3, 63) = 5.385$, p = .002. Condition and item set did not interact significantly, $F(6, 63) = .919$, $p = .488$.

Discussion

We obtained a clear effect of presentation format on strategy frequencies, however, our results differ from the prediction of the original format hypothesis. TTB was only slightly less frequent in the image condition than in the verbal and the symbols condition. Thus, the original format hypothesis that predicts a lower frequency of TTB compared to compensatory strategies (DR and FR) in the image condition was not clearly supported. However, DR alone was more frequent in the image condition than in the verbal and symbols conditions and only a single participant in the image condition was classified as using FR.

DR is the compensatory strategy that ignores cue validity. FR is compensatory as well, but factors in cue validity. Thus it seems that those participants in our image condition who considered several cues (those not classified as TTB users) did not weigh cues according to validity. This is consistent with prototype matching as it would be expected with holistic images that do not signal cue validity. A shift towards DR indicates prototype matching with such images even more convincingly than a shift towards DR and FR, because pure image matching does not include a consideration of cue validity. The low frequency of TTB users in the verbal and symbols conditions presumably was a consequence of the learning procedure.

Overall, our participants were less often classified as TTB users than Bröder and Schiffer's (2003b) participants. Consequently, TTB was not more frequent than DR and FR combined in the verbal and in the symbols condition. We presume that our learning procedure, which - other than in Bröder and Schiffer's experiments - did not induce sequential retrieval in descending order of cue validity, reduced the frequency of TTB users (cf. Newell & Shanks, 2003). If cue values are less strictly retrieved in the order of validity, cues lower in validity might receive more attention. In turn, the frequency of compensatory strategies DR and FR increases and the frequency of TTB decreases. The decision times of TTB users did also suggest that memory retrieval did not strictly follow the order of cue validities.

Bröder and Gaissmaier (in press) report decision times for TTB users in the experiments reported in Bröder and Schiffer (2003b). Decision times were the longer the lower in validity the first discriminating cue was (see also Bergert & Nosofsky, 2007; Glöckner & Hodges, 2006). They also report an experiment, in which learning order did not conform to validity as in our experiment. In this experiment, decision times did not increase with descending validity of the first discriminating cue. This finding is consistent with decision times of TTB users in our experiment. Participants classified as TTB users thus did not perform sequential cue retrieval in the exact order of validity and with the strict stopping rule of TTB. The high decision times of TTB users in the image condition suggest that those participants who learned images and

tried to compare cues one by one were at a disadvantage compared with participants who chose a similar strategy with learned lists.

To summarize, we found more holistic information integration with a disregard for cue validity when decision alternatives were learned as integrated images. This result is consistent with the basic idea of Bröder and Schiffer's format hypothesis, however, our participants did not receive support from the learning procedure in the consideration of cue validity. Hence, the overall frequency of TTB classifications was lower in our experiment.

Conclusions

We presented evidence that the presentation format as well as the learning procedure affect information integration in multi-attribute decision making. Both factors influence the accessibility of information and thus the ease with which information can be retrieved and the cognitive processes by which it is processed. We have demonstrated that the extent to which presented information enters multi-attribute decisions can be increased or diminished. Thus, the strategies and heuristics in human decision making not only vary depending on how well they are adapted to the information structure of a domain given the boundaries of human information processing (Gigerenzer et al., 1999), but also depending on how the information about a domain is learned and represented in memory. With regard to the possibility of improving the quality of human decisions, this deserves attention, because memory representations are easier to optimize than the information structure of a real-world domain.

References

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 107-129.

Bröder, A., & Gaissmaier, W. (in press). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review.*

Bröder, A., & Schiffer, S. (2003a). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making, 16*, 193-213.

Bröder, A., & Schiffer, S. (2003b). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General, 132*, 277-293.

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

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

Glöckner, A., & Hodges, S. D. (2006). Strategy selection in memory based decisions: Simplifying fast and frugal heuristics versus weighted compensatory strategies based on automatic information integration. Manuscript submitted for publication.

Juslin, P., & Persson, M. (2002). PROBabilities from EXemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science, 26*, 563-607.

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 53-65.

Appendix: Bayesian Strategy Classification

As Bröder & Schiffer (2003b), we presented participants with binary choice items each consisting of two alternatives. The alternatives differ with respect to their values on four cue dimensions. Table 2 presents three types of decision items and the predicted choices of TTB, DR, FR, and Guessing. The cues are listed in descending order of cue validity.

The core assumption of the classification method is that one of the three strategies or random guessing is used, but that deviations from a strategies' predictions occur with an unknown error rate ε that does not change across item types.

The likelihood function of a decision pattern given a strategy k (k = TTB, DR, FR, or Guessing) can then be computed by a joint binomial function. Let n_j be the number of items of each item type j. Let n_{jk} be the number of choices in item type j predicted by strategy k. The number of choices in item type j not predicted by strategy k is then (n_j - n_{jk}). Let ε_k be the error probability of choosing the option not predicted by strategy k. Then the likelihood of a decision pattern given a strategy k is determined by

$$L_k = p\left(n_{jk} \mid k, \varepsilon_k\right) = \prod_{j=1}^{3} \binom{n_j}{n_{jk}} \times \left(1 - \varepsilon_k\right)^{njk} \times \varepsilon_k^{(nj-njk)}$$

(A1)

The unknown error probability ε_k can be determined by

$$\hat{\varepsilon}_k = \left[\sum_{j=1}^{J}\left(n_j - n_{jk}\right)\right] \div \left[\sum_{j=1}^{J} n_j\right]$$ (A2)

Because DR has to guess with item type 3, the index j in Equation A2 only runs from 1 to 2, if k = DR. For the same reason, ε_k = 0.5 in Equation A1 for the case k = DR and j = 3. For k = Guessing, all error probabilities are set to .5 in Equation A1 and Equation A2 is not needed to compute them.

Likelihoods for all three strategies and Guessing are determined in this way. A decision pattern is then classified as presumably being generated by a particular strategy, if all likelihood ratios in favor of this strategy are larger than 1. Otherwise, the pattern remains unclassified.