Base-rate Respect by Intuition: Approximating Rational Choices in Base-rate Tasks with Multiple Cues

Andreas Glöckner /
Stephan Dickert

MAX PLANCK SOCIETY

Base-rate Respect by Intuition: Approximating Rational Choices in Base-rate Tasks with Multiple Cues

Andreas Glöckner / Stephan Dickert

December 2008

Base-rate Respect by Intuition: Approximating Rational Choices in Base-rate Tasks with Multiple Cues

Andreas Glöckner∗ / Stephan Dickert‡

Abstract

Although intuitive-automatic processes sometimes lead to systematic biases in judgment and choice, in many situations especially this kind of processes enables people to approximate rational choices. In complex base-rate tasks with repeated outcome feedback we observed choices which were in line with the Bayes' solution in 86% of the cases and which were made within a relatively short time (i.e., 2.2 seconds). The results indicate reliance on extremely well-calibrated intuition. This view is supported by the findings that choice proportions are almost perfectly predicted by posterior probabilities ($r = .93$), and that error rates, response times and confidence ratings are highly correlated with inconsistency in the provided information. Our results support the hypothesis that parallel constraint satisfaction models may account for the processes underlying intuition and make the application of simple heuristics and deliberate strategies very unlikely. Taking an interdisciplinary perspective, implications for economic and psychological modeling are outlined.

Keywords: Decision Making, Intuition, Base-rate Neglect, Heuristics, Parallel Constraint Satisfaction

∗   Correspondence concerning this paper should be addressed to: Andreas Glöckner, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn, Phone: +49-(0) 228 / 9 14 16 857, E-mail: gloeckner@coll.mpg.de

‡   Max Planck Institute for Research on Collective Goods, Bonn

Base-rate Neglect and Intuition

To be successful in a fundamentally uncertain world, individuals have to make a multitude of judgments and decisions based on probabilistic information. Research on judgment and decision making has repeatedly shown systematic biases in probability assessment and in applying the rules of probability (Gilovich, Griffin, & Kahneman, 2002; Kahneman, Slovic, & Tversky, 1982; Reyna, 2004; Slovic, 1987). One of the most prominent and well established findings is individuals' neglect of base-rate probabilities.[1] This has been, for instance, demonstrated in the following task (e.g., Bar-Hillel, 1980):

> Two cab companies, Blue and Green, operate in a given city. Eighty-five percent of the cabs in the city are Blue, and the remaining 15% are Green. A cab was involved in a hit-and-run accident at night. A witness later identified the cab as a Green cab. The court tested the witness' ability to distinguish between Blue and Green cabs under nighttime visibility conditions. The witness was able to identify each color correctly about 80% of the time, but confused it with the other color about 20% of the time. What is the probability that the accident was caused by the Green company?

To answer this question, both the base rate of green cabs (15%) and the validity of the eyewitness (80%) have to be taken into account. However, most people consider only the latter and give the answer that signals neglect of the base rate, namely 80%. There have been several attempts to explain base-rate neglect and to provide solutions to solve the problem (for an overview see Barbey & Sloman, 2007). In line with earlier approaches (e.g., Tversky & Kahneman, 1983), Barbey and Sloman proposed that base-rate neglect is due to associative / intuitive (i.e., system 1) judgment strategies that fail to adequately represent the structure of the problem. Our research empirically tests and expands this intriguing theory. In line with Hammond, Hamm, Grassia, and Pearson (1987) and with recent findings in gambling decisions (Glöckner & Betsch, 2008a) and probabilistic inference tasks (Glöckner & Betsch, 2008c) we argue that choices based on intuition approximate a weighted compensatory information integration that can be explained by parallel constraint satisfaction (Glöckner & Betsch, 2008b; Simon, Snow, & Read, 2004). Contrary to research that focuses on highlighting the shortcomings in human cognitive capacity, we aim to show empirically that although sometimes being mislead by the context, these processes can enable individuals to make quick choices that approximate a rational solution according to Bayes' theorem, even in rather complex tasks and without knowledge how to deliberately apply it.

Based on our findings and taking an interdisciplinary perspective, in the discussion we critically address the misleading classification of deliberate processes as "rational" and intuitive processes as "irrational" and provide a different account of the long-standing controversy between (classic) economic modeling and psychological approaches to human decision making concerning the

---

[1] Note that "base-rate neglect" does not mean that base-rates are completely ignored but that they are given not sufficient weight in a judgment or decision (cf. Koehler, 1996).

homo oeconomicus assumption (e.g., Lopes, 1994; Smith, 2005). Finally, we make a suggestion that might allow for more precise behaviorally informed economic modeling of base-rate tasks.

Findings on Base-rate Neglect

One of the early insights in research on base-rate neglect was that people order information by its perceived degree of relevance, and let high-relevance information sometimes dominate low-relevance information (e.g., Bar-Hillel, 1980). The context influences the perceived relevance of base-rates and of other predictive information. In the cab company example, the relevance of the base-rate (i.e., probability of green cabs) for the problem is by many persons perceived to be low. Thus, the mental representation of the problem (Glöckner & Betsch, 2008b; Johnson-Laird, 1983) might be constructed based on the eyewitness information only. However, the relevance of the base-rate can be easily increased, for instance by increasing the specificity of the base-rate information, which reduces base-rate neglect at least partially (Bar-Hillel, 1980). Base-rate neglect seems to be independent of response format. Although base-rate neglect has been classically investigated using judgments of posterior probabilities as dependent variables (e.g., Bar-Hillel, 1980; De Neys & Glumicic, 2008; Reyna & Brainerd, 1995), it has also been observed for choices (e.g., Goodie & Fantino, 1999; Kleiter et al., 1997).

Several factors influence the prevalence and the size of base-rate neglect (Koehler, 1996). For example, monetary incentives can decrease the extent of base-rate neglect (Brase, Fiddick, & Harries, 2006), as well as providing information in natural frequency format (Gigerenzer & Hoffrage, 1995; but see Kleiter et al., 1997). However, the reasons for the latter is likely not a natural capability for frequency processing, as originally proposed, but the mere fact that frequencies sometimes make it easier to construct adequate mental representations (Fiedler, Brinkmann, Betsch, & Wild, 2000; Kleiter et al., 1997). Furthermore, some evidence exists that individuals might be trained sufficiently to take base-rates into account more properly (Sedlmeier & Gigerenzer, 2001). Finally, base-rate neglect has also been investigated in experiments in which individuals were not explicitly provided with probabilistic information but learned base-rates and cue-validities by direct experience (Goodie & Fantino, 1999).

Research Paradigm

Our goal in the current study is to investigate individuals' intuitive choice behavior and the underlying mental mechanisms. Therefore, we use complex tasks with multiple cues and direct feedback. We assume that for such tasks individuals are usually unable to deliberately calculate correct solutions without sufficient prior training and hence have to rely on their intuition or might alternatively use simple deliberate heuristics such a lexicographic or equal weight strategies (Payne, Bettman, & Johnson, 1988). We investigate choices between two options with dif-

ferent base-rates and additional information from three probabilistic cues with explicitly stated predictive accuracy (*a priori cue validity*[2]):

> There is a lottery with two outcomes A and B. Outcome A has a 85% probability of occurring, outcome B occurs with 15% probability. There are three forecasters that make predictions of whether A or B will turn up.
>
> The predictions of the first forecaster are correct in 75% of the cases, the predictions of the second forecaster are correct in 65% of the cases, and the predictions of the third forecaster are correct in 55% of the cases. All predictions are made independently and without knowledge of the distribution of outcomes in the lottery. The predictions are:
>
> > Forecaster 1: B
> >
> > Forecaster 2: B
> >
> > Forecaster 3: A
>
> Which outcome will show up?

The posterior probability for option A, p(A | cues, base-rate), according to Bayes' theorem is

55.4%.[3] Thus, outcome A is more likely and should be selected. In our study, participants worked on a set of such tasks and then were provided with instant feedback whether they selected the correct answer or not.

Hypotheses

According to previous findings in probabilistic inference decisions (Bröder, 2003; Glöckner & Betsch, 2008c) we expected that individuals have the cognitive capability to integrate cues and base-rate information appropriately using intuitive processes. We expected greater base-rate neglect in the early trials but a calibration of the mental representations after repeated outcome feedback. According to the parallel constraint satisfaction approach to intuitive judgment and decision making (Glöckner & Betsch, 2008b; Holyoak & Simon, 1999; Simon, Krawczyk, & Holyoak, 2004) choices should follow a weighted linear integration of base-rate information and cues, which approximates the normative solution according to Bayes' theorem quite well.[4] Response times should peak and confidence should reach its minimum in tasks in which the evi-

---

2 We assume that the a priori cue validity is equal for both outcomes and defined as p(cue predicts A | A)= p(cue predicts B | B).

3 The solution can be calculated by:

$$p(A \,|\, c_1, c_2, c_3, BR) = \frac{p(BR)*(1-p_{c1})*(1-p_{c2})*(p_{c3})}{p(BR)*(1-p_{c1})*(1-p_{c2})*(p_{c3}) + (1-p(BR))*(p_{c1})*(p_{c2})*(1-p_{c3})}$$

$$= \frac{.85*.25*.35*.55}{.85*.25*.35*.55 + .15*.75*.65*.45} = \underline{\underline{.554}}$$

4 Note that other models, such as decision field theory (Busemeyer & Townsend, 1993) or random support theory (Brenner, Griffin, & Koehler, 2005), allow for similar predictions concerning choices.

dence is most inconsistent. Highest inconsistency is reached if base-rate and cue information taken together speak equally strong for option A and B. Hence, the *consistency* of a task was operationalized as the absolute deviation of the posterior probability from .50.

Experiment

Method

Participants and design. One hundred and one students (76 female; mean age = 21.4 years) from the University of Erfurt participated in a 35-minute-long experiment, which was part of a one-hour experimental battery of thematically unrelated studies. Students received a flat payment of 6 Euros (approximately $8.80) for their participation and additionally performance-contingent payments between 0.78 and 2.76 Euros. Participants were randomly assigned to either a deterministic or a probabilistic learning condition (Learning). In the deterministic learning condition, individuals received perfect Bayesian feedback, which means that their choice was considered to be correct if the posterior probability of the chosen option was larger than .50. In the probabilistic feedback condition, the outcome was determined using a random procedure and compared with the choice. If the posterior probability was .80 for option A, option A was drawn with a probability of .80 but option B was drawn with a probability of .20. Accordingly, it was possible to lose even when the normatively correct option was selected (i.e., non-fitting choice). Participants earned 0.02 Euro for each correct/fitting choice and lost the same amount of money for a wrong/non-fitting choice. As introduced in the example above, tasks with a base-rate and three probabilistic cues with dichotomous cue values were used. Tasks were manipulated within participants along the factors Base-rate and Cue Pattern. For Base-rate, all levels from .05 to .95 in steps of .05 were selected. For Cue Pattern, all eight (i.e., 2^3) possible combinations of dichotomous cue predictions were used. This resulted in a 2 (Learning) x 19 (Base-rate) x 8 (Cue Pattern) mixed-model design with the second and third factor being manipulated within participants.

Materials and procedure. The 152 tasks and the respective posterior probabilities are shown in Figure 1. Green (red) cells indicate clear evidence for option A (option B). Yellow cells indicate choices with inconsistent information in which the posterior probability was between .33 and .66. First, our participants were introduced to the task (Figure 2) and to the incentive scheme. They were informed about the (a priori) validity of the forecasters (forecaster 1: 75%, forecaster 2: 65%, forecaster 3: 55%) and that a cue with a validity of 50% has no predictive power. The validity and the presentation order of the forecasters remained stable throughout the experiment. Choices were made by mouse click and the time interval from stimulus onset to the selection of the gamble was recorded (i.e., response time).

Figure 1

Base-Rate A	Prediction Pattern of the three Cues (c1=.75, c2=.65, c3=.55)							
	A-A-A	B-A-A	A-B-A	A-A-B	B-B-A	B-A-B	A-B-B	B-B-B
0.05	0.26	0.04	0.09	0.19	0.01	0.03	0.07	0.01
0.10	0.43	0.08	0.18	0.34	0.02	0.05	0.13	0.02
0.15	0.55	0.12	0.26	0.45	0.04	0.08	0.19	0.03
0.20	0.63	0.16	0.33	0.53	0.05	0.11	0.25	0.04
0.25	0.69	0.20	0.40	0.60	0.07	0.14	0.31	0.05
0.30	0.74	0.24	0.46	0.66	0.09	0.18	0.36	0.06
0.35	0.79	0.29	0.52	0.71	0.11	0.21	0.42	0.07
0.40	0.82	0.34	0.57	0.75	0.13	0.25	0.47	0.09
0.45	0.85	0.38	0.62	0.79	0.15	0.29	0.52	0.11
0.50	0.87	0.43	0.66	0.82	0.18	0.34	0.57	0.13
0.55	0.89	0.48	0.71	0.85	0.21	0.38	0.62	0.15
0.60	0.91	0.53	0.75	0.87	0.25	0.43	0.66	0.18
0.65	0.93	0.58	0.79	0.89	0.29	0.48	0.71	0.21
0.70	0.94	0.64	0.82	0.91	0.34	0.54	0.76	0.26
0.75	0.95	0.69	0.86	0.93	0.40	0.60	0.80	0.31
0.80	0.96	0.75	0.89	0.95	0.47	0.67	0.84	0.37
0.85	0.97	0.81	0.92	0.96	0.55	0.74	0.88	0.45
0.90	0.98	0.87	0.95	0.98	0.66	0.82	0.92	0.57
0.95	0.99	0.93	0.97	0.99	0.81	0.91	0.96	0.74

Figure 1. Choice Tasks and Posterior Probabilities. Manipulation of the base-rates (in rows) and cue pattern (in main columns) result in 152 choice tasks with different posterior probabilities. Tasks with clear evidence for option A (p(A|evidence) > .66) are indicated green, tasks with clear evidence for option B (p(A|evidence) < .33) are indicated red. Tasks with inconsistent evidence (.33 ≤ p(A|evidence) ≤ .66) are indicated yellow.

After each choice, individuals rated their confidence on a horizontal scroll bar ranging from -100 (not confident at all) to +100 (very confident) and they were provided with feedback. The 152 tasks were computer-based and presented in individually randomized order (without practice trials). In a post-test, participants were informed that they could earn an additional 2 Euros by providing the correct answer to another task. In this test, we wanted to measure if individuals were able to determine the exact posterior probability after repeated feedback. The post-test was essentially equivalent to the choice phase except that individuals judged the probability that the outcome of the gamble was A (not B). Individuals had sufficient time and were informed that they could use pen and paper to do any calculations, if they wished. We used slightly modified cue validities, which made it easier to calculate the posteriors because forecaster 3 could be ignored (Base-rate A = .40; forecaster1$_{p=.70}$ = B; forecaster2$_{p=.60}$ = B; forecaster3$_{p=.50}$ = A).

Figure 2

	Gambles	
	A 95% [Select]	**B** 5% [Select]
Fore-caster 1	B	
Fore-caster 2	A	
Fore-caster 3	B	

Figure 2. Presentation format of choice task.

Results and Interpretation

For each choice, a Bayes_Correct score (0 = wrong, 1 = correct) was calculated, which coded whether the choice was in line with the Bayes' solution or not (i.e., if the gamble with the higher posterior probability was selected). For each of the 152 tasks, the posterior probability for option A was coded in the variable Bayes. Furthermore, a task consistency score (Consistency) was calculated that indicates whether base-rate and cues clearly speak for one or the other option: Consistency = | 0.5 – Bayes |. Hence, the Consistency score of each task is the distance between the posterior probability of the options and maximal uncertainty (.5) (see also Figure 1). High consistency scores indicate that one option should be clearly preferred. According to a PCS approach, Consistency should be inversely related to the difficulty of the task and correlate with Bayes_Correct, response time and confidence.

Choice accuracy. The overall level of choice accuracy (i.e., proportion of choices in line with the predictions of Bayes' theorem) was .86 (*SE* = .003). All participants showed an astonishingly good performance of at least .73 (Figure 3, left) across all trials.
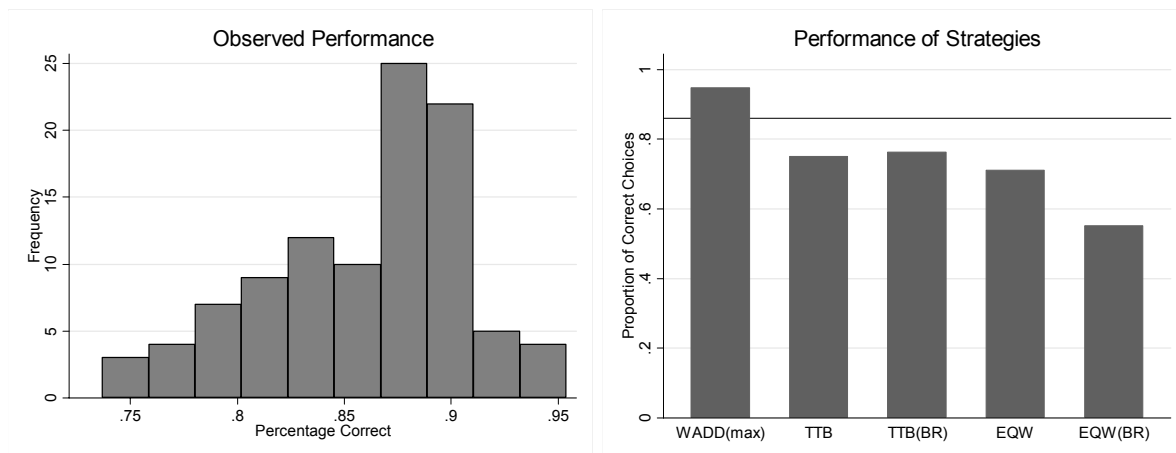
Figure 3

Figure 3. Observed performance of individuals (left) and performance predictions of different strategies (right). The line in the right graph indicates mean observed performance of individuals.

To test whether individuals might have applied simple heuristics or (as we assumed) used intuition which approximates weighted linear information integration we simulated the performance of prominent strategies considered in the literature (Payne et al., 1988; Rieskamp & Hoffrage, 1999) for the 152 tasks. We simulated a lexicographic take-the-best model based on the most valid cue (TTB) and the base-rate (TTB(BR)), equal weight strategies based on cues only (EQW) and including the base-rate as cue (EQW(BR)) (without counting undecided cases), as well as an optimal weighted linear model (WADD) based on a linear regression with base-rate and the three cues as predictors (without constant). Note that the estimated performance of WADD denotes the maximal performance with optimal weights. Individuals which apply other weights score worse. As expected, we found that the performance of WADD was very high indicating that an optimally weighted linear strategy can approximate the correct Bayes' solution in 95% of the cases. The performance of the other strategies was considerably lower (Figure 3, right).

When comparing participants' performance and the performance predictions of the strategies it can be seen that almost all participants performed better than the prediction of the best simple heuristic. This supports our hypothesis that individuals have the cognitive capacity to solve complex Bayes' tasks based on intuition approximating linear information integration.

To jointly analyze the development of performance over time and the effects of our manipulations, we conducted a repeated measurement logistic regression with Bayes_Correct as dependent variable and Learning, Consistency and Trial Number as predictors (Table 1, top).[5]

---

5   In this and in all following analyses we used the STATA option Cluster to account for dependencies in observations caused by repeated measurement. Robust standard errors were used to account for violations of homoscedasticity.

Table 1

Summary of Logistic Regression Analysis for Variables Predicting Choice Accuracy (top) and Choices for Option A (bottom) (*N* = 101)

	Odds Ratio	SE	z	P>|z|
Choice Accuracy				
Learning Condition	1.01	0.099	0.06	0.950
Trial Number	1.00	0.001	0.67	0.500
Consistency	22082.73	5747.17	38.43	0.000
Choices for Option A				
Learning Condition	0.97	0.050	-0.68	0.497
Trial Number	0.9987	0.001	-1.97	0.049
Base-rate	289806.40	132639.8	27.48	0.000
Forecaster 1 (75%)	50.25	10.49	18.76	0.000
Forecaster 2 (65%)	6.19	0.65	17.26	0.000
Forecaster 3 (55%)	2.99	0.24	13.44	0.000

Note. Choice accuracy (0-wrong, 1-correct) refers to choices in line with Bayes' theorem, Pseudo R^2 = .23, Wald χ^2 (3) = 1487.35, $p < .001$ (top). In the regression of choices for option A (0-no, 1-yes) Pseudo R^2 = .62, Wald χ^2 (6) = 927.85, $p < .001$ (bottom).

In line with our expectations, participants showed more accurate choices with increasing Consistency. There was no effect of our manipulation of the learning condition, suggesting that the type of feedback (probabilistic vs. deterministic) made no difference in performance. There was no effect of trial number indicating no learning effects over time (Figure 4). Interestingly, there was not even a significant difference in performance if one compares the proportion of correct answers in the first and the last of the 152 trials, *Odds-Ratio*=1.21, *z*=.69, *p*=.49. Performance of our participants did not improve over time but was very high from the beginning, possibly indicating a ceiling effect.
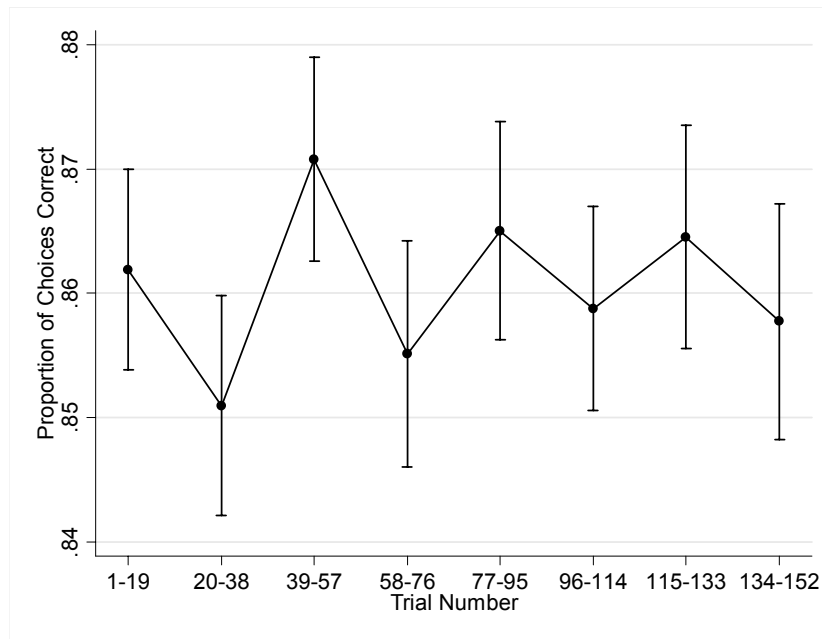
Figure 4

Figure 4. Development of performance over time.

Choices. The proportion of choices for option A and the proportions of choices in line with the rational Bayes' solution are shown in Figure 5. The choice proportions almost exactly matched the posterior probabilities of the tasks (cf. Figure 1). The correlation between choice proportion and Bayes' posterior was $r = .93$, $p < .001$. Deviations from the rational solution were mainly observed in the inconsistent cases (Figure 5, right) as indicated by a correlation of $r = .71$, $p < .001$, between consistency and proportion of correct choices.

Figure 5

A) Choice Proportions for Option A

B) Choices in line with Bayes' Posterior

	Prediction Pattern of the three Cues ($c1$=.75, $c2$=.65, $c3$=.55)								Prediction Pattern of the three Cues ($c1$=.75, $c2$=.65, $c3$=.55)							
Base-Rate A	A-A-A	B-A-A	A-B-A	A-A-B	B-B-A	B-A-B	A-B-B	B-B-B	A-A-A	B-A-A	A-B-A	A-A-B	B-B-A	B-A-B	A-B-B	B-B-B
0.05	0.17	0.00	0.04	0.08	0.00	0.00	0.01	0.00	0.83	1.00	0.96	0.92	1.00	1.00	0.99	1.00
0.10	0.24	0.00	0.04	0.12	0.01	0.00	0.00	0.01	0.76	1.00	0.96	0.88	0.99	1.00	1.00	0.99
0.15	0.30	0.01	0.06	0.24	0.00	0.01	0.00	0.00	0.30	0.99	0.94	0.76	1.00	0.99	1.00	1.00
0.20	0.41	0.03	0.12	0.23	0.00	0.01	0.03	0.00	0.41	0.97	0.88	0.23	1.00	0.99	0.97	1.00
0.25	0.56	0.04	0.16	0.35	0.00	0.01	0.00	0.00	0.56	0.96	0.84	0.35	1.00	0.99	1.00	1.00
0.30	0.54	0.04	0.25	0.41	0.01	0.01	0.07	0.00	0.54	0.96	0.75	0.41	0.99	0.99	0.93	1.00
0.35	0.64	0.03	0.41	0.52	0.00	0.01	0.11	0.00	0.64	0.97	0.41	0.52	1.00	0.99	0.89	1.00
0.40	0.79	0.08	0.58	0.69	0.00	0.00	0.19	0.01	0.79	0.92	0.58	0.69	1.00	1.00	0.81	0.99
0.45	0.86	0.14	0.73	0.81	0.01	0.02	0.30	0.00	0.86	0.86	0.73	0.81	0.99	0.98	0.30	1.00
0.50	0.98	0.38	0.97	0.97	0.04	0.05	0.64	0.02	0.98	0.62	0.97	0.97	0.96	0.95	0.64	0.98
0.55	0.98	0.66	1.00	0.99	0.15	0.23	0.83	0.14	0.98	0.34	1.00	0.99	0.85	0.77	0.83	0.86
0.60	0.99	0.82	0.99	1.00	0.26	0.48	0.89	0.17	0.99	0.82	0.99	1.00	0.74	0.52	0.89	0.83
0.65	1.00	0.85	1.00	1.00	0.42	0.60	0.93	0.28	1.00	0.85	1.00	1.00	0.58	0.40	0.93	0.72
0.70	1.00	0.93	1.00	1.00	0.57	0.72	0.96	0.45	1.00	0.93	1.00	1.00	0.43	0.72	0.96	0.55
0.75	1.00	0.96	1.00	1.00	0.65	0.80	0.97	0.53	1.00	0.96	1.00	1.00	0.35	0.80	0.97	0.47
0.80	1.00	0.96	1.00	1.00	0.72	0.90	0.98	0.59	1.00	0.96	1.00	1.00	0.28	0.90	0.98	0.41
0.85	1.00	0.99	1.00	1.00	0.81	0.90	0.98	0.61	1.00	0.99	1.00	1.00	0.81	0.90	0.98	0.39
0.90	1.00	0.99	1.00	1.00	0.89	0.97	1.00	0.77	1.00	0.99	1.00	1.00	0.89	0.97	1.00	0.77
0.95	1.00	0.99	1.00	0.99	0.94	0.98	0.98	0.85	1.00	0.99	1.00	0.99	0.94	0.98	0.98	0.85

Figure 5. Choice proportions for option A (left) and deviation from rational norm of Bayes' theorem (right). In the choice graph (left), tasks with choices mainly for option A (p(A) > .66) are indicated green, tasks with choices mainly for option B (p(A) < .33) are indicated red and tasks with inconsistent choices (.33 ≤ p(A) ≤ .66) are indicated yellow. In the error graph (right) green, yellow, and red indicate high (p_{cor} > .80), medium (.60 ≤ p_{cor} ≤ .80) and low (p_{cor} < .60) proportions of correct choices.

To test our hypothesis that information was integrated in a weighted compensatory manner we conducted a logistic regression on choices using Learning, Trial Number, the base-rate for option A, and the forecasts of cue 1 to 3 (1 = option A, 0 = option B) as predictors (Table 1, bottom). We observed that base-rates had a strong effect on choices and that cue information from all three cues was considered. The decreasing odds-ratios from cue 1 to 3 indicate that cues were weighted differently according to their cue validity (all three coefficients were significantly different from each other at $p < .001$). This provides additional support for the hypothesis that choices are based on weighted compensatory information integration. We can also reject the hypothesis that base-rates or cue information are neglected by participants.

Response times and confidence. Overall, we observed very short response times. Individuals integrated the base-rate probability and information from three probabilistic cues (according to their validity) in 2.2 seconds (median). This makes it unlikely that choices are based solely on deliberate information integration (Glöckner & Betsch, 2008c; for estimations of the time necessary for deliberate decision making see also Lohse & Johnson, 1996) and supports the hypothesis that intuitive integration processes might have been used. In line with the PCS hypothesis, particularly long response times were observed for inconsistent tasks (Figure 6, left). The log-means of individuals' response time for the 152 tasks are correlated with Consistency at $r = -0.88$, $p < .001$. Participants showed moderately positive confidence ratings with a relatively high variance ($M = 27.1$, $SD = 44.1$). The mean confidence ratings for the 152 tasks are shown in Figure 6 (right). In line with the PCS hypothesis, confidence was particularly low in the inconsistent cases as indicated by a highly positive correlation between confidence and consistency, $r = .87$, $p < .001$.

Figure 6

A) Response Times

Base-Rate A	A-A-A	B-A-A	A-B-A	A-A-B	B-B-A	B-A-B	A-B-B	B-B-B
0.05	2.64	2.18	2.27	2.39	1.90	1.90	2.33	1.77
0.10	2.52	2.34	2.80	2.63	1.82	2.12	2.23	1.73
0.15	2.87	2.34	2.65	2.89	1.98	2.11	2.34	1.89
0.20	2.75	2.35	2.75	3.13	2.03	2.12	2.45	1.83
0.25	2.88	2.37	2.85	2.96	2.18	2.25	2.48	1.82
0.30	2.73	2.51	3.24	3.17	2.05	2.18	2.83	1.77
0.35	2.80	2.57	3.39	3.22	2.29	2.55	2.84	2.03
0.40	2.49	2.60	3.29	3.00	2.29	2.30	2.89	1.97
0.45	2.42	3.03	3.45	2.79	2.40	2.50	3.17	2.03
0.50	2.27	3.80	3.33	2.43	2.49	2.89	3.45	2.32
0.55	2.12	3.39	2.74	2.34	2.78	3.75	3.15	2.46
0.60	2.09	2.98	2.64	2.07	3.11	3.47	2.59	2.67
0.65	1.91	3.26	2.51	2.37	3.06	4.08	2.65	2.91
0.70	1.72	2.71	2.20	1.99	3.18	3.40	2.41	2.79
0.75	1.92	2.80	2.24	2.13	3.06	3.26	2.32	2.65
0.80	1.81	2.51	2.17	1.92	3.21	2.93	2.30	2.91
0.85	1.76	2.75	2.26	1.96	2.90	3.02	2.34	2.69
0.90	1.72	2.44	2.05	1.90	2.75	2.59	2.00	2.68
0.95	1.71	2.33	1.86	1.89	2.63	2.45	2.15	2.50

Prediction Pattern of the three Cues (c1=.75, c2=.65, c3=.55)

B) Confidence Judgments

Base-Rate A	A-A-A	B-A-A	A-B-A	A-A-B	B-B-A	B-A-B	A-B-B	B-B-B
0.05	9.15	51.22	44.33	28.43	75.66	68.93	49.42	84.82
0.10	8.34	48.85	26.78	20.61	65.46	58.26	46.70	81.50
0.15	-1.29	37.02	22.66	0.92	60.55	46.42	31.58	74.89
0.20	-1.73	39.48	10.64	2.91	55.19	47.74	33.96	71.08
0.25	-9.25	31.61	7.06	1.50	49.32	41.76	26.85	64.83
0.30	-1.63	27.31	1.73	-1.65	50.37	42.54	21.43	62.68
0.35	-5.30	19.81	-2.60	-9.19	38.23	31.60	9.15	53.96
0.40	-0.48	13.76	-7.11	-3.52	38.54	27.28	7.10	50.09
0.45	10.57	11.56	0.43	-1.64	25.87	18.05	2.80	38.59
0.50	20.04	-9.19	6.34	15.66	17.95	10.66	-6.61	25.45
0.55	42.87	-1.72	13.14	20.05	1.07	-3.97	0.69	9.23
0.60	49.75	3.31	25.03	32.62	0.46	-3.77	14.66	6.27
0.65	53.96	0.31	29.03	32.84	-3.31	-2.84	13.55	-1.12
0.70	63.50	12.35	39.59	44.60	-7.91	2.13	28.79	-7.50
0.75	64.95	16.80	43.76	52.89	-2.98	5.66	29.25	-6.70
0.80	69.93	22.74	46.08	56.69	3.96	12.72	32.22	-1.57
0.85	74.83	24.40	52.62	52.71	7.10	12.96	35.27	-1.83
0.90	80.36	39.54	60.37	69.10	10.45	30.77	42.29	13.65
0.95	84.10	44.94	65.30	74.79	26.87	32.64	52.40	15.79

Prediction Pattern of the three Cues (c1=.75, c2=.65, c3=.55)

Figure 6. Response times (left) and confidence judgments (right) for the 152 tasks. In the response time graph (left) green, yellow, and red indicate low ($t < 2.8$s), medium (2.8s $\leq t \leq 3.2$s) and high ($t > 3.2$s) times. In the confidence graph (right) green, yellow, and red indicate high ($c > 40$), medium ($10 \leq c \leq 40$) and low ($c < 10$) confidence.   *Post-test.* Probability estimations from the higher incentivized post-test are shown in Figure 7. Although participants took considerable time to work on the post-test and some of them took notes to calculate the result, there was no participant who found the correct answer (i.e., .16). Hence, even after 152 feedback trials none of our participants was able to apply Bayes' rule deliberately.
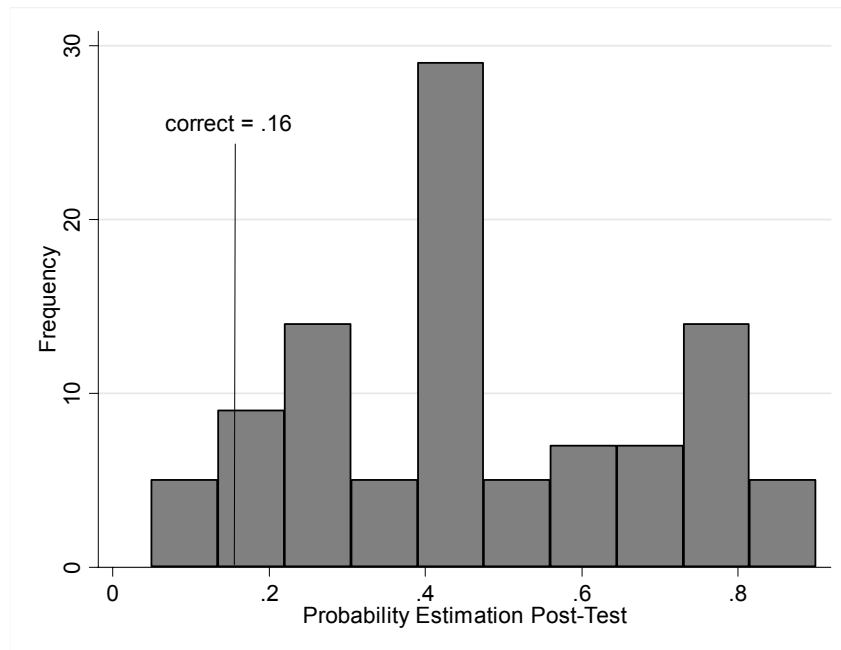
Figure 7

Figure 7. Probability estimation in high incentivized post-test.

We determined individuals' hypothetical choices by dichotomizing probability ratings in estimations below and above .50. We found that according to their probability rating .37 of our participants would have preferred option A, which is clearly the non-preferable option. This error rate is higher than the error rate observed in tasks in the choice phase that had comparable posterior probabilities. Particularly, the proportion of erroneous choices for option A in the (nine) tasks with posterior probabilities between .15 and .20 was .06 which was significantly lower than the proportion in the post-test, $\chi^2(1, N=1007) = 106.7$, $p < .001$. Obviously, participants estimated worse when the response mode was judgment of probability (in contrast to choice) and the stakes were higher. Keeping in mind that participants worked on slightly different tasks, this might indicate that both changes induced more deliberation and hampered the performance in tasks which had before been solved better by relying on intuition (cf. Wilson & Schooler, 1991).

Discussion

The results show that individuals are able to integrate base-rates and multiple probabilistic cues in complex tasks rather quickly. In line with our expectations, participants' choices approximated the rational norms provided by Bayes' theorem quite well, and the posterior probability of an option is highly correlated with the choice proportion for this option ($r = .93$). Regression analyses showed that individuals took into account base-rates and information from all three cues according to their validity. Hence, we observed neither base-rate neglect nor neglect of cue information. The median decision time of 2.2 seconds makes it unlikely that these pieces of information were integrated using extensive deliberate calculations. Specifically, the response time rules out that individuals deliberately applied Bayes' theorem or explicit linear weighted

integration rules (e.g., weighted additive strategy, Payne et al., 1988). Thus, our results indicate that individuals might indeed rely, at least in part, on intuitive processes in base-rate tasks.

In line with hypotheses put forward by Hammond et al. (1987) and derived from PCS models, these intuitive processes take into account information in a weighted compensatory manner that approximates rational choice surprisingly well. Converging evidence for the view that individuals indeed applied intuition based on PCS mechanisms is provided by the significant correlation of consistency with error rate, response time and confidence. Individuals almost instantly recognize if the provided information is very clear (consistent) or equivocal (inconsistent) and adapt their response time and confidence ratings accordingly.

The results of the post-test indicate that increasing deliberation can reduce choice accuracy. According to the PCS model this could be explained by the fact that deliberation, although often being beneficial, can sometimes highlight specific but not necessarily relevant pieces of information and thereby impair calibration of the mental representation (cf. Dijksterhuis, Bos, Nordgren, & van Baaren, 2006).

Learning Effects and Adaptation of Mental Representations

Possibly due to ceiling effects, we did not observe learning effects over time. The calibration of the mental representation underlying individuals' intuition was good from the beginning. It was far above the maximal performance that could be reached with simple heuristics and quite close to the optimal linear approximation. The good initial calibration might have been caused by the low context lotteries. A worse calibration might be expected for materials that have been used in previous studies demonstrating base-rate neglect. Note that our findings do not necessarily conflict with these findings. We would, however, predict that bad calibration and base-rate neglect will quickly disappear after a few feedback trials.[6]

Approximating Rationality based on Intuition

Our findings allow building bridges from psychology to (behavioral) economics. The results suggest a pragmatic solution to the homo oeconomicus/bounded rationality controversy between psychological decision researchers and economists (cf. Loewenstein, Rick, & Cohen, 2008; Smith, 2005). The former postulate that an individual behaves approximately rational and chooses the option that maximizes its utility; the later state that for many problems humans' cognitive capacity might not suffice to conduct the necessary complex calculations. We show that individuals are indeed able to approximate rationality on an individual level without having the

---

6   In a recent study we replicated and extended the findings (Glöckner & Dickert, in preparation) and indeed observed lower performance in the first trials using the cab-accident-problem instead of the lottery task. These differences, however, were surprisingly small and disappeared quickly.

ability for deliberately calculating the solution to the problem. Well calibrated intuition seems to make this possible. Hence, the rational choice assumption underlying economic modeling seems more psychologically plausible than usually assumed by decision researchers.

In economics and psychology, judgments and decisions are often categorized into rational ones that are based on deliberation and irrational ones that are based on intuition and/or heuristics. Our results indicate that this categorization should be rethought: deliberate and intuitive decision making might both lead to (approximately) rational behavior or deviations from it. It might be the core advantage of humans to be able to combine both kinds of processes to make good decisions (approximate rationality) (Glöckner, 2008). On the other hand, in line with the bulk of findings in psychological decision research (e.g., Kahneman et al., 1982) our results show that individuals are not perfectly rational. In many tasks, a considerable portion of individuals systematically selects the less likely option. The expected error can, however, be quantitatively predicted by the distance of the posterior probability from .5. Hence, behaviorally informed economic modeling could, instead of assuming completely rational choices, estimate choice proportions and error rates based on posterior probabilities using Bayes' theorem.

In summary, our results show that intuition enables individuals to quickly come to choices in complex probabilistic tasks for which they are not able to (deliberately) produce a normative numerical solution. Choices approximate rationality surprisingly well and choice proportions are predicted by the normative posterior probability according to Bayes' theorem. Response times, error rates and confidence judgments are compatible with the predictions of PCS models, which we suggest to be a promising approach for specifying intuitive information processing in complex probabilistic tasks.

References

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44*(3), 211-233.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences, 30*(3), 241-254.

Brase, G. L., Fiddick, L., & Harries, C. (2006). Participant recruitment methods and statistical reasoning performance. *The Quarterly Journal of Experimental Psychology, 59*, 965–976.

Brenner, L., Griffin, D., & Koehler, D. J. (2005). Modeling patterns of probability calibration with random support theory: Diagnosing case-based judgment. *Organizational Behavior and Human Decision Processes, 97*(1), 64-81.

Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 611-625.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432-459.

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

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On Making the Right Choice: The Deliberation-Without-Attention Effect. *Science, 311*(5763), 1005-1007.

Fiedler, K., Brinkmann, B., Betsch, T., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General, 129*(3), 399-418.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684-704.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Heuristics and biases: The psychology of intuitive judgment xvi, 857 pp New York, NY, US: Cambridge University Press.

Glöckner, A. (2008). How evolution outwits bounded rationality: The efficient interaction of automatic and deliberate processes in decision making and implications for institutions. In C. Engel & W. Singer (Eds.), *Better than conscious? Decision making, the human mind, and implications for institutions* (pp. 259-284). Cambridge, MA: MIT Press.

Glöckner, A., & Betsch, T. (2008a). Do people make decisions under risk based on ignorance? An empirical test of the Priority Heuristic against Cumulative Prospect Theory. *Organizational Behavior and Human Decision Processes, 107*(1), 75–95.

Glöckner, A., & Betsch, T. (2008b). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making, 3*(3), 215–228.

Glöckner, A., & Betsch, T. (2008c). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1055-1075.

Goodie, A. S., & Fantino, E. (1999). What does and does not alleviate base-rate neglect under direct experience. *Journal of Behavioral Decision Making, 12*(4), 307-335.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions on Systems, Man, & Cybernetics, 17*(5), 753-770.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128*(1), 3-31.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.

Kleiter, G. D., Krebs, M., Doherty, M. E., Garavan, H., Chadwick, R., & Brake, G. (1997). Do subjects understand base rates? *Organizational Behavior and Human Decision Processes, 72*(1), 25-61.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*(1), 1-53.

Loewenstein, G., Rick, S., & Cohen, J. D. (2008). Neuroeconomics. *Annual Review of Psychology. Vol 59 2008, 647-672.*

Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes, 68*(1), 28-43.

Lopes, L. L. (1994). Psychology and economics: Perspectives on risk, cooperation, and the marketplace. *Annual Review of Psychology. Vol 45 1994, 197-227.*

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534-552.

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science, 13*(2), 60-66.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1-75.

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell? In *Simple heuristics that make us smart* (pp. 141-167). New York, NY: Oxford University Press.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130*(3), 380-400.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of Preferences by Constraint Satisfaction. *Psychological Science, 15*(5), 331-336.

Simon, D., Snow, C. J., & Read, S. J. (2004). The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction. *Journal of Personality and Social Psychology, 86*(6), 814-837.

Slovic, P. (1987). Perception of risk. *Science, 236*(4799), 280-285.

Smith, V. L. (2005). Behavioral economics research and the foundations of economics. *The Journal of Socio-Economics, 34*(2), 135-150.

Tversky, A., & Kahneman, D. (1983). Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review, 90*(4), 293-315.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*(2), 181-192.