Information Aggregation in Groups: The Approach of Simple Group Heuristics (SIGH)

Torsten Reimer (t.reimer@mpib-berlin.mpg.de)
Center for Adaptive Behavior and Cognition (ABC),
Max Planck Institute for Human Development, Lentzeallee 94, D-14195 Berlin, Germany

Ulrich Hoffrage (hoffrage@mpib-berlin.mpg.de)
Center for Adaptive Behavior and Cognition (ABC),
Max Planck Institute for Human Development, Lentzeallee 94, D-14195 Berlin, Germany

Abstract

A new framework is introduced that models group decision making by using simple group heuristics (SIGH). We report results of a set of simulations that systematically varied (a) the group members' strategies (compensatory unit weight model, UWM, and a noncompensatory lexicographic heuristic, LEX), (b) the distribution of cue validities (J-shaped vs. linear), and (c) the quantity and quality of shared information. Individual decisions were aggregated by using a majority decision rule (proportionality in case of ties). (1) The simulations revealed strong effects of the distribution of cue validities on group performance. When validities were linearly distributed, UWM gained an 8% better accuracy than LEX by considering all cues. Yet, if cue validities followed a J-shaped distribution, the much more frugal LEX surpassed the UWM by achieving a 16% higher accuracy. (2) This effect was robust across different quantities of shared information. (3) Systematic allocation of information in favour of valid or invalid cues showed that the performance of UWM mainly depended on mean validity, whereas the performance of LEX was more strongly affected by the degree to which the most valid cues were shared.

Introduction

A widely spread assumption in research on group decision making is that good decisions require exhaustive information processing. For example, studies on the hidden-profile effect revealed that group decisions may be strongly affected by the quantity of shared information, that is, by the amount of information that is known by all individual group members at the onset of a group decision process (cf. Stasser & Titus, 1985; see Wittenbaum & Stasser, 1996, for an overview). However, it may be questioned whether this robust empirical finding supports the general claim that less information yields inferior group decisions: In research on the hidden-profile effect, the correct choice is not defined by an outside criterion, but by a particular strategy–namely by a unit weight linear model, UWM, which sums up the cue values of each object and chooses the alternative with the highest sum score (Dawes & Corrigan, 1974). A consequence of taking UWM as the gold standard is that (a) group decisions cannot get better if they are based on a partial set of information, and (b) judgements that deviate from this standard are not only classified as distinct but as inferior.

In the simulations reported in the present paper, we deviated from this tradition and introduced an outside criterion. Each alternative in a choice task was described by several cues that were probabilistically related to the criterion. This allowed us to compare the performance of several strategies for cue based inferences, where performance was measured in terms of correct predictions of the outside criterion rather than their matches of UWM's choices.

Contrary to the claim that using more information yields a better performance, recent comparisons between compensatory strategies and noncompensatory heuristics suggested that this is not necessarily true (Czerlinski, Gigerenzer, & Goldstein, 1999). Noncompensatory heuristics such as Take The Best (TTB) search sequentially for information and stop as soon as a cue is found that discriminates between a given pair of alternatives. The decision is then based only on this cue (one reason decision making). TTB can successfully compete with much more complex strategies (such as multiple regression or unit weight models): It is not only more frugal but sometimes also reaches an even higher percentage of correct inferences (see also Martignon & Hoffrage, 2002).

In the present work, we apply this framework of fast and frugal heuristics as proposed by Gigerenzer, Todd, and the ABC Research Group (1999) to group decision making, a field that has not yet been treated within this framework. Conversely, research on group decision making has not yet paid much attention to decision strategies used by the individual group members or to the validities of the available cues (for an exception, see Gigone & Hastie, 1996, who conceptualized group decision making in the framework of the Brunswikian lens model). Hence, the present approach of Simple Group Heuristics (SIGH) bridges a gap and promises to make a contribution to both fields.

Task

The task is as follows: Consider a four-member personnel committee that has to decide which of three candidates is best suited for a position (cf. Davis, 1973; Stasser & Titus, 1985). These three candidates are randomly drawn from a reference class that consists of 20 potential job applicants (for the concept of a reference class, see Gigerenzer, Hoffrage, & Kleinbölting, 1991). Each candidate has a particular value on an objective criterion that allows to rank order the candidates according to their qualification. Thus, there exists a correct decision for every potential triple of candidates with whom the group may be faced. Group members do not know the criterion but have information on 20 dichotomous cues (cue values are '+1' or '-1') that are positively related to the criterion (i.e., a positive value indicates a higher qualification than a negative cue value does). Each group member has access to a certain amount of information on the three candidates (unknown cue values are indicated by '0'). Based on the information a group member has, he or she is able to form an individual decision first. Subsequently, the group integrates the individual decisions into one group decision.

Decision strategies for individuals In the simulations reported below, we compared the performance of three decision strategies a group member may use: (1) a *unit weight model* (UWM), (2) a simple heuristic that eliminates candidates on the basis of randomly drawn cues (*elimination by random cue*, ERC), and (3) a *lexicographic heuristic* (LEX) that looks up cues according to their validity.

UWM, which serves as the gold standard in research on the hidden-profile effect, is based on sum scores. UWM has been used as a model for individual (e.g., Bröder, 2000; Rieskamp & Hoffrage, 1999) as well as for group decision making (e.g., Stasser & Titus, 1985; for a model that considers unequal weights, see Gigone & Hastie, 1996). For each alternative (here, candidate), UWM computes an overall score by summing up the values on all available cues (which amounts to subtracting the number of negative cue values from the number of positive cue values, while ignoring the number of unknown cue values). UWM then predicts that the candidate with the higher sum score has the higher criterion value. Note that this strategy uses all available information and is compensatory in that cues with positive values can be compensated by those with negative values.

In contrast, the heuristics defined by Gigerenzer et al. (1999) are simpler to execute because they do not require any computation, and they are more frugal because they stop information search as soon as one cue that discriminates between the alternatives has been found. The elimination by random cue heuristic (ERC) draws cues randomly. ERC consists of the following building blocks:

Search rule: Draw a cue randomly (among those that have not yet been used) and look up the cue values of all candidates who are still in the choice set.

Stopping rule: Eliminate all candidates who have a lower value than the candidate with the maximum value. (Thereby, unknown cue values are treated as a third category, i.e., both the pairs '1/0' and '-1/0' are assumed to discriminate). If only one candidate remains or if all cues have been looked up already, then stop search and proceed with the next step; otherwise search for another cue.

Individual decision rule: If all but one candidate are eliminated after search has been stopped, predict that this candidate is the one with the highest criterion value. If more than one candidate is left but search cannot be continued because all cues have already been looked up, then choose randomly among the remaining candidates.

If a decision maker has an idea about the validity of the cues, he or she may draw the cues not randomly but in an order established by their (perceived) validity. The lexicographic heuristic (LEX) first compares the three candidates on the basis of the most valid cue (following Gigerenzer et al., 1991, validity is defined as the proportion of correct inferences in a complete paired comparison between the objects of the reference class; for other definitions see Martignon and Hoffrage, 2002). If this cue discriminates by allowing an unequivocal decision, no further information is considered and the candidate with the highest value is chosen. If two candidates share the highest cue value, the third competitor is excluded from the choice set, and competition among these two continues (for other heuristics that may be used to choose among more than two alternatives, see Rieskamp & Hoffrage, 1999).

Decision strategies for the integration of individual decisions into a group decision How can groups integrate the individual decisions into a joint group decision? (1) Groups may ignore the individuals' decisions and randomly draw an alternative (equal probability), (2) draw alternatives with probabilities proportional to the number of group members who favour them (proportionality), (3) decide for a correct decision if it is favoured by at least one group member (truth wins), or (4) use a voting strategy (majority).

Research on the social decision scheme approach (see Davis, 1973; Laughlin & Ellis, 1986; Hinsz, Tindale, & Vollrath, 1997) has consistently shown that groups adapt their social combination rule according to the task at hand. In general, if a task has a correct solution (intellective task), which is known by at least one group member who is able to demonstrate its correctness (Laughlin & Ellis, 1986), then the likelihood is high that this group member will dominate the group decision (truth-wins scheme). On the contrary, if a task has no demonstrable correct solution (judgmental task), groups are more likely to apply a voting rule (majority scheme). Because the current simulation focussed on the individual decision strategies, we restricted ourself to the majority rule. This rule seems to be plausible given that the current task is a judgmental task that has an objectively correct solution that is not demonstrable, because none of the group members has knowledge about the candidates' criterion values.

Group decision rule: Predict that the candidate with the most votes has the highest criterion value. If there is a tie with respect to the number of votes, then choose the decision of one group member by random.

Environments: Distributions of cue validities

The performance of a particular strategy and hence the result of a comparison between strategies may depend on the environment in which this performance is evaluated (Martignon & Hoffrage, 2002). To check for the robustness of our results across different environments and to explore the ecological rationality of the strategies introduced above, we ran the simulations in four different types of environments (generated with a random error method, constrained such that different distributions of cue validities result; see Figure 1).

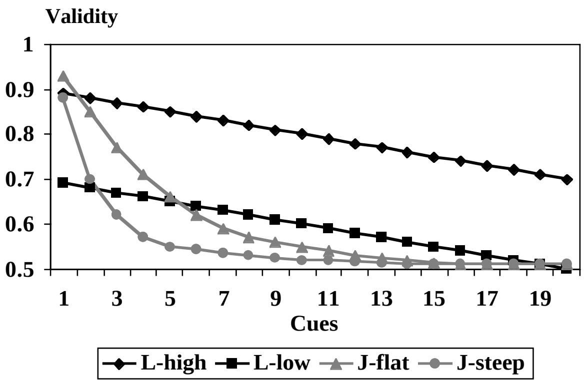

Figure 1: Distributions of cue validities.

In two of the four environments, the distribution of cue validities is linear (L), and in the other two they follow a J-shaped distribution (J). J-shaped distributions are ubiquitous: Not only the values of many continuous variables of the objects in many environments follow such a distribution (Hertwig, Hoffrage, & Martignon, 1999), but the validities of dichotomous cues also tend to be J-shaped, as a reanalysis of the environments used in Czerlinski et al. (1999) has shown. The two linear distributions that we generated for the present simulation differ with respect to their overall means (L-high vs. L-low), whereas the two J-shaped distributions differ in their skewness, which mainly affects the validity of the most valid cues (J-flat vs. J-steep).

As shown in Table 1, the L-high distribution has a much higher mean validity than the other three. Secondly, the L- and J-distributions differ systematically in their standard deviations (*SDs*) according to their skewness. The validities in the L-low and the J-flat distributions have equal means, but their SDs differ by a factor of two.

Table 1. Mean cue validities based on pairs (V_P) and triples (V_T), and correlations between both.

	Distributions of cue validities			
	L-high	L-low	J-flat	J-steep
	Validity based on pairs of candidates (V_P)			
M (V_P)	.80	.60	.60	.56
SD (V_P)	.06	.06	.12	.09
	Validity based on triples of candidates (V_T)			
M (V_T)	.68	.43	.42	.37
SD (V_T)	.08	.09	.17	.13
r(V_P x V_T)	.93	.90	.99	.96

In each environment and for each cue, the discrimination rate was kept constant at the possible maximum (i.e., half of the candidates had a positive and half had a negative value on each cue throughout). The validities were computed on the basis of all *pairs* of candidates: The validity of a cue was determined by dividing the number of pairs in which the candidate with the higher criterion value also has the higher cue value by the number of pairs in which two candidates have different cue values. Because this task consisted of choosing among three candidates, we additionally computed cue validities on the basis of all possible *triples* by dividing the number of triples in which the best candidate had the highest cue value (correct decisions) by the number of triples in which any candidate had a higher cue value than the remaining two candidates (number of discriminating cases).

As shown in Table 1, the pattern of triplewise validities (V_T) turned out to be very similar to the pattern of pairwise validities (V_P), and the correlations between both are very high (at least .90). The main difference between these two measures consists in their mean values. The V_T yields lower means but higher SDs throughout.

Because the accuracy of LEX was not strongly affected by the validity measure according to which cues have been ordered, only the results based on the ordering established by V_P will be reported in the remainder.

Overview of the simulations

The first simulation compared the accuracy of individuals' decision strategies and the resulting group decision for the four distributions of cue validities when each group member knew all available information. The second simulation systematically varied the quantity of shared information by providing group members with a certain amount of randomly chosen cue values. Finally, the third simulation tested whether the quality rather than quantity of shared information affected group performance.

Group members did not have contradictory information in any of the simulations (i.e., all had either veridical or no information) and the group as a whole always had access to all available information (i.e., each piece of information was known by at least one group member). For each of the four environments, all possible triples of candidates (n=1140) were generated. Then, the available information for each triple was randomly distributed among group members according to the constraints of the respective condition (of Simulation 2 and 3). For each single condition (defined by the distribution of cue validities, the respective triple, and the number of shared information), ten runs were realized. When applying LEX, each group member used the same order of cues, namely the one based on the validities computed in the environment.

Simulation 1: Does the distribution of cue validities matter?

The first simulation compared the three strategies in each of the four environments. In this simulation, group members shared all available information. Accuracy is reported both for the decisions of the individual group members (Ind) as well as for group decisions based on the majority rule (Maj).

Linear distributions We first turn to the two linear distributions. As Table 2 shows, the strategies differ in the number of cues they consider. Whereas the UWM used all 20 cues, the simple heuristics used on average not more than 3 cues to form a decision. The small difference between ERC and LEX in the L-high condition (2.4 vs 3.0) is due to the fact that within the set of cues with very high validities, candidates were more likely to have the same cue values as compared to the set of cues with moderate or low validities. What costs do the simple heuristics have to pay for their frugality? In the condition of high (low) cue validities, UWM gains a 12% (5%) increase in accuracy by considering all cues. LEX and ERC achieve almost equal accuracy if cues have a high validity (1% difference). Regardless of what strategy is used, performance is higher in the L-high than in the L-low distributions—the differences vary between 22% (LEX) and 28% (UWM). Finally, group and individual decision making yields almost identical results unless ERC is used. If cues are drawn randomly, the majority rule yields somewhat more accurate decisions (4% difference on average).

Table 2. Performance of three decision strategies, averaged across all individuals (Ind) and across group decisions (Maj) in two types of environments with linear distributions of cue validities.

	L-high		L-low	
	Frugality	Accuracy	Frugality	Accuracy
	Unit weight model (UWM)			
Ind	20.0	88	20.0	60
Maj	20.0	89	20.0	61
	Elimination by random cue (ERC)			
Ind	2.4	70	2.2	46
Maj	2.4	77	2.2	51
	Lexicographic (LEX)			
Ind	3.0	78	2.2	56
Maj	3.0	78	2.1	56

Note. Performance is measured in terms of frugality (average number of cues looked up) and accuracy (% correct).

J-shaped distributions If decisions are made in environments with cues that follow a J-shaped distribution with respect to their validities, the results are different (see Table 3).

Table 3. Performance of three decision strategies, averaged across all individuals (Ind) and across group decisions (Maj) in two types of environments with J-shaped distributions of cue validities.

	J-flat		J-steep	
	Frugality	Accuracy	Frugality	Accuracy
	Unit weight model (UWM)			
Ind	20.0	55	20.0	46
Maj	20.0	55	20.0	46
	Elimination by random cue (ERC)			
Ind	2.3	43	2.3	38
Maj	2.2	47	2.2	39
	Lexicographic (LEX)			
Ind	2.4	73	2.2	61
Maj	2.4	73	2.2	61

Overall, the flat J-shaped distribution leads to more accurate decisions than the steep distribution in which cues are less valid on average (cf. Table 1). And, again, the noncompensatory heuristics are much more frugal than UWM. However, in these J-shaped environments, LEX *gains* from ignoring most information by yielding an 18% (15%) higher accuracy in the flat (steep) J-shaped distribution than UWM, which differs from ERC by only 8%.

Simulation 2: Does the quantity of shared information matter?

The second simulation tested whether this difference between the strategies is robust across different percentages of shared information. Recall that in Simulation 1, each group member had access to all available information. Does performance decrease if group members have less information? The second simulation systematically varied the quantity of shared information under the restriction that each piece of information was always known by at least one group member. Thus, in the most extreme case, in which no single piece of information was shared by group members, each of the four members received 15 (25%) of the 60 cue values. This number was systematically increased in ten steps by adding 5 cue values per step (20, 25, 30, ... up to 60). Table 4 shows the percentage of correct decisions (a) when each group member had access to half of the information (50%, i.e., when each piece of information was, on average, shared by two group members); (b) when all information was unshared (25%); and (c) across the nine cases, in which group members only had access to a partial set of information (M(25-92%)).

Table 4. Accuracy (% correct) of group decisions according to quantity of shared information.

	Distributions of cue validities			
	L-high	L-low	J-flat	J-steep
	Unit weight model (UWM)			
M (25-92%)	85	59	53	45
50%	85	58	53	45
25%	82	57	51	43
	Elimination by random cue (ERC)			
M (25-92%)	66	45	42	38
50%	64	45	42	37
25%	62	43	42	37
	Lexicographic (LEX)			
M (25-92%)	68	52	69	60
50%	65	51	67	60
25%	68	54	66	56

As the results show, the quantity of shared information does not matter much (the largest difference amounts to 4%) even though performance is somewhat better if all information is shared (see Tables 2 and 3), in particular in the L-high condition. Thus, the relationships between the decision strategies and the environments reported above remain stable and are robust across different amounts of shared information.

Simulation 3: Does the quality of shared information matter?

By now, group members' distributions of cue validities have, on average, matched the distributions of cue validities in the environment. To see what happens if this match is distorted, we ran another set of simulations in which information was distributed in a biased way such that information either on the most valid or on the least valid cues had a higher chance of being shared (see Table 5).

(3a) In this simulation, the available information was first randomly distributed among group members as before. Each group member then filled up his or her set of known cue values to 50% by randomly choosing additional information from the ten most or ten least valid cues (most vs. least valid cues). As can be seen in Table 5, this variation did not exert strong effects on the performance of ERC and LEX. However, it strongly affected the performance of UWM in the two environments, in which average cue validities were moderately high (i.e., in the L-low and J-flat environment; see the framed numbers in Table 5, Simulation 3a).

Table 5. Accuracy of group decisions according to the validity of shared information.

	Distributions of cue validities			
	L-high	L-low	J-flat	J-steep
Simulation 3a				
	Unit weight model (UWM)			
Most valid cues	85	69	60	44
Least valid cues	84	48	46	43
	Elimination by random cue (ERC)			
Most valid cues	67	47	46	40
Least valid cues	64	42	39	36
	Lexicographic (LEX)			
Most valid cues	68	52	70	62
Least valid cues	67	54	67	56
Simulation 3b				
	Unit weight model (UWM)			
Ten most valid cues	83	73	65	46
Ten least valid cues	81	44	42	40
	Elimination by random cue (ERC)			
Ten most valid cues	78	56	52	42
Ten least valid cues	74	44	40	37
	Lexicographic (LEX)			
Ten most valid cues	78	56	73	61
Ten least valid cues	70	42	31	39

(3b) Because LEX looks up cues in an order established by their validity, this heuristic should be mainly affected by the degree to which the most valid cues are shared. In order to demonstrate this relationship, in Simulation 3b, three group members had all information either on the ten most valid or on the ten least valid cues (ten most vs. ten least valid cues). The remaining cue values were allocated to the fourth group member. Whereas in Simulation 3a and 3b, UWM performed almost equally well, this is only partly true for the noncompensatory heuristics. If a majority of group members only had access to the ten least valid cues, performance of LEX dropped dramatically in line with the assumption that the accuracy of LEX depends on the degree to which the most valid cues are shared (see the framed numbers in Table 5, Simulation 3b).

Conclusion

Contrary to a widespread claim in research on group decision making, our simulations revealed that good group decisions do not necessarily require exhaustive information processing. The performance of the strategies and therefore also the result of the comparison between the compensatory UWM and the noncompensatory heuristics depended on how cue validities were distributed in the environment. Overall, the accuracy of UWM was mainly affected by mean cue validity, whereas the performance of LEX was more strongly affected by the degree to which the most valid cues were shared. As a consequence, when validities were linearly distributed, UWM gained an 8% higher accuracy than the more frugal LEX. However, if cue validities followed a J-shaped distribution, LEX surpassed the UWM by achieving a 16% higher accuracy on average. The quantity of shared information did not strongly affect accuracy which also supports the claim that less information does not necessarily yield poorer decisions.

Despite the attraction that the frugal, simple noncompensatory heuristics may offer to decision making groups (Reimer & Hoffrage, 2003), it remains to be seen in empirical studies to what extent groups use simple heuristics when forming a decision (for an example, see Reimer & Katsikopoulos, 2003). Thereby, the distributions of cue validities should be taken into consideration given that the performance of strategies depends on such environmental properties.

References

Bröder, A. (2000). Assessing the empirical validity of the 'Take The Best' heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1332–1346.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, and the ABC Research Group, *Simple heuristics that make us smart* (pp. 97–118). New York: Oxford University Press.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review*, *80*, 97–125.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *American Psychologist*, *34*, 571–582.

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.

Gigone, D., & Hastie, R. (1996). The impact of information on group judgment: A model and computer simulation. In E. H. Witte & J. H. Davis (Eds.), *Understanding group behavior: Consensual action by small groups* (Vol. 1, pp. 221–251). Mahwah, NJ: Erlbaum.

Hertwig, R., Hoffrage, U., & Martignon, L. (1999). Quick estimation: Letting the environment do the work. In G. Gigerenzer, P. M. Todd, and the ABC Research Group, *Simple heuristics that make us smart* (pp. 209–234). New York: Oxford University Press.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*, 43–64.

Laughlin, R. P., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology*, *22*, 177–189.

Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, *52*, 29–71

Reimer, T., & Hoffrage, U. (2003). Simulating group decision making by noncompensatory heuristics. *Unpublished manuscript*.

Reimer, T., & Katsikopoulos, K. (2003). Group decision making, recognition heuristic, and less-is-more effect. *Unpublished manuscript*.

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

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*, 1467–1478.

Wittenbaum, G. M., & Stasser, G. (1996). Management of information in small groups. In J. L. Nye & A. B. Brower (Eds.), *What's Social About Social Cognition?* (pp. 3–28). London: Sage.