

The influence of decision heuristics and overconfidence on multiattribute choice: A process-tracing study

Selart, Marcus; Kuvaas, Bård; Boe, Ole; Takemura, Kazuhisa

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Selart, M., Kuvaas, B., Boe, O., & Takemura, K. (2006). The influence of decision heuristics and overconfidence on multiattribute choice: A process-tracing study. *European Journal of Cognitive Psychology*, 18(3), 437-453. <https://doi.org/10.1080/09541440500173054>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-SA Lizenz (Namensnennung-Weitergabe unter gleichen Bedingungen) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-sa/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-SA Licence (Attribution-ShareAlike). For more information see: <https://creativecommons.org/licenses/by-sa/4.0>

**The influence of decision heuristics and
overconfidence on multiattribute choice:
A process-tracing study**

Marcus Selart

Norwegian School of Economics and Business Administration, Norway

Bård Kuvaas

Norwegian School of Management, Norway

Ole Boe

Göteborg University, Sweden

Kazuhisa Takemura

Waseda University, Japan

In the present study it was shown that decision heuristics and confidence judgements play important roles in the building of preferences. Based on a *dual-process* account of thinking, the study compared people who did well versus poorly on a series of decision heuristics and overconfidence judgement tasks. The two groups were found to differ with regard to their information search behaviour in introduced multiattribute choice tasks. High performers on the judgemental tasks were less influenced in their decision processes by numerical information format (probabilities vs. frequencies) compared to low performers. They also looked at more attributes and spent more time on the multiattribute choice tasks. The results reveal that performance on decision heuristics and overconfidence tasks has a bearing both on *heuristic* and *analytic* processes in multiattribute decision making.

Correspondence should be addressed to Marcus Selart, Department of Strategy and Management, Norwegian School of Economics and Business Administration, Breiviksveien 40, N-5045 Bergen, Norway. E-mail: Marcus.Selart@nhh.no

This research was financially supported by grant no. F0637/97 to the first author from the Swedish Council for Research in the Humanities and Social Sciences.

© 2005 Psychology Press Ltd

<http://www.psypress.com/ecp>

DOI: 10.1080/09541440500173054

THE FREQUENTIST APPROACH TO RATIONAL THOUGHT

There is a growing body of literature showing that frequentist representations cause various cognitive biases to disappear related to confidence judgements (Cosmides & Tooby, 1996) and to decision heuristics. Confidence and frequency judgements allude to the broader concept of *metacognition* (for overviews, see Chambers, Izaute, & Marescaux, 2002; Metcalfe, 2000; Metcalfe & Shimamura, 1994; Perfect & Schwartz, 2002; Yzerbyt, Lories, & Dardenne, 1998). In a generic sense, metacognition designates our knowledge of our own knowledge and memory or, more specifically, the monitoring and control of the processes and outputs of our cognitions (Fernandez-Duque, Baird, & Posner, 2000; Johansson, 2004; Nelson, 1996; Nelson & Narens, 1990).

It has been claimed by Gigerenzer (1991, 1993, 1994) that applying probability as any form of rational norm, in decision and judgemental research, may be considered controversial. The reason for this is that although probability may be looked upon as a subjective measure or belief, the concept of "probability" may also be interpreted as a series of long-run relative frequencies. This latter position implies, among other things, a refusal to assign probabilities to unique events. In several studies, Gigerenzer has revealed how well-established judgemental errors may be eliminated if questions are asked in terms of frequencies rather than in terms of probabilities. In addition, he has shown that such judgemental errors may be avoided if a procedure of random sample is strictly applied (Gigerenzer & Murray, 1987).

However, there also exist research findings that do not support the frequentist perspective. For instance, it has recently been demonstrated that subjects who made confidence judgements concerning unique events and subjects who estimated relative frequencies produced essentially the same responses (Brenner, Koehler, Liberman, & Tversky, 1996). In the same study, it was shown that both confidence judgements and frequency estimates exhibited substantial overconfidence, and that these measures were both highly correlated with independent judgements of representativeness (Brenner et al., 1996). Still, it must be noted that only one type of decision heuristic was used in this study.

The role of frequencies and probabilities in confidence judgements and decision heuristics has until now been the focus of much research. For instance, it has previously been revealed by Gigerenzer that presentation mode of uncertainty has an impact on performance on confidence judgements and on decision heuristic tasks themselves.

However, little is yet known about how performance in these tasks governs to what degree people are sensitive to presentation format (frequencies/probabilities) in multiattribute choice. We therefore investigate

if performance on judgemental tasks (such as decision heuristics and overconfidence tasks) has an influence on peoples' sensitivity to probability-based or frequency-based presentation of uncertainty in multiattribute choice tasks. The new feature of the present study is that a numerical presentation format is assumed also to play a role in multiattribute choice. More specifically, it is assumed that general performance on decision heuristic tasks and overconfidence tasks governs this influence.

Heuristic processes

Emphasising that performance on judgemental tasks should also matter to reasoning in multiattribute choice tasks, we build on recent dual-processes accounts on thinking. Such accounts reveal that there exist individual differences in rational thought on quite a general basis (Evans & Over, 1996; Sloman, 1996; Stanovich & West, 1997, 1998, 2000). These individual differences are particularly present in processes that are characterised as fast, automatic, and largely unconscious to their nature (often referred to as System1 processes). As a common feature, these processes are relatively undemanding of computational capacity. They are also considered rather personal and contextualised in that they take individual goals and pragmatic considerations into account. Based on these characteristics they are often termed *heuristic*. Deriving from a common process similarity in this respect (confidence vs. frequency judgements), it could be argued that performance on decision heuristics and confidence judgement tasks would have a bearing on performance on multiattribute choice tasks (Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1996; Tversky & Kahneman, 1974). The reason is that these judgement types share the same heuristic process specificity with the multiattribute choice tasks.

We thus hypothesise:

H1. High performers on decision heuristic tasks and on overconfidence judgement tasks will be less influenced by the numerical presentation format (frequencies vs. probabilities) in multiattribute choice, than low performers on such tasks.

Analytic processes

The other processes proposed by dual-process theorists (System2) are considered as slower, strategic, and conscious to their nature (Evans & Over, 1996; Sloman, 1996; Stanovich & West, 1997, 1998, 2000). These processes pose higher demands on processing capacity and operate in an impersonal and decontextualised way, obeying logic or some other normative system. They are therefore often labelled as *analytic*. The dual-

process theorists agree that System1 processes have a huge influence on everyday judgement but that abstract and hypothetical System2 processes can guide our reasoning on specific tasks toward normative answers (Verschueren, Schaeken, & d'Ydewalle, 2004). Hence, high performers on heuristics tasks and overconfidence tasks are also assumed to use more accurate decision strategies in their decision processing (Payne, Bettman, & Johnson, 1993). This means that high performers on judgemental tasks to a higher extent are supposed to use rational decision processes. Such processes display attention patterns that can be tied to systematic weighting and utility maximisation. Thus, it is also hypothesised that:

H2. High performers on decision heuristic tasks as well as on overconfidence judgement tasks are expected to use more optimal decision strategies, compensatory in nature (characterised by a refusal to make tradeoffs), in multiattribute choice.

PROCESS TRACING METHOD

In order to test the two proposed hypotheses, a process tracing method was applied using eye tracking equipment. In this connection, attention-based latency time was used as a means of measuring the cognitive processes occurring in multiattribute choice tasks (see Lohse & Johnson, 1996). Compared with other more traditional process tracing methods like verbal protocols and information boards, eye tracking has in recent years received an increasing interest in, for instance, the organisational behaviour and marketing research fields.

METHOD

In the present study, participants were initially presented with 24 tasks measuring biases connected to the representativeness, availability, and anchoring and adjustment heuristics (Tversky & Kahneman, 1974). In addition, the tasks were designed to measure biases related to the use of the attribution heuristic (Plous, 1993). Each task was designed so that participants were first instructed to choose one of two options, knowing that one of the answers was correct and the other one incorrect. Subsequently, they were instructed to make a confidence judgement on a half-range scale, indicating how sure they were of having chosen the correct answer (see, for example, Fischhoff, Slovic, & Lichtenstein, 1977; Lichtenstein & Fischhoff, 1977; Oskamp, 1965). The main reason for adding the dimension of confidence to the fulfilment of the heuristic tasks was that it would add critical information about participants' ability to carry out accurate judgements.

Based on the outcome of this test, it was possible to subdivide the participants according to their performance. High performers were characterised by a low degree of biases with regard to the achievement on the heuristic tasks (availability, representativeness, anchoring and adjustment, attribution). Another feature was that they were also quite well calibrated. Low performers were on the other hand producing a high degree of biases on the heuristic test, and were also characterised by being not so well calibrated.

After having completed this initial paper and pencil block of tasks, participants were asked to perform a number of computerised multiattribute preference tasks. A paper and pencil condition was also conducted for control purposes. All the tasks involved deciding about candidates for a job position, and the participants were asked to take the role of a consulting adviser for a company involved in personnel recruitment. In these tasks, uncertainty was either expressed in terms of probabilities or in terms of frequencies. During the fulfilment of the tasks, both participants' decision processes and their preferential outcomes were registered. A main question of interest was to determine whether participants performing well on the heuristics tasks would also behave more accurately than low performers in the fulfilment of the preference tasks.

Participants

One hundred and ninety-two undergraduates (96 men and 96 women) at Göteborg University participated in the experiment in return for the equivalent of \$7. These participants had on prior occasions indicated that they were willing to take part in the experiment. The mean age of the participants was 24.4 years ($SD = 3.6$), which fell in range between 18 and 39 years.

Materials

All participants were first requested to complete a paper and pencil test booklet. It consisted of a diagnostic test measuring the degree of biases in different decision heuristics. The test booklet consisted of 24 questions. These questions were adopted from earlier research on heuristics and biases (Kahneman & Tversky, 1973; Tversky & Kahneman, 1973, 1974, 1982). The 24 questions were divided into four decision heuristic categories dependent upon which decision heuristic they measured. Each of the four decision heuristic categories consisted of six questions. The four decision heuristic categories included in the booklet were representativeness, availability, attribution, or anchoring and adjustment. Each question always had one correct and one incorrect answer. Subsequent to having given an answer, participants were requested to give a confidence rating on a scale ranging

from 50 (making a guess) to 100 (absolutely sure) as to how sure they were that they had given the correct answer. The participants were then administered either a computerised version of a job recruitment task or a paper and pencil version of the same task. The Eyegaze System (cf. Boe, Selart, & Takemura, 2000; Lohse & Johnson, 1996) eye tracking equipment was used together with the computerised version of the job recruitment task. Each scenario contained information about job candidates expressed with reference to eight different attributes. Four of these attributes concerned profit goals (e.g., improving the company's production, share of the market and profit, and increasing sales) (Cyert & March, 1963). The remaining four attributes were related to environmental goals (e.g., decreasing the company's effluent level, improving the company's working environment, environmental policy, and energy saving). An example of one of the four multiple event scenarios given to participants in the conditions with probability-based information can be found in the Appendix. The eight candidates' ability to obtain these goals was for half of the participants expressed on a probability scale ranging from 1% to 100%. The other half of the participants were instead presented with information in terms of how frequent it was that the candidates would achieve the goals, for instance in 3 cases out of 10. All participants were given a total of eight different problems.

Design

The design was mixed factorial with frequency-based versus probability-based information as one of the between-subjects factors. A second between-subjects factor consisted of an experimental control condition measuring participants' responsibility for their performance. A within-subjects factor (profit vs. environmental attributes) was used as a control variable with no connection to the hypotheses.

Procedure

Participants attended the experiment individually in laboratory conditions. They were seated in a private booth in front of a computer screen and were requested to first fill out the test that measured the degree of biases in different decision heuristics. After having completed the test, participants were randomly assigned to either an eye-gaze condition or a paper and pencil version of the same task. For the participants in the eye-gaze condition, a calibration procedure was thoroughly performed so that the Eyegaze System could be used in the experiment. This calibration procedure usually took about 2 min for each new participant.

The participants were then given general instructions on how to perform the experiment. They were also instructed that their task in the experiment was to act as a job recruiter and that they had a variety of different organisations in trade and industry as their clients. Participants were told that their task was to make decisions about job candidates (in some cases groups of candidates). In this respect, it was made clear that their decisions were to be based on as thorough a judgement as possible. Participants were also instructed that the different candidates or groups of candidates would differ in the degree to which they could fulfil a certain company's goal. The task in each situation was to select the best four candidates or groups of candidates for a post in an organisation. The participants were also informed that the four candidates they selected would continue to further interviews or analyses.

In the general instruction it was stressed that participants did not have to rank order the chosen alternatives. Moreover, participants were told that it did not matter in which order they were selected. All participants were explicitly instructed to carefully consider all information presented on the screen while making their choices.

After having considered the information they pressed the return button and typed in their choices. Thereafter, they pressed the return button again and another scenario was presented. Participants assigned to the paper and pencil version simply wrote down their choices on the bottom of each page before continuing to the next page. The different environmental or profit attributes as well as the positions of the different groups of candidates were randomised for each scenario. Each participant also received a randomised presentation of the scenarios. After having participated in the sessions, participants were debriefed and paid. The sessions lasted for approximately 50 min.

Measures

Measures of the heuristics tasks. The performance measure was constructed by summing the number of times the respondents answered correctly across the 24 different heuristic questions. The correct answers were coded as 1, and the incorrect as 0. If participants chose the correct answer, the corresponding confidence rating was treated as positive; otherwise it was coded as a negative value. An index measure of confidence was obtained by taking the mean values of the confidence ratings of the same 24 questions. All participants performing above the mean value on the choices ($M = 11.20$, $SD = 2.60$) and on the confidence ratings ($M = -4.53$, $SD = 17.47$) were coded as high achievers, and those performing below or equal to the mean values were coded as low achievers. In this way, it was possible to create two groups of participants, one consisting of high achievers with high accuracy and calibration, and another group of low achievers with low accuracy and calibration.

Measures of the preference task. The selected alternatives were in both the conditions (the process tracing and the paper and pencil condition) for each participant and task assigned the score of 1 while the rejected alternatives were assigned the score of 0. No reactivity was found for the process tracing condition in the sense that *t*-tests revealed that the mean response scores for a clear majority of the alternatives did not reliably differ between conditions.

Recoding of the eye fixations data. In order to investigate whether high achievers used more compensatory decision strategies than low achievers, analyses of the Eyegaze recordings were made. Depth of search refers to the total amount of information that is searched (Ford, Schmitt, Schechtman, Hults, & Doherty, 1989; Klayman, 1983; Payne, 1976; Svenson, 1979). In the present experiment, only depth of search was used to examine whether participants used compensatory or noncompensatory decision strategies, due to limitations in the Eyegaze recorder's processing software. Another strategy measure used in the present study was the participants' response latency time (in ms).

A mean value of latency time for each attribute in ms was constructed. Each attribute was measured eight times since there was a total of eight problems for which each of the attributes could be attended to in each problem. A second mean value of the four profit attributes and a third mean value of the four environmental attributes were likewise constructed (both in ms).

The time required to acquire information using eye fixations varies between 200 ms and 400 ms (Card, Moran, & Newell, 1983; Russo, 1978). An index was therefore constructed based upon the mean value (300 ms) of these two endpoints. All the participant's attention that required less than 300 ms upon an attribute was coded as 0, whereas if it required more than 300 ms it was coded as 1. Summing the four profit attributes (now recoded as 1 or 0) that participants had been attending to, a measure indicating the number of attributes that had received attention was created. This ranged from 0 to 4. The same procedure was used to construct another measure for the four environmental attributes (also recoded as 1 or 0). In this way, it was possible to investigate the degree to which participants focused upon profit or environmental attributes.

RESULTS

The presentation mode of uncertainty

High achievers were expected to be less affected than low achievers by whether the information was probability or frequency based. A 2 (group: high achievers vs. low achievers) \times 2 (condition: probability-based vs. frequency-based information) \times 2 (attributes: profit vs. environmental attributes) mixed

ANOVA with repeated measures on the last factor performed on the number of attributes yielded a significant main effect of attributes, $F(1, 64) = 8.63$, $p < .01$, $MSE = 6.55$. This effect again confirmed that participants searched for more information concerning profit goals than concerning environmental goals. A main effect of group was also found, $F(1, 64) = 5.04$, $p < .05$, $MSE = 15.26$, revealing that high achievers attended to the attributes reliably more than low achievers. Separate Bonferonni-corrected t -tests at $p = .05$ performed on the high achievers revealed that no significant differences existed between the probability- or frequency-based information conditions regarding the time spent on searching for information concerning the two types of attributes or concerning the number of attendance paid to these. There were no significant differences between high achievers in the probability- or frequency-based conditions for any of the single profit or environmental attributes. In line with this, low achievers showed no significant differences in the time spent on the environmental or on the profit attributes. However, whether or not low achievers had been searching for information about an attribute was found to have some effects on the time used. Additional separate Bonferonni-corrected t -tests at $p = .05$ revealed that low achievers' search for information concerning some of the attributes (the profit attribute of improved share of the market, and the environmental attributes of improved energy saving, decreased effluent level, and improved working environment) differed in the frequency- and probability-based information conditions. Evidently, in the majority of these cases low achievers revealed a reliably higher value for searches of probability-based information, but one effect was in the other direction. The results therefore suggest that qualitative differences may exist in the cognitive processes applied by low performers depending on the nature of the attribute/dimension. Table 1 shows the mean percentages of low achievers' searches for information for the above-mentioned attributes. The other environmental and profit attributes, as well as the number of environmental or profit attributes that were searched for revealed no significant differences between the two conditions.

Bonferonni-corrected t -tests at $p = .05$ on the preference data confirmed Hypothesis 1. As shown in Table 2, low achievers differed significantly between the frequency- and probability-based information groups. Such a difference was not observed for the high achievers.

The decision heuristic test

High achievers were expected to use more compensatory decision strategies when processing information than low achievers. Table 3 shows the mean numbers of attributes attended to by type of goal for high and low achievers. As can be seen, high achievers searched for the information by using more

TABLE 1
Mean values (scale ranging from 0 to 1) of high and low achievers' searches of attributes and *t*-test effects for participants in the frequency- and probability-based information conditions

| <i>Attributes</i> | <i>Probabilities</i> | <i>Frequencies</i> | <i>Effects</i> |
|-------------------------------|----------------------|--------------------|--------------------------|
| Low achievers | | | |
| Improved productivity | .25 | 0.01 | $t(24) = 1.25, p = .22$ |
| Increased sales | .42 | 0.50 | $t(24) = -0.41, p = .69$ |
| Improved share of the market | .25 | 0.00 | $t(24) = 2.08, p < .05$ |
| Improved profit | .25 | 0.21 | $t(24) = 0.21, p = .84$ |
| Improved environmental policy | .00 | 0.01 | $t(24) = -0.92, p = .37$ |
| Improved energy saving | .25 | 0.00 | $t(24) = 2.08, p < .05$ |
| Decreased effluent level | .25 | 0.00 | $t(24) = 2.08, p < .05$ |
| Improved working environment | .00 | 0.29 | $t(24) = -2.10, p < .05$ |
| High achievers | | | |
| Improved productivity | .48 | 0.33 | $t(40) = 0.93, p = .36$ |
| Increased sales | .38 | 0.38 | $t(40) = 0.00, p = 1.00$ |
| Improved share of the market | .43 | 0.33 | $t(40) = 0.62, p = .54$ |
| Improved profit | .33 | 0.48 | $t(40) = -0.93, p = .36$ |
| Improved environmental policy | .19 | 0.14 | $t(40) = 0.41, p = .69$ |
| Improved energy saving | .48 | 0.33 | $t(40) = 0.93, p = .36$ |
| Decreased effluent level | .48 | 0.33 | $t(40) = 0.93, p = .36$ |
| Improved working environment | .24 | 0.24 | $t(40) = 0.00, p = 1.00$ |

attributes than did low achievers. A 2 (group: high achievers vs. low achievers) \times 2 (attributes: profit vs. environmental attributes) mixed ANOVA with repeated measures on the last factor performed on the number of attributes yielded a significant main effect of attributes, $F(1, 55) = 4.97, p < .05, MSE = 3.70$. This effect revealed that participants searched for more information concerning profit goals than environmental goals. A main effect of group was also found, $F(1, 55) = 5.30, p < .05, MSE = 5.30$. In line with the expectations this effect showed that high achievers attended to more attributes than did low achievers. Separate Bonferonni-corrected *t*-tests at

TABLE 2
Mean percentages (scale 0–100) and *t*-test effects of high and low achieving participants' choices by groups of alternatives

| <i>Choices of candidates</i> | <i>Probabilities</i> | <i>Frequencies</i> | <i>Effects</i> | <i>N</i> |
|-------------------------------------|----------------------|--------------------|--------------------------|----------|
| Low achievers | | | | |
| Profit _{max} alternatives | 60.2 | 49.9 | $t(90) = 2.56, p < .05$ | 96 |
| Environ _{max} alternatives | 39.8 | 50.1 | $t(90) = -2.50, p < .05$ | 96 |
| High achievers | | | | |
| Profit _{max} alternatives | 58.4 | 53.2 | $t(90) = 1.15, p = .25$ | 96 |
| Environ _{max} alternatives | 41.6 | 46.8 | $t(90) = -1.12, p = .27$ | 96 |

TABLE 3
Mean values of number of attributes attended to (scale 0-4) by type of goal for high and low achievers

| | <i>Number of attributes</i> | |
|----------------|-----------------------------|----------------------------|
| | <i>Profit goals</i> | <i>Environmental goals</i> |
| High achievers | 1.44 | 1.15 |
| Low achievers | 0.83 | 0.34 |

$p = .05$ revealed that this difference was reliable for the environmental attributes, $t(66) = 2.67$, $p < .05$, but not for the profit attributes.

Table 4 reveals the mean values (in ms) for time spent on the attributes by type of goal for high and low achievers. Another 2 (group: high achievers vs. low achievers) \times 2 (attributes: profit vs. environmental attributes) mixed ANOVA with repeated measures on the last factor performed on the time spent on searching the attributes resulted in a significant main effect of attributes, $F(1, 55) = 5.52$, $p < .05$, $MSE = 9.03$. Again it was revealed that participants attended more to profit attributes than to environmental attributes. However, the main effect of group did not reach significance, $F(1, 55) = 3.30$, $p = .074$, $MSE = 10.03$. Separate Bonferonni-corrected t -tests at $p = .05$ showed that participants searched the profit attribute information significantly more often than they searched the environmental attribute information, $t(77) = 3.03$, $p < .01$. Furthermore, high achievers were found to reveal a tendency to attend more to the environmental attributes as compared to low achievers, $t(66) = 2.67$, $p = .069$. No significant differences were found for the time attended to the profit attributes.

DISCUSSION

The results revealed that participants performing well on heuristics tasks (availability, representativeness, anchoring and adjustment, attribution) and on confidence judgement tasks also behaved quite accurately in the fulfilment of preference tasks.

TABLE 4
Mean values of time spent on the attributes (in ms) by type of goal for high and low achievers

| | <i>Time spent on the attributes</i> | |
|----------------|-------------------------------------|----------------------------|
| | <i>Profit goals</i> | <i>Environmental goals</i> |
| High achievers | 1615 | 774 |
| Low achievers | 721 | 406 |

Hypothesis 1

For instance, it was revealed that the high performers were not as influenced as the low performers by whether uncertainty was presented in terms of probabilities or in terms of frequencies. Moreover, the high performers spent a more equal amount of time between the two conditions (probability/frequency) searching for information, compared with the low performers. Also, the high performers were not to the same extent biased by the conditions (probability/frequency) in their preferences for any of the groups of alternatives compared with the low performers. This seems particularly to have been the case with people who were performing below the average on decision heuristic and confidence judgement tasks. All these results support recent dual-processes accounts on thinking revealing that there exist individual differences in rational thought on a quite general basis (Evans & Over, 1996; Sloman, 1996; Stanovich & West, 1997, 1998, 2000). They also give some support to the frequentist view (Gigerenzer, 1991, 1993, 1994) in the sense that a frequency format not only matters for performance in decision heuristic tasks but also for confidence judgement tasks. Additionally, the fact that performance on decision heuristics and confidence judgement tasks had a bearing on performance on multiattribute choice tasks corroborates the validity of previous research on decision heuristics (Kahneman et al., 1982; Kahneman & Tversky, 1996; Tversky & Kahneman, 1974). It is hereby also suggested that *heuristic thinking* may be able to explain the different information search behaviour observed between the two groups with regard to presentation format.

Hypothesis 2

Furthermore, it was shown that high performers invested more time in searching for information than did low performers (Evans & Over, 1996; Sloman, 1996; Stanovich & West, 1997, 1998, 2000). It was also established that high performers investigated more attributes in their search for information in comparison with low performers. All these results suggest that the ability to reason in a logically and statistically correct way, combined with good self-calibration, clearly has an impact on the extent to which people use optimal decision strategies in multiattribute decision situations (Verschuere et al., 2004). The findings therefore add credit to the opinion that accurate decision behaviour may be determined by choice of decision strategies (the weighted additive strategy being used as a normative yardstick for accurate decision behaviour; see Payne et al., 1993). Hereby, it is suggested that *analytical thinking* also plays an important role for the observed differences between the two groups.

Limitations

In the present study, we have, in line with Payne et al. (1993), focused on preferences, such as between hypothetical job candidates, rather than on inferences about the real world, such as which soccer team will win or which of two cities is larger. This procedure has been criticised by Gigerenzer, Todd, and The ABC Group (1999). They claim that it is difficult with such a procedure to measure the accuracy of strategies in terms of their ability to predict real-world outcomes. Instead, such a procedure restricts itself to measure accuracy by how closely a strategy can match the predictions of a weighted additive rule, which is the traditional gold standard for rational preferences. According to Gigerenzer et al., fast and frugal heuristics may actually sometimes outperform strategies built on systematic weighting and utility maximisation. It should also be mentioned that Gigerenzer (1991, 1993, 1994) stresses the use of frequencies in a natural sampling framework (i.e., "natural frequencies"), which would entail the frequencies of candidates that achieve particular goals, nested within frequencies of candidates that achieve other particular goals. The results of the present study are still generally supportive of the Gigerenzer et al. paradigm, because they propose that frequencies are more generally a format with which the mind is better able to work.

On the other hand, it may be inferred that there are validity problems tied to the ecological perspective on accuracy presented by Gigerenzer et al. (1999). For instance, it may be the case that the most well-known stocks are those best at predicting financial success on the market in good times, but that this fact may be altered in times of market crisis.

Practical implications

It is evident to Gigerenzer (2002) that statistics are often presented in highly confusing ways, but that our difficulty in thinking about numbers can easily be overcome. With a few helpful techniques we can learn to uncloud our minds, demand helpfully presented information and turn ignorance into insight. In this respect, the choice of presentation format plays an extremely important role. It has been revealed in, for instance, both medical diagnostics and legal practice, that the way vital information is presented has a bearing on judgements and decisions. Here, the knowledge that presentation format (frequencies/probabilities) has different impact on decision-making behaviour for high-performing vs. low-performing rational-thinking groups is a new important fact for practitioners. Also the fact that performance on rational thinking has an impact on information search behaviour in multiattribute choice situations is a result that is highly important for real-world applications.

Future research

A question that remains to be addressed in future research is whether high performers are doing better on the nonfrequency tasks because (a) they are using “System2” procedures and thus directly are making normatively correct inferences, or (b) they are better at mentally converting the stimuli to alternative formats (e.g., frequency) formats and then applying appropriate “System1” procedures. PrEview proof published online month/year

Original manuscript received April 2004
Revised manuscript received December 2004
PrEview proof published online May 2005

REFERENCES

- Boe, O., Selart, M., & Takemura, K. (2000). *Perspectives on information acquisition: Rethinking its role in the construction of reason-based preferences* (Göteborg Psychological Rep., Vol. 30, No. 2). Göteborg, Sweden: Göteborg University, Department of Psychology.
- Brenner, L. A., Koehler, D. J., Liberman, V., & Tversky, A. (1996). Overconfidence in probability and frequency judgment: A critical examination. *Organizational Behavior and Human Decision Processes*, *65*, 212–219.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Chambers, P., Izaute, M., & Marescaux, J.-P. (2002). *Metacognition: Process, function, and use*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1–73.
- Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice Hall.
- Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.
- Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition*, *9*, 288–307.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 552–564.
- Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Performance*, *43*, 75–117.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases”. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 83–115). Chichester, UK: Wiley.
- Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K. I. Manktelow & D. E. Over (Eds.), *Rationality: Psychological and philosophical perspectives* (pp. 284–313). London: Routledge.
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York: Wiley.
- Gigerenzer, G. (2002). *Reckoning with risk*. London: Penguin Books.

- Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Gigerenzer, G., Todd, P. M., & The ABC Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Johansson, M. (2004). *Realism in metacognitive judgments: Effects of social factors*. Unpublished doctoral dissertation, Lund University, Sweden.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591.
- Klayman, J. (1983). Analysis of predecisional information search patterns. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analyzing and aiding decision processes*. New York: North-Holland.
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, *20*, 159–183.
- Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes*, *1*, 28–43.
- Metcalfe, J. (2000). Metamemory: Theory and data. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197–211). London: Oxford University Press.
- Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, *51*, 102–116.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 1–45). New York: Academic Press.
- Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology*, *29*, 261–265.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, *16*, 366–387.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.
- Perfect, T. J., & Schwartz, B. L. (2002). *Applied metacognition*. Cambridge, UK: Cambridge University Press.
- Plous, S. (1993). *The psychology of judgment and decision making*. New York: McGraw-Hill.
- Russo, J. E. (1978). Adaption of cognitive processes to eye movement systems. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and higher psychological functions* (pp. 89–109). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Slooman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.
- Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357.
- Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 161–188.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioural and Brain Sciences*, *23*, 645–726.
- Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, *23*, 86–112.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 153–160). Cambridge, UK: Cambridge University Press.
- Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2004). *A dual-process theory of everyday conditional reasoning* (Tech. Rep. No. 307). Leuven, Belgium: Laboratory of Experimental Psychology.
- Yzerbyt, V. Y., Lories, G., & Dardenne, B. (1998). *Metacognition: Cognitive and social dimensions*. London: Sage Publications.

APPENDIX

An example of a multiple events information scenario given to participants in the select conditions with probability-based information

Imagine that you as an outside consultant are going to choose among eight different groups of candidates applying for a post in an information division in a chemical company. The eight different groups of candidates differ in the degree to which they can be expected to promote certain aims that the company has. The candidate groups' ability to achieve these aims is expressed on a scale ranging from 1 to 100%. We want you to select the four best groups of candidates who will progress to an interview. As an outside consultant you do not have to justify your decision to others and the identity of decision-makers will in the company remain anonymous.

| <i>Groups of candidates</i> | <i>Decreased effluent level</i> | <i>Improved share of the market</i> | <i>Increased sales</i> | <i>Improved working environment</i> | <i>Improved environmental policy</i> | <i>Improved energy saving</i> | <i>Improved productivity</i> | <i>Improved profit</i> |
|-----------------------------|---------------------------------|-------------------------------------|------------------------|-------------------------------------|--------------------------------------|-------------------------------|------------------------------|------------------------|
| Group D | 60% | 50% | 60% | 40% | 40% | 50% | 50% | 50% |
| Group A | 20% | 20% | 30% | 80% | 30% | 80% | 80% | 90% |
| Group E | 70% | 70% | 80% | 20% | 80% | 30% | 30% | 30% |
| Group B | 80% | 30% | 90% | 30% | 70% | 20% | 70% | 30% |
| Group F | 80% | 70% | 20% | 70% | 30% | 20% | 90% | 20% |
| Group C | 40% | 50% | 40% | 50% | 50% | 60% | 60% | 60% |
| Group G | 40% | 40% | 50% | 50% | 50% | 60% | 60% | 50% |
| Group H | 50% | 40% | 60% | 40% | 50% | 60% | 50% | 60% |

Which four candidates do you select?