Title
Processing Time Evidence for a Default-Interventionist Model of Probability Judgments

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Gillard, Ellen
Schaeken, Walter
Van Doorden, Wim
et al.

Publication Date
2009

Peer reviewed

Processing Time Evidence for a Default-Interventionist Model of Probability Judgments

Ellen Gillard (ellen.gillard@ped.kuleuven.be)
Centre for Instructional Psychology and Technology, Vesaliusstraat 2
B-3000 Leuven, Belgium

Wim Van Dooren (wim.vandooren@ped.kuleuven.be)
Centre for Instructional Psychology and Technology, Vesaliusstraat 2
B-3000 Leuven, Belgium

Walter Schaeken (walter.schaeken@psy.kuleuven.be)
Laboratory of Experimental Psychology, Tiensestraat 102
B-3000 Leuven, Belgium

Lieven Verschaffel (lieven.verschaffel@ped.kuleuven.be)
Centre for Instructional Psychology and Technology, Vesaliusstraat 2
B-3000 Leuven, Belgium

Abstract

We report an experiment examining dual-process models of probability judgments. Participants had to judge which of two samples provided the best chances of picking a target. Often, participants based their judgment on a comparison of the absolute frequencies of the targets instead of on the ratios. The comparison of the absolute frequencies of the targets was either congruent with the comparison of the ratios (congruent condition), or not (incongruent condition). In this study, the proportion of incongruent (versus congruent) trials was manipulated. Accuracy rates and reaction times were recorded. Accuracy rates for incongruent items decreased but remained unchanged for congruent items when more congruent items were presented. The most striking reaction time result is that the RTs for correct responses to congruent items drastically increased when fewer such items were presented, whereas the RTs for correct and incorrect responses to incongruent items remained unchanged. This increase in RTs for correct responses to congruent items runs counter to the parallel-competitive model, but can in our opinion be reconciled with a default-interventionist model.

Keywords: dual-process models, probability judgment, reasoning.

Introduction

In this paper we investigated probability judgments. Participants had to judge which of two samples provided the best chances of picking a target. Participants are known to have the tendency to base their responses on a comparison of the absolute frequencies of the target instead of on the ratios (Babai, Brecher, Stavy, & Tirosh, 2006; Falk, Falk, & Levin, 1980; Green, 1983). In other words, they neglect the denominators in their judgment. Sometimes the judgment based on the absolute frequencies of the targets (e.g., black marbles in a box containing both black and white marbles) will be *congruent* with the judgment based on the ratios (e.g., Figure 1a), but the task can easily be modified to create an *incongruent* item (e.g., Figure 1b). Comparison of the absolute frequencies of the targets then no longer results in correct probability judgments. The basic congruency effect (i.e., the lower accuracy for incongruent trials than for congruent ones) has been demonstrated before, but the extension of this experiment comparing dual-process accounts is novel. We will first introduce the general dual-process framework.

According to the dual-process view, people have two systems for decision-making: one fast, automatic, and heuristic based, and the other slower, constrained by working memory (WM) capacity, and analytic (e.g., Evans 2008). Even though heuristic processing often leads to correct judgments, it can also bias reasoning in some situations (the incongruent conditions). Dual-process researchers posit that participants have a tendency to rely on heuristic processes instead of on analytic processes. This way, they can account for participants' failure to provide the normatively correct responses on a range of reasoning tasks.

The dual-process framework is frequently used to interpret performance when judging probabilities. That, is the tendency to rely on a comparison of the absolute frequencies of the targets would result from heuristic or intuitive processes (e.g. Babai, et al., 2006; Reyna & Brainerd, 2008; Stavy & Tirosh, 2000). A judgment based on a comparison of the ratios on the other hand would be the outcome of analytic processes.

Evidence for a dual-process account of performance on this task was found in that correct judgments on congruent items and incorrect judgments on incongruent items have been observed to correspond with smaller latencies than correct responses to incongruent items (Babai et al., 2006). Also, under time pressure the congruency effect increased (Gillard, Van Dooren, Schaeken, & Verschaffel, 2008). That is, more errors were made on incongruent items, whereas accuracy remained unchanged for congruent ones. In the same study, a similar response pattern for accuracy rates was observed when participants were asked to solve the task while simultaneously solving an attention demanding secondary task. Therefore it was concluded that the tendency to neglect the denominators exhibits two quintessential features of heuristic processes: fastness and effortlessness.

Our first goal was to investigate whether the tendency to neglect the denominators is sensitive to a manipulation of *expectations*. We manipulated the proportion of incongruent (versus congruent) trials: One group of participants received a task with 80% incongruent trials whereas the other group received a task where only 20% of the trials were incongruent. If neglecting the denominators is sensitive to the expectation that a comparison of the absolute frequencies of the targets is not sufficient to yield a correct response, we expect a *higher proportion of incongruent (versus congruent) items* to result in a *smaller congruency effect*. That is, we expect fewer errors on incongruent items, but we expect accuracy for congruent items to remain unchanged.

The exact way in which heuristic and analytic processes interact in this type of task has not yet been investigated. Therefore the second goal was to clarify how the processes compete to determine responding by registering the reaction times. In the dual process literature, different models have been proposed about this interaction. *Pre-emptive models* (e.g., Klaczynski, 2000) assume that people employ decision strategies that decide at onset whether the heuristic or analytic route will be followed in order to prevent potential conflicts. *Parallel-competitive models* (e.g., Epstein, 1994; Sloman, 1996), on the other hand, assume that the two pathways are always (simultaneously) activated. If both processes result in the same outcome, that response is given. If not, the conflict must be resolved in order to give the correct response. Another possibility is the *default-interventionist model*, where it is assumed that the heuristic response is the default, unless intervening analytic processes override it (e.g., De Neys & Glumicic, 2008; Evans, 2007; Kahneman & Frederick, 2005). We believe the expectancy manipulation and the latency recordings can help to disentangle the way in which the heuristic and the analytic processes interact in this probability task. We will discuss now each of the three models in more detail.

Pre-emptive models are models of *conflict avoidance* rather than of *conflict resolution* (Evans, 2007). They assume that there is some superficial feature of the problem that cues either heuristic or analytic processes when scrutinizing the problem. It is decided at onset, whether the analytic or heuristic route will be followed. An example is the selective scrutiny model (Evans, Barston, & Pollard, 1983) for syllogistic reasoning. In the case of the probability task under investigation, a pre-emptive model seems psychologically implausible to us because the different items do not differ from each other in any immediately visible way. For this reason, it seems unlikely that by scrutinizing the problem, a decision can be made whether heuristic or analytic processes will be applied.

A parallel-competitive account for this task has been provided by some authors (e.g., Babai et al., 2006). They have suggested that as the absolute frequency of the black marbles is the salient feature, it is automatically processed and compared. This processing would run parallel to the comparison of the black-to-white marbles ratios. When the processing of the absolute frequencies of the black marbles and of the ratios result in the same conclusion (congruent condition), this is the end of processing. If they result in two different conclusions (incongruent condition), the conflict has to be resolved. The analytic processes have the capacity to inhibit the heuristic-based response. But the heuristics can also be so pervasive that people fail to inhibit the heuristic-based response.

Because both processes run simultaneously according to the parallel-competitive account, even when people err on incongruent items, this model predicts that people should be "aware" of the conflict between the outcome of comparing the absolute frequencies of the targets and the ratios (De Neys & Glumicic, 2008; Evans 2007). Such conflict awareness would obviously arise for incongruent items only. This should lead to longer reaction times (RTs) of incorrect responses to incongruent items than for correct responses to congruent items. These latter items should always be very fast because of the absence of conflict. But the RTs for incorrect responses to incongruent items should still be shorter than RTs for correct responses to those items, because in the latter cases the heuristic-based response was successfully inhibited.

The parallel-competitive processing model would predict a smaller congruency effect in the 80%-incongruent condition. More frequent confrontations with the conflict between heuristic and analytic processes might facilitate the inhibition of the heuristic-based response and hence increase accuracy for incongruent items.

As we said before, according to the parallel-competitive model, when no conflict between heuristic and analytic processing occurs, no additional time is needed. For this reason, the model would not predict a difference in RTs for congruent items. Also for the other response types, the model does not predict a difference in RTs between conditions. Hence, the same RT pattern is expected within both conditions: namely that correct responses for incongruent items take longer than incorrect responses on those items, and that incorrect responses for incongruent items take longer than correct responses for congruent ones.

In a default-interventionist model, heuristic and analytic processes are sequentially engaged. An example of a default-interventionist model is Evans' (2006) revised heuristic-analytic model. In such a model, a problem is always processed heuristically first and a default response is provided. After that, analytic processes may or may not be engaged. If engaged, these analytic processes evaluate the heuristic outcome. They can refute the heuristic outcome and alter it, but often they tend to satisfy the heuristic outcome. That is, they accept the heuristic response if it is plausible.

In the case of the probability task we use in our experiment, a default-interventionist model would assume that participants start with comparing the absolute frequencies of the targets. Their judgment will be based on this comparison if no analytic processes are engaged or if the analytic processes merely satisfy the heuristic outcome. Only when the analytic processes refute the heuristic outcome and alter it by comparing the ratios, a correct response will be made.

A default-interventionist model would further predict that heuristic-based responses require less time than analytic-based responses. Unlike in the parallel-competitive model, there is no reason to assume that incorrect responses to incongruent items should be more time-consuming than correct responses to congruent items. A default-interventionist model does, however, predict that correct responses to incongruent items should take longer than incorrect responses to those items.

A default-interventionist model would also predict a smaller congruency effect in the 80%-incongruent condition. The experience that going down the analytic route is worth the effort in most cases (because the default response was incorrect) would motivate the participant to engage more in analytic processes. Evans (2006) names motivation as one of the possible factors that might influence the engagement of analytic processes.

According to a default-interventionist model, manipulating the proportion of incongruent (versus congruent) items should not only affect the congruency effect with regard to accuracy rates but also with regard to the RTs. Since the model predicts that correct responses for congruent items will more often result from analytic processes when more incongruent (versus congruent) items are present, the RTs for correctly solved congruent items should also be longer compared to the condition where fewer incongruent (versus congruent) items are present.

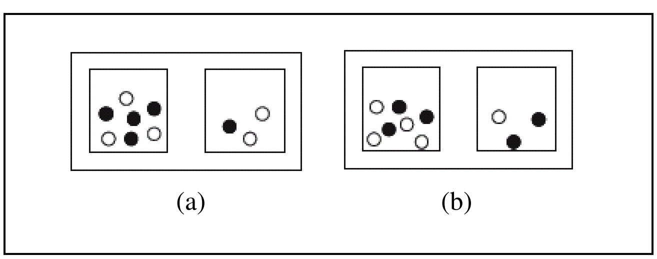

Figure 1: Example of (a) a congruent and (b) an incongruent item.

Method

Participants and design

The participants were 88 first-year (undergraduate) students from the University of Leuven, who participated in return for course credit. They solved congruent and incongruent items. The proportion of incongruent (versus congruent) trials was manipulated. For one group (n = 44) 80% of the trials were congruent trials (and 20% were incongruent ones). The other group (n = 44) received 20% congruent trials (and 80% incongruent ones).

Materials

On each trial, participants were presented with two boxes, both of which contained black and white marbles. Participants were asked to decide which of both boxes had the highest probability of picking a black marble. There were two types of items, congruent and incongruent ones. Figure 1 presents an example of each. For congruent items the box with the highest frequency of black marbles also had the highest probability of picking a black marble. For incongruent items, the box with the highest frequency of targets provided a lower probability of picking a black marble. There were 120 trials. For all items, the difference between both boxes in the true probability of picking a black marble was always higher than 20% but lower than 30%. We also took care that the total numbers of marbles, and the amount of white marbles was not always highest or lowest in the box that was the correct answer.

Procedure

General task instructions were given to familiarize participants with the task and procedure. The first 10 trials were practice trials, thereafter 120 experimental items were presented. For all trials (practice and experimental trials), the proportion of incongruent trials (versus congruent trials) was either 20 or 80%. The inter-stimulus interval was 500 ms (a fixation cross was presented). Participants pressed a key to indicate their response, first box (alphanumeric ‘1’ key) or second box (‘2’ key). No feedback was given.

The trials were semi-randomly presented with the restrictions that (a) the correct response could not be the same on more than three consecutive trials; (b) the intuitive response could not be the same on more than three consecutive trials. After completion of the first block, participants continued with the second block without a pause or warning.

Results

Accuracy Rates

Table 1 presents the basic findings. Participants erred significantly more often on incongruent items. The crucial finding is the interaction with the manipulation of expectation: the congruency effect is larger in the condition with fewer incongruent trials: The difference in accuracy between congruent and incongruent items increases from 18% when 80% of the trials are incongruent to 33% when 20% of the trials are incongruent.

The repeated measurements ANOVA produced a significant main effect of item type: $F(1,86) = 40.64$, $p <$

.01, $p\eta^2 = 0.32$. Condition was marginally significant: $F(1, 86) = 3.86$, $p = .05$, $p\eta^2 = 0.04$. Also the interaction between item type and block was marginally significant: $F(1, 86) = 3.38$, $p = .07$, $p\eta^2 = 0.03$.

Table 1: Accuracy results for congruent and incongruent trials for each condition.

Condition	Congruency	Mean *(stdev)*	
20% incongruent	C	97,10	*(9,99)*
	I	64,11	*(40,14)*
80% incongruent	C	97,92	*(3,87)*
	I	79,71	*(35,34)*

Reaction Times

Figure 2 presents the RTs in each condition for three types of responses: correct responses to congruent items, incorrect responses to incongruent items, and correct responses to incongruent items. In the '20%-incongruent' condition it is clear that correct responses to incongruent items take considerably longer than the other two response types. In the '80%-incongruent' condition, however, it is clear that not only the correct responses to incongruent items but also correct responses to congruent items take considerably longer.

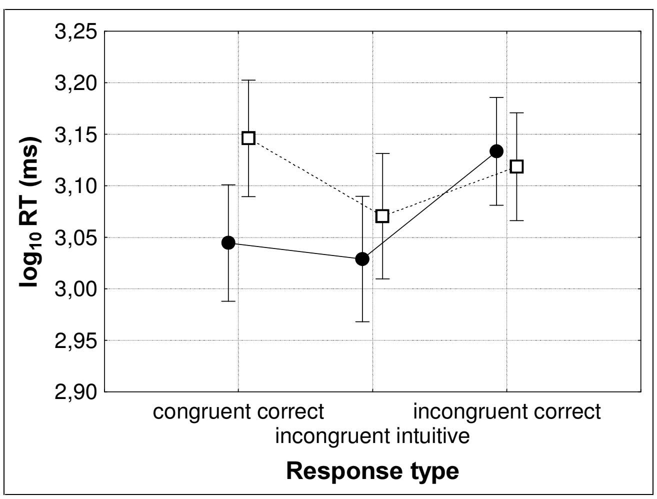

Figure 2: RTs for three response types: correct responses to congruent items (congruent correct), incorrect responses to incongruent items (incongruent incorrect), and correct responses to incongruent items (incongruent correct), for the 80%-incongruent (dotted line) and 20%-incongruent condition (full line).

The repeated measurements ANOVA produced a significant main effect of response type, $F(2, 172) = 9.76$, $p < .01$, $p\eta^2 = 0.10$. There was no main effect of condition $F(1, 86) = 1.50$, $p = .22$, $p\eta^2 = 0.02$. The interaction between response type and condition was also significant, $F(2, 172) = 5.61$, $p < .01$, $p\eta^2 = 0.06$. Missing data in each condition were replaced by the log RT mean of that response type. It should be noted that the RTs for correct responses are based on more trials (97% for congruent items and 79% or 64% for incongruent items) than for incorrect trials (20% or 35% for incongruent items).

Post-hoc comparisons (Duncan's MRT) revealed that the RTs for correct responses to incongruent items in the 20%-incongruent condition are significantly higher than RTs for incorrect responses to those items, $p < .01$. They were also higher than correct responses to congruent items, $p < .01$. Correct responses to congruent items and incorrect responses to incongruent items did not significantly differ in their RTs, $p = .53$.

In the 80%-incongruent condition, the RTs for correct responses to congruent items did not statistically differ from those for correct responses to incongruent items, $p = .29$. Incorrect responses to incongruent items were significantly faster than both correct responses to those items, $p = .05$, and than correct responses to congruent items, $p < .01$.

The significant condition × response type interaction effect was mainly due to a significant increase in RTs of correct responses to congruent items, $p = .02$. For incorrect responses to incongruent items there was no statistically reliable difference between both conditions, $p = .33$. The same was true for correct responses to incongruent items, $p = .71$.

Discussion

Both the parallel-competitive model and the default-interventionist model predicted the increase of the congruency effect in the presence of more congruent items. In the parallel-competitive model, the effect can be explained by facilitation in inhibiting the heuristic-based response when a conflict occurs. A default-interventionist model can explain the finding by stating that in the 20%-incongruent condition, participants more often experience that going down the analytic route yields the same response as staying with the heuristic-based outcome. This makes the monitoring component of the analytic system lax and cause participants to rely more on the heuristic processes in this condition.

The analysis of the RTs favors according to us a default-interventionist model over a parallel-competitive model for two reasons. First, the finding that in the 20%-incongruent condition, the RTs for incorrect responses to incongruent items are similar to the RTs of correct responses to congruent items runs counter to the prediction of the parallel-competitive model. In that model, RTs of the first type of response should be slower than RTs of the second type of response (but still faster than those of correct responses to incongruent items), because for incongruent items, heuristic and analytic processes cue different responses. The resulting conflict should slow down the respondent, even when they fail to resolve the conflict in favor of the analytic outcome. The default-interventionist model, on the other hand, includes a straight heuristic pathway and can therefore account for this RT finding.

Second, the most striking RT result is that RTs of correct responses to congruent items drastically increase when more incongruent (versus congruent) items are present. Again, this finding runs counter to a parallel-competitive model. According to this model, there is no reason why the processing of congruent items should take longer even when those items are less frequent. In a parallel-competitive model heuristic and analytic processes run simultaneously and because for congruent items the processes do not conflict, there would be no need for a deeper processing under any circumstances.

A default-interventionist model on the other hand can account for this finding by claiming that the more frequent experience that going down the analytic route merited the effort (because the outcome was not the same as the heuristic-based outcome), causes participants to also follow this analytic path when solving the less frequent congruent items. This explanation makes sense under the assumption that congruent and incongruent items cannot be distinguished before having compared the ratios of the two samples.

We now give our interpretation of the probability judgment task under consideration according to a default-interventionist model. In such a model, the role of the heuristic processes is in selecting the (seemingly) most relevant information. In this case this would be the absolute frequency of the targets. Many others consider the absolute frequency the most salient property, the 'gist' of the problem information (e.g., Reyna & Brainerd, 2008; Stavy and Tirosh, 2000). The monitoring component of the analytic system then evaluates whether the heuristic-based outcome is reasonable, or whether it should be altered. Often, people stop processing the problem at that point. This means that they will go straight down the heuristic route and base their judgment on the comparison of the absolute frequencies of the targets (i.e., they neglect the denominator, the total number of marbles in each box). The experience that refuting the heuristic-based response and going down the analytic route is not worth the effort, in the sense that the outcome is the same (as is most frequently the case when a majority of the items are congruent), makes people more lax in their responding (and therefore faster but more error-prone), according to a default-interventionist model.

Conclusion

The expectancy manipulation shows that despite the tendency to base ones judgment on a comparison of the absolute frequencies of the targets, this behavior can be controlled to a certain extent. When the proportion of incongruent (versus congruent) items is large, one tends to question the judgment based on the absolute frequencies more often, and alter the heuristic-based outcome by comparing the ratios. That is, by taking the denominator into account. This leads to higher accuracy scores on incongruent items and longer RTs for congruent items. The longer RTs for congruent items can be explained by the fact that congruent and incongruent items cannot be distinguished before having compared the ratios of the two samples. Hence, the heuristic-based outcome is refuted by the monitoring component of the analytic system for these items as well, and the judgment is based on the comparison of the ratios. These findings run counter to a parallel-competitive model but can according to us be accounted for by a default interventionist model.

Acknowledgments

This research was partially supported by Grant GOA 2006/01 "Developing adaptive expertise in mathematics education" from the Research Fund of the K.U.Leuven, Belgium. Wim Van Dooren is a post-doctoral researcher of the Research Fund K.U.Leuven.

References

Babai, R., Brecher, T., Stavy, R., & Tirosh, D., (2006). Intuitive interference in probabilistic reasoning. International *Journal of Science and Mathematics Education, 4,* 627-639.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual-process theories of reasoning. *Cognition, 106,* 1248-1299.

Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49,* 709-724.

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review, 13,* 378-395.

Evans, J. St. B. T. (2007). On the resolution of conflict in dual-process theories of reasoning. *Thinking and Reasoning, 13,* 321-329.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgement and social cognition. *Annual Review of Psychology, 59,* 255-278.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11,* 295-306.

Falk, R., Falk, R., & Levin, I. (1980). A potential for learning probability in young children. *Educational Studies in Mathematics, 11,* 181-204.

Gillard, E., Van Dooren, W., Schaeken, W., & Verschaffel, L. (2008). *The automatic nature of neglecting the denominator in probability judgments: Time pressure and Dual-task considerations.* Manuscript under preparation.

Green, D. R. (1983). A survey of probability concepts in 3000 pupils aged 11–16 years. In D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (pp. 766-783). Sheffield, U.K.: Teaching Statistics Trust.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K.J. Holyoak & R.G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-293). New York : Cambridge University Press.

Klaczynski, P. A. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development. 71,* 1347-66.

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments or risk and probability. *Learning and Individual Differences, 18,* 89-107.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3-22.

Stavy, R., & Tirosh, D. (2000). *How students (Mis-) Understand Science and Mathematics: Intuitive Rules.* Teachers College Press.