

# Now you Bayes, now you don't: effects of set-problem and frequency-format mental representations on statistical reasoning

Miroslav Sirota · Lenka Kostovičová ·  
Frédéric Vallée-Tourangeau

Published online: 25 February 2015  
© Psychonomic Society, Inc. 2015

**Abstract** People appear to be Bayesian when statistical information is presented in terms of natural frequencies and non-Bayesian when presented in terms of single-event probabilities, unless the probabilities resemble natural frequencies, for example, as chances. The isomorphic format of chances, however, does not always facilitate performance to the extent that the format of natural frequencies does. Prior research has not addressed the underlying mechanism that accounts for this gap despite its theoretical significance. The mechanism explaining this external format gap could lie in the interpretation of the problem as a set-problem, which cues relevant problem model and arithmetic operations (*the problem interpretation hypothesis*) and/or in the interpretation of the format as frequencies, which may be easier to process (*the format interpretation hypothesis*). In two parallel experiments, we found support for the problem interpretation hypothesis only: set representations mediated solely the isomorphic format gap (Experiment 1: part A) and accounted for the transfer effect to natural frequencies (Experiment 1: part B); priming set representations improved performance with chances (Experiment 2). We discuss how the supported explanation corroborates the nested-sets rather than the ecological rationality account of statistical reasoning and how it helps explain individual differences in Bayesian reasoning.

**Keywords** Bayesian reasoning · Chances · Natural frequencies · Problem mental representation · Format mental representation

---

M. Sirota (✉) · F. Vallée-Tourangeau  
Department of Psychology, Kingston University, Penrhyn Road,  
Kingston upon Thames, Surrey, KT1 2EE London, UK  
e-mail: miroslav.sirota@kingston.ac.uk

L. Kostovičová  
Institute of Experimental Psychology, Slovak Academy of Sciences,  
Bratislava, Slovakia

## Introduction

Presenting statistical information in the external format of natural frequencies improves Bayesian reasoning compared with single-event probabilities in traditional Bayesian textbook problems (e.g., Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). The nature of the cognitive mechanism underpinning such facilitation has kindled an intense and ongoing debate (e.g., Barbey & Sloman, 2007; Brase, 2008; Gigerenzer & Hoffrage, 1995; Girotto & Gonzalez, 2001; Lesage, Navarrete, & De Neys, 2013; Sirota, Kostovičová, & Vallée-Tourangeau, 2015; Sloman, Over, Slovak, & Stibel, 2003).

Natural frequencies—originally formulated by Kleiter (1994) who stressed their natural sampling structure, which makes the base-rate information redundant—offer a computationally simpler format than single-event probabilities. However, more controversially, whether or not computational simplicity is the only element contributing to the facilitative effect remains an open question. In fact, according to the *ecological rationality account*, the computational simplicity as well as the specificity of information inputs provided by naturally sampled frequencies (i.e., natural frequencies) together drive the facilitation (e.g., Brase, 2008; Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). The human mind adapted to process natural frequencies—since they reflect the way the information has been encountered naturally during evolution—and therefore reasons better with them than with any other formats. In contrast, according to the *nested-sets account*, the partitive structure of the problem leads to its nested-sets representation and enables a computationally simple solution of the problem (e.g., Girotto & Gonzalez, 2001; Sloman et al., 2003; Tversky & Kahneman 1983): Any format that reveals adequately the same partitive structure of the

statistical information as natural frequencies would cause facilitation.

To test these two accounts, Girotto and Gonzalez (2001, 2002) contrasted natural frequencies with an isomorphic format that does not feature frequencies, namely a chances format that expresses single-event probabilities with natural sampling structure (e.g., 10 chances out of 1,000 that a woman has cancer). The two formats yielded similar performance supporting the computational simplicity prediction of the nested-sets account (Girotto & Gonzalez, 2001). However, further evidence weakened this conclusion since the chances format did not facilitate Bayesian reasoning to the same extent as did natural frequencies (Brase, 2008); we term this discrepancy the format gap. Prior research failed to address the underlying mechanism of such a gap despite its theoretical significance. An explanation for the gap would likely unveil the additional facilitatory ingredient of natural frequencies.

#### Interpreting the format gap: the format and problem interpretation hypotheses

We propose two non-exclusive interpretations of the gap: the format and problem interpretation hypotheses. The first interpretation posits that the mental representation of the statistical format accounts for the format gap; hence, the *format interpretation hypothesis*. Some researchers asserted that chances with natural sampling are really just natural frequencies disguised as probabilities (e.g., Hoffrage, Gigerenzer, Krauss, & Martignon, 2002). The mental representation of chances as frequencies rather than as probabilities will facilitate reasoning because frequencies are easier to process, for example, they tap easily into a frequency coding mechanism (e.g., Brase, 2008). Therefore, any discrepancies in performance between the two formats should be explained through the representation of chances as frequencies.

The second interpretation posits that the mental representation of the problem type—for example, classified in terms of required mathematical operations—accounts for the format gap; hence, the *problem interpretation hypothesis*. Some have proposed that every mathematical word problem, such as a Bayesian textbook problem, activates corresponding information networks in memory that cue related mathematical operators to map out a path to solution (Kintsch, 1988; Kintsch & Greeno, 1985). Consequently, the superior performance of natural frequencies may occur because participants interpret the task as a problem to which set operations should be applied rather than as a problem featuring frequencies.

The two mechanisms proposed by the format interpretation and problem interpretation hypotheses reflect to some extent two theoretical accounts of Bayesian reasoning—the ecological rationality account and the nested-sets account—but, strictly speaking, are not necessarily mutually exclusive. A frequency-based mental representation can trigger a set

representation of the problem and therefore affect performance or vice versa. Only the joint investigation of the two mechanisms can differentiate between the two theoretical accounts of Bayesian reasoning. According to some proponents of the nested-sets account, the problem set representation unveils the nested-sets structure of a problem and triggers appropriate operations employed to solve the problem (e.g., Sloman et al., 2003). Frequencies may induce set operations, which lead to increased performance; however, it is the set representation and not the frequency format per se that spurs the facilitation.

In contrast, according to some proponents of the ecological rationality account, the information structure and question format affect Bayesian reasoning through their frequency representation. For example, Brase (2008) asserted that the effect of question format found in Girotto and Gonzalez (2001) could be explained “... by the degree to which [it] elicits a frequency representation” (Brase, 2008, p. 287). This implies that a frequency mental representation is needed for facilitation, above and beyond the appropriate set representation of the problem. Therefore, we hypothesized that unique support for the problem interpretation hypothesis corroborates some versions of the nested-sets account, whereas unique support for the format interpretation hypothesis corroborates some versions of the ecological rationality account.

#### The present experiments

In this paper, we report three tests of the *problem interpretation hypothesis* and the *format interpretation hypothesis* in two parallel experiments. The first part of Experiment 1 (part A) tested both hypotheses by investigating the mediation pathways via the problem and format mental representations of the effect of statistical, external, format on Bayesian performance. The second part of Experiment 1 (part B) tested both hypotheses by investigating the mediation pathways via the problem and format mental representations of the transfer effect on Bayesian reasoning. That is, a transfer effect from the previous Bayesian problem featuring either chances or natural frequencies to a new (medical) problem featuring natural frequencies. Experiment 2 tested directly the problem representation processes: it tested the causal effect of the availability of set representations—induced by priming set or normalization operations—on Bayesian performance. It also tested the independence of these processes from those proposed in the format interpretation hypothesis by measuring the effect of priming on the format of mental representations.

The problem interpretation hypothesis would be fully supported only if the problem representations mediate the format gap and subsequent transfer effect, and affect performance by priming set-problem representations. The format interpretation hypothesis would be fully supported only if the format mental representations mediate the format gap and subsequent

transfer effect, and if priming affects performance via the changes in the format mental representations.

**Method**

**Participants and design**

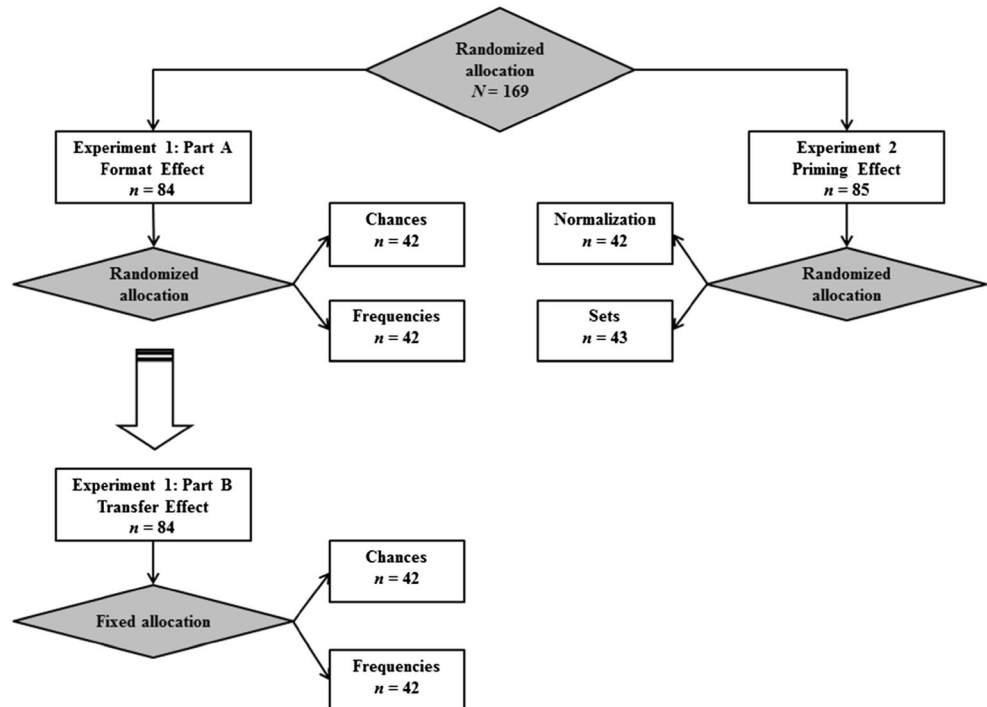
We aimed to include at least 82 participants in each of the two experiments, since we powered them to detect at least a medium size effect,  $w = 0.4$  (a conservative estimate given the prior literature, e.g., Brase, 2008), when assuming  $\alpha = .05$ ,  $1 - \beta = .95$ , and a two-sided test. Based on such power calculation, 169 social science undergraduates (45 males, age range 18–28 years,  $M = 20.6$ ,  $SD = 1.4$ ) from a national-level university in Slovakia participated in two parallel-design experiments (Fig. 1). In the first part of Experiment 1, designed to measure a format effect (part A), we allocated randomly the first (randomly selected) half of the participants ( $n = 84$ ) to one of two format conditions featuring a Bayesian problem either in natural frequencies or chances. In the second part of Experiment 1, designed to assess a transfer effect (part B), the same participants ( $n = 84$ ) answered a different (medical) Bayesian problem expressed in natural frequencies. In Experiment 2, designed to measure a priming effect, we allocated randomly the remaining participants ( $n = 85$ ) to one of two priming conditions: priming normalization or priming sets. This priming manipulation prompted participants to perform either normalization- or set-related

operations on structurally unrelated tasks *before* completing the critical Bayesian problem featuring chances.

**Materials and procedure**

Participants in the first part of Experiment 1 (format effect) solved a Bayesian problem expressed either in chances or natural frequencies (namely, the “Admission” problem, as used in Brase, 2008; adopted from Girotto & Gonzalez, 2001; see both versions in Table 1). We coded the responses numerically corresponding to the Bayesian solution (i.e., 3 out of 22 chances/applicants) as Bayesian and the rest as non-Bayesian answers. Upon completion, participants assessed the ways they thought about the problem and format (i.e., the problem and format representations). The problem representation was measured by the operations reported by participants on a 7-point Likert scale [1: “I considered the solution to the task as one involving a transformation into percentages, probabilities or other normalizations (such as ‘to hundreds’)”; 4: “Both equally”; 7: “I considered the solution to the task as one involving sets, subsets and relations between them”]. The format representation was measured by the nature of the statistical information participants attributed to the format of the task on a 7-point Likert scale (1: “The information in the task has the character of frequencies”; 4: “Both equally”; 7: “The information in the task has the character of probabilities”). Finally, participants answered some socio-demographic questions.

**Fig. 1** Flowchart of the design and sample size of the two parallel experiments



**Table 1** Bayesian problems used in the two experiments

“Admission” problem in natural frequencies (Brase, 2008, p. 284)

The applicants for admission to a prestigious university have to pass an entrance examination that involves both an oral test and a written test. Here is some information about the results of last year’s examination

5 out of 100 applicants were accepted. 3 of the 5 applicants who were accepted passed the oral test. However, 19 of the 95 applicants who were rejected also passed the oral test.

Imagine 100 applicants are taking the entrance examination. Out of the \_\_\_\_\_ applicant(s) that will pass the oral test, \_\_\_\_\_ applicant(s) will be accepted

“Admission” problem in chances (Brase, 2008, p. 285)

The applicants for admission to a prestigious university have to pass an entrance examination that involves both an oral test and a written test. Here is some information about the results of last year’s examination

An applicant had 5 chances out of 100 of being accepted. 3 of the 5 chances of being accepted were associated with passing the oral test. However, 19 of the remaining 95 chances of being rejected were also associated with passing the oral test.

Imagine Jean is an applicant taking the entrance examination. Out of the \_\_\_\_\_ chance(s) that Jean will pass the oral test, there are \_\_\_\_\_ chance(s) she will be accepted

“German measles” in natural frequencies\* (adapted from Gigerenzer & Hoffrage, 1995)

Based on the results of medical studies we know the following:

100 out of 20,000 children are born with severe prenatal damage, and thus 19,900 of the 20,000 are born without severe prenatal damage. 40 of the 100 children with severe prenatal damage are born to mothers who had German measles during early pregnancy.

However, 199 of the 19,900 children without severe prenatal damage are born to mothers who also had German measles during early pregnancy.

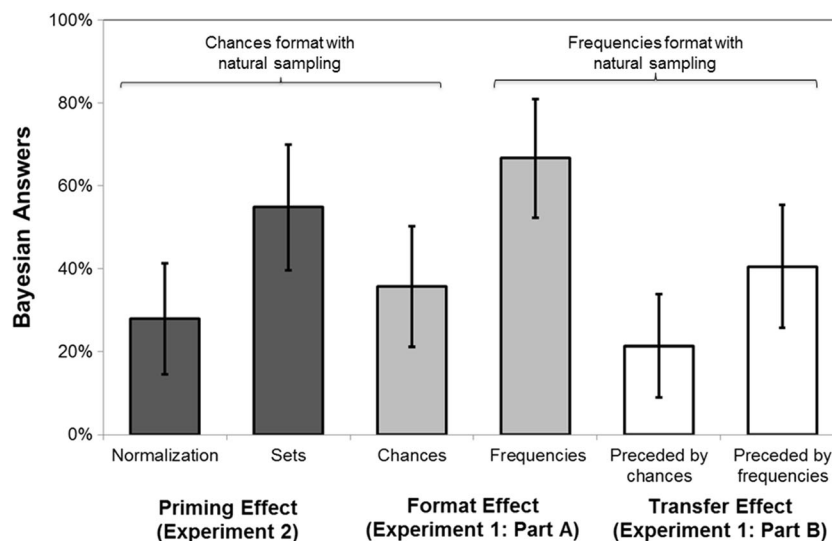
Imagine a sample of 20,000 future mothers with the diagnosis of German measles during early pregnancy. How many of these mothers will give a birth to a child with severe prenatal damage? \_\_\_ out of \_\_\_

\*Note. Gigerenzer and Hoffrage (1995) presented only a single-event probability version of the “German measles” problem in their seminal paper. Therefore, we have translated the probabilities into natural frequencies; the numbers respected the rule that all nodes consist of integers. We have also shortened the natural frequency version of the problem presented here by omitting the explanatory introduction

Unexpectedly, after completing the socio-demographic questions, we asked participants in the second part of Experiment 1 (transfer effect), to solve a different Bayesian problem (the “German measles” problem in natural frequencies, as used in Gigerenzer & Hoffrage, 1995; see Table 1) in order to measure transfer. We coded the responses numerically

corresponding to the Bayesian solution (i.e., 40 out of 239 mothers) as Bayesian and the rest as non-Bayesian answers.

Participants in Experiment 2 (priming effect) first solved a set of simple mathematical problems (15 tasks, see Table 2) structurally dissimilar to the Bayesian problem (the “Admission” problem in chances), which they solved afterwards.



**Fig. 2** The effect of problem priming (Experiment 2), external formats (Experiment 1: part A), and problem transfer (Experiment 1: part B) on Bayesian performance in problems with an equal information structure, namely chances sampling with natural sampling (*left half*) and frequencies with natural sampling (*right half*). Note that we used the “Admission”

problem to measure Bayesian performance in the priming and external format effects and the “German measles” problem to measure Bayesian performance in the transfer effect experiments. *Error bars* represent 95 % confidence intervals

**Table 2** The priming of two distinct mathematical types of problems: normalization and sets<sup>a</sup>

Priming normalization	Priming sets
<p>1. Given these probabilities: a) 8.10 %, b) 8.15 %, c) 13.3 %, d) 7 out of 47, e) 8.03 % <i>Answer the following:</i> 1.1 Which of the given probabilities is the highest? _____ 1.2 Which of the given probabilities is the second highest? _____</p> <p>2. Given these frequencies: a) 26.17 %, b) 26.09 %, c) 25.80 %, d) 12 out of 50, e) 26.7 % <i>Answer the following:</i> 2.1 Which of the outcomes is the most frequent? _____ 2.2 Which of the outcomes is the least frequent? _____</p> <p>3. Given these probabilities: a) 0.008, b) 0.13, c) 0.11, d) 6 out of 50, e) 0.03 <i>Answer the following:</i> 3.1 Which of the given probabilities is the highest? _____ 3.2 Which of the given probabilities is the second highest? _____</p> <p>4. Given these frequencies: a) 4 out of 100, b) 1 out of 100, c) 3 %, d) 2 out of 100, e) 3 out of 100 <i>Answer the following:</i> 4.1 Which of the outcomes is the most frequent? _____ 4.2 Which of the outcomes is the least frequent? _____</p> <p>5. <i>Compute the following:</i> 5.1) 4 out of 40 = 0. _____ 5.2) 18/36 = _____ out of 100 5.3) 35 out of 100 = _____ % 5.4) 60/120 = 0. _____ 5.5) 8 out of 40 = _____ out of 100 5.6) 5 out of 1000 = _____ % 5.7) 3/100 = _____ %</p>	<p>1. Sets A and B are given, where <math>A = \{1, 2, 3, 4, 5, 6, 7\}</math>, <math>B = \{2, 4, 6, 8, 10, 12\}</math> <i>Answer the following:</i> 1.1 State the elements of set A and B: _____ 1.2 State the elements of set A, which are included in set B: _____ 1.3 State the elements of set B, which are not included in set A: _____</p> <p>2. Sets A, B and C are given, where <math>A = \{a, b, c, d\}</math>, <math>B = \{a, b, c, d, e, f\}</math>, <math>C = \{a, b, c, d, g, h\}</math> <i>Answer the following:</i> 2.1 Is set A a subset of set B? _____ 2.2 What is the intersection of sets B and C? _____ 2.3 State the elements of set C, which are not included in set B: _____</p> <p>3. Sets I, J and K are given. If set I is a subset of set J and set K is a subset of set I, which of the following is necessarily true statement? <i>Circle the correct answers.</i> 3.1 All elements of set I are elements of set J. a) yes b) no 3.2 Some elements of set J are maybe not involved in sets I or K. a) yes b) no 3.3 Sets I and K are union sets. a) yes b) no 3.4 Set J includes elements of sets I and K. a) yes b) no</p> <p>4. Given sets A, B and C, where <math>A = \{1, 2, 3, 4, 5, 6, 7, 8\}</math>, <math>B = \{2, 3, 4, 5\}</math>, <math>C = \{4, 5, 6, 7, 8\}</math> <i>Answer the following:</i> 4.1 Are sets B and C subsets of set A? _____ 4.2 State the elements, which are the intersection of sets B and C: _____ 4.3 State the elements of set C, which are not included in set B: _____ 4.4 State the elements of set A, which are not included in set B: _____</p> <p>5. State the elements, which are the intersection of sets G and H, where G is a set of all odd natural numbers smaller than 10 and <math>H = \{-1, 1, 3, 5, 7\}</math>: _____</p>

<sup>a</sup> In the instructions, participants were asked to solve the problems as best they could. They were not allowed to use calculators and there was no time limit

Upon completion, participants answered questions on the operations employed using the same Likert scales as in Experiment 1; the same socio-demographic data were collected. Overall, participants were very successful in solving the set of mathematical problems ( $Mdn = 14$ ,  $IQR = 3.5$ ). They were statistically less successful,  $M-W U = 341$ ,  $P < .001$ ,  $r_m = .55$ , in the priming normalization tasks ( $Mdn = 12$ ,  $IQR = 4$ ) compared with the priming sets tasks ( $Mdn = 14$ ,  $IQR = 1$ ). Despite this difference, we think that the activation of the corresponding operations was similarly strong in terms of affecting Bayesian performance. This is mainly because the majority of the participants solved most of the priming tasks in both conditions. It is reasonable to believe that corresponding arithmetic concepts were activated effectively by solving a couple of tasks only. In addition, if higher priming tasks accuracy rates lead to stronger activations then the Bayesian

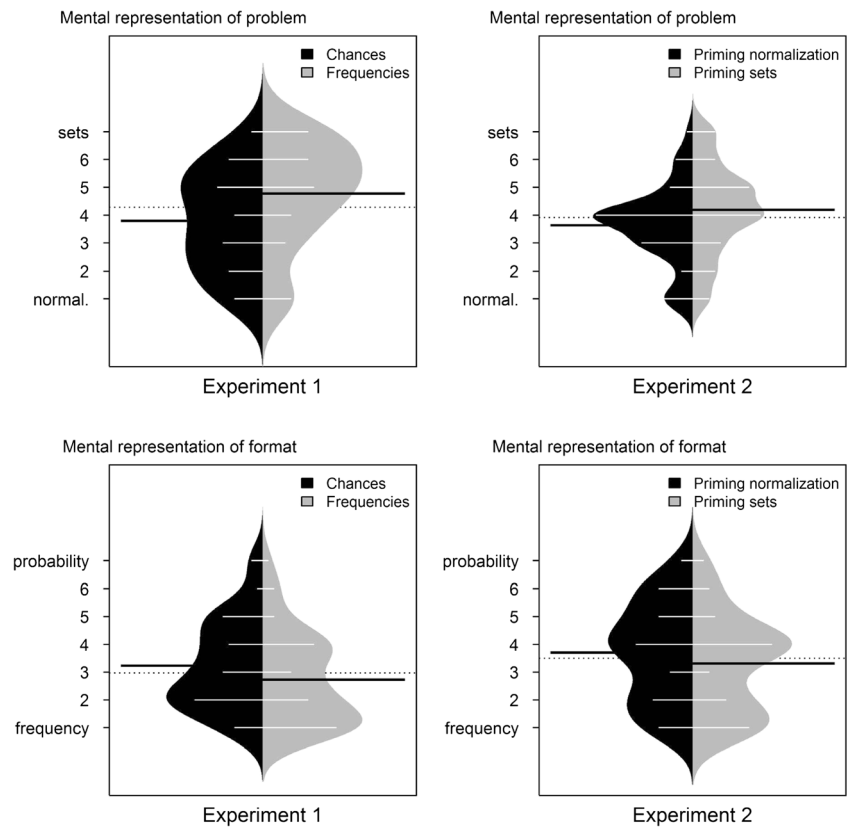
performance should drop down proportionally in the priming normalization condition and increase in the priming sets condition as a function of the accuracy rates; this was not the case (respectively,  $r_{pb} = .022$ ,  $P = .889$ ;  $r_{pb} = .092$ ,  $P = .561$ ).

## Results

Experiment 1 (part A): mediation of the format effect on Bayesian performance

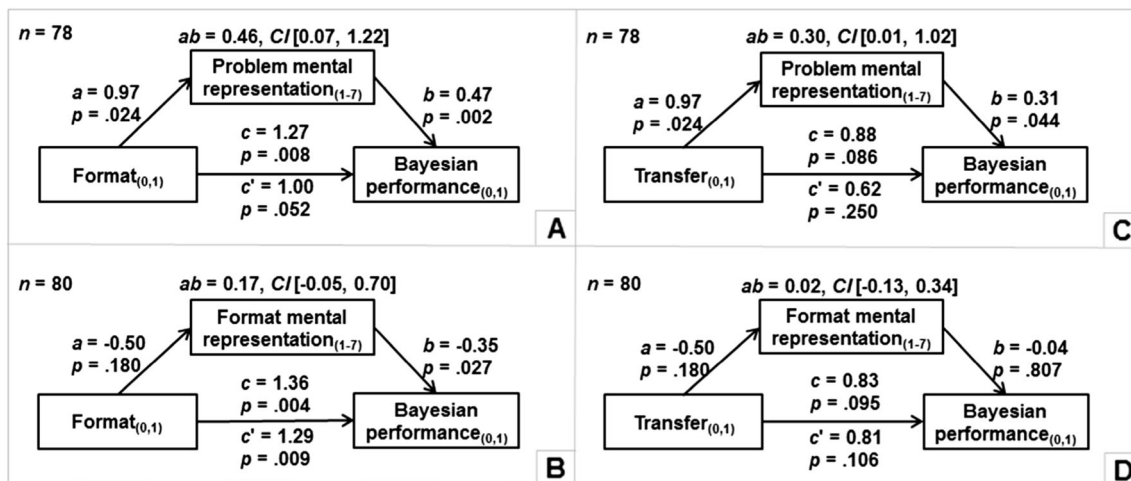
We observed an external format gap: participants' answers were more frequently Bayesian when the problem format involved natural frequencies than when it involved chances,  $\chi^2(1, N = 84) = 8.05$ ,  $P = .005$ ,  $\phi = .31$  (Fig. 2, middle panel). The external format manipulation—chances vs natural

**Fig. 3** Effect of manipulations in Experiment 1 (chances vs frequencies) and in Experiment 2 (priming normalization vs priming sets) on mental representation of problem and format (*asymmetric bean plot*). *Bold lines* Overall mean value of the scale per group, *polygons* density shapes determined by a normal density trace computed with a default bandwidth method using the Sheather-Jones method to select a bandwidth per batch and then averaging bandwidths over all batches (see Kampstra, 2008)



frequencies—affected the mental representations of the problem but not the mental representations of the format: natural frequencies led to more set representations, but not to more

frequency representations, compared with chances (see Fig. 3, Experiment 1). We conducted two simple mediation analyses to test the mediation of the external format gap via the



**Fig. 4** Mediation analyses of the external format (a, b Experiment 1: part A) and transfer effects (c, d Experiment 1: part B) on Bayesian reasoning (measured as a dichotomous variable) via mental problem representations ('Problem mental representation'; a, c) and mental format representations ('Format mental representation'; b, d). The mediation models consist of the dichotomous variable of format or transfer effect manipulation as the independent variable, the 7-point Likert scale variable of problem or format mental representation as the mediator, and the dichotomous variable of Bayesian performance

as the dependent variable. Thus, the path coefficients represent either linear or logistic regression coefficients. The reported confidence intervals represent 95 % bootstrapped confidence intervals. The numbers differ slightly due to some missing values in the mental representation questions. *a* path coefficient from independent variable to mediator, *b* path coefficient from mediator to dependent variable, *ab* path coefficient from independent variable to dependent variable via mediator (indirect effect), *c'* path coefficient from independent variable to dependent variable (direct effect), *c* path coefficient from independent variable to dependent variable (total effect)

different type of mental representations. The mental problem representations mediated the effect of external formats on performance (the indirect effect was significant and the direct effect of format on performance was not), whereas the mental format representations yielded a non-significant indirect effect of external formats on performance (see Fig. 4a,b).

These findings support the problem interpretation hypothesis rather than the format interpretation hypothesis because set-problem representations explained the increased performance with natural frequencies regardless of the frequency interpretation of the information. An additional test of these two hypotheses would be to scrutinize the mediation pathways via the problem and format mental representations in the performance transferred to a new problem formulated in terms of natural frequencies. Such a test follows from the argument that any natural-frequencies-training effects on Bayesian reasoning are linked to the mental problem re-structuring rather than the mental format re-structuring (e.g., Sirota et al., 2015). We tested the underlying mediation pattern of a transfer effect of working through the “Admission” problem featuring chances or natural frequencies (from the part A of Experiment 1) on a medical problem featuring natural frequencies (the “German measles” problem).

#### Experiment 1 (part B): mediation of the transfer effect on Bayesian performance with natural frequencies

Performance with natural frequencies or chances in a preceding “Admission” problem transferred to a great extent to the performance with natural frequencies in a medical problem. The transfer effect resulted in substantially higher performance if the preceding problem featured natural frequencies compared with chances, although the difference did not reach statistical significance,  $\chi^2(1, N = 84) = 3.57, P = .059, \phi = .21$  (Fig. 2, right panel). The transfer effect is less surprising, given the format gap found in the previous problem; its underlying mediation pattern, on the other hand, is surprising and informative. In general, the mediation patterns can be informative even in the situation of a non-significant total effect, because they indicate potentially interesting underlying mechanisms (e.g., Hayes, 2009). The problem—not format—representations (i.e., set-representations) employed in the preceding problem accounted for the performance discrepancy in the current problem: the significant indirect effect via problem representations accounted for substantial variance of the total effect, and the non-significant indirect effect via format representations accounted for only marginal variance of the total effect (see Fig 4c,d).

In summary, representing a problem as a set-problem, regardless of its frequency representation, accounted for the performance discrepancy in natural frequencies transferred from the previous task; such findings support the problem rather than the format interpretation hypothesis. However, the

supporting evidence remains only correlational, since we did not manipulate the mediator (see Spencer, Zanna, & Fong, 2005). In Experiment 2, we planned to overcome this shortcoming and provide supportive causal evidence by manipulating the problem representations through priming either normalization-related or set-related representations.

#### Experiment 2: priming effect on Bayesian performance with chances

The priming affected the mental representation of the problem type, but not that of the format. Priming sets yielded significantly more set than normalization representations of the Bayesian problem,  $M-W U = 673, P = .038$  (1-tailed exact significance test),  $r_m = .19$ , but did not trigger substantially more frequency representations,  $M-W U = 673, P = .148$  (1-tailed exact significance test),  $r_m = .12$  (see Fig. 3, Experiment 2). This effective manipulation check suggests that any effects on performance occurred likely due to the changes in the problem representation.

In line with our expectations, priming sets caused a significantly higher proportion of Bayesian answers compared with normalization priming,  $\chi^2(1) = 6.33, P = .012, \phi = .27$  (see Fig. 2, left panel). The effect of priming manipulation mirrored the performance of chances and natural frequencies in Experiment 1. Indeed, the performance in the set priming condition was higher than in the control chances condition although not statistically significantly so,  $\chi^2(1) = 3.08, P = .079, \phi = .19$ , and very similar to the performance in the natural frequencies condition,  $\chi^2(1) = 1.25, P = .264, \phi = .12$ . In contrast, performance with normalization priming resembled remarkably the performance in the control chances condition,  $\chi^2(1) = 0.60, P = .440, \phi = .08$ , being significantly lower than the performance in the natural frequencies condition,  $\chi^2(1) = 12.81, P < .001, \phi = .39$ . Priming set representations boosted performance, regardless of the frequency representations, further supporting the problem rather than the format interpretation hypothesis.

### General discussion

The three empirical tests reported here evaluated two hypotheses proposing distinct psychological mechanisms for the facilitative effect occurring with Bayesian problems featuring partitive formats. According to the problem interpretation hypothesis, a proper problem mental representation (i.e., in terms of sets) activates appropriate knowledge, problem schemas and corresponding mathematical operations and, therefore, improves Bayesian reasoning. According to the format interpretation hypothesis, a proper format mental representation (i.e., in terms of frequencies) activates a frequency-coding mechanism and, in turn, improves Bayesian reasoning. We

found here strong support for the problem interpretation hypothesis only: the set, not frequency representations, mediated the format gap (Experiment 1: part A); the set, not frequency representations, accounted for the performance with natural frequencies transferred from a preceding task (Experiment 1: part B); priming set representations, regardless of the format representations, boosted Bayesian performance with chances (Experiment 2).

The exclusive support for the problem interpretation hypothesis corroborates the nested-sets (Giroto & Gonzalez, 2001; Sloman et al., 2003) rather than the ecological rationality account (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Indeed, if set mental representations of Bayesian problems boost performance with chances *independently* of their interpretation in terms of frequencies, then such interpreted representations cannot be an input for, or output of, a frequency-coding mechanism wrought by evolutionary forces as suggested by the ecological rationality account. Rather, the set representations make the nested-sets structure of the problem more visible as suggested by the nested-sets account. The robustness of the effects reported here suggests that a general rather than a specific cognitive mechanism is responsible for the interpretation and integration of the task elements (e.g., format, explicit verbal triggers) and the knowledge associated with them (e.g., schemas, problem-typical mathematical operations). Indeed, the specific role of natural frequencies in the facilitation disappears when the comparative task features an identical partitive structure accompanied with unambiguous cues leading to similar problem representations (e.g., set operations; see Sirota, Juanchich, & Hagmayer, 2014). The existence of a general cognitive mechanism is further supported by other evidence (e.g., Lesage et al., 2013; Sirota et al., 2014a; Sirota, Kostovičová, & Juanchich, 2014) and conceptual arguments (e.g., the argument of isomorphism between set-subsets relationships in natural sampling structures and some more fundamental logical operations as developed in Over, 2007).

Despite the demonstrated robustness of the problem representation effect, future research should overcome some of the methodological limitations of the experiments reported here. For example, researchers should develop more elaborated mental representation measures and assess the content of mental representations concurrently rather than retrospectively. Given the reported data patterns, however, these limitations likely played a minimal role and constrain our theoretical conclusions only marginally. Future research should also test the problem interpretation explanation for the other findings reported in the literature on Bayesian reasoning. Indeed, developing the correct problem representation may require deeper, more extensive and more active networks of pertinent numerical-problems knowledge, which would explain the positive link between numeracy and Bayesian reasoning (e.g., Chapman & Liu, 2009; Sirota & Juanchich, 2011) and

the positive link between numeracy and causal problem structure facilitation (McNair & Feeny, 2015). Such a correct problem representation may also require better inhibitory control of inappropriate problem-type interpretations, which could explain the positive link between cognitive reflection ability and Bayesian reasoning (e.g., Lesage et al., 2013; Sirota & Juanchich, 2011). Finally, developing the correct problem representation may require sufficient cognitive capacity to build the adequate problem model, which could explain the effect of cognitive load on Bayesian reasoning and its link with general intelligence (e.g., Lesage et al., 2013; Sirota et al., 2014a).

## Conclusion

In two experiments, we found exclusive support for the problem interpretation hypothesis, since only a problem set representation accounted for the gap between two isomorphic formats. This hypothesis offers a novel explanation of why some partitive formats encourage more normative reasoning than others and supports the nested-sets rather than the ecological rationality account of Bayesian reasoning.

## References

- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: from ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*, 241–297. doi:10.1017/S0140525X07001653
- Brase, G. L. (2008). Frequency interpretation of ambiguous statistical information facilitates Bayesian reasoning. *Psychonomic Bulletin & Review*, *15*, 284–289. doi:10.3758/PBR.15.2.284
- Chapman, G. B., & Liu, J. (2009). Numeracy, frequency, and Bayesian reasoning. *Judgment and Decision Making*, *4*, 34–40.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgement under uncertainty. *Cognition*, *58*, 1–73. doi:10.1016/0010-0277(95)00664-8
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, *102*, 684–704. doi:10.1037/0033-295X.102.4.684
- Giroto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: a matter of information structure and question form. *Cognition*, *78*, 247–276. doi:10.1016/S0010-0277(00)00133-5
- Giroto, V., & Gonzalez, M. (2002). Chances and frequencies in probabilistic reasoning: rejoinder to Hoffrage, Gigerenzer, Krauss, and Martignon. *Cognition*, *84*, 353–359. doi:10.1016/S0010-0277(02)00051-3
- Hayes, A. F. (2009). Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Communication Monographs*, *76*, 408–420. doi:10.1080/03637750903310360
- Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: what natural frequencies are and what they are not. *Cognition*, *84*, 343–352. doi:10.1016/S0010-0277(02)00050-1



- Kampstra, P. (2008). Beanplot: a boxplot alternative for visual comparison of distributions. *Journal of Statistical Software*, 28, 1–9.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, 95, 163–182. doi:10.1037/0033-295X.95.2.163
- Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, 92, 109–129. doi:10.1037/0033-295X.92.1.109
- Kleiter, G. D. (1994). Natural sampling: rationality without base rates. In G. Fischer & D. Laming (Eds.), *Contributions to mathematical psychology, psychometrics, and methodology* (pp. 375–388). New York: Springer.
- Lesage, E., Navarrete, G., & De Neys, W. (2013). Evolutionary modules and Bayesian facilitation: the role of general cognitive resources. *Thinking & Reasoning*, 19, 27–53. doi:10.1080/13546783.2012.713177
- McNair, S., & Feeney, A. (2015). Whose statistical reasoning is facilitated by a causal structure intervention? *Psychonomic Bulletin & Review*, 22, 258–264. doi:10.3758/s13423-014-0645-y
- Over, D. (2007). The logic of natural sampling. *Behavioral and Brain Sciences*, 30, 277. doi:10.1017/S0140525X07001859
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–891. doi:10.3758/BRM.40.3.879
- Sirota, M., & Juanchich, M. (2011). Role of numeracy and cognitive reflection in Bayesian reasoning with natural frequencies. *Studia Psychologica*, 53, 151–161.
- Sirota, M., Juanchich, M., & Hagemayer, Y. (2014a). Ecological rationality or nested sets? Individual differences in cognitive processing predict Bayesian reasoning. *Psychonomic Bulletin & Review*, 21, 198–204. doi:10.3758/s13423-013-0464-6
- Sirota, M., Kostovičová, L., & Juanchich, M. (2014b). The effect of iconicity of visual displays on statistical reasoning: evidence in favor of the null hypothesis. *Psychonomic Bulletin & Review*, 21, 961–968. doi:10.3758/s13423-013-0555-4
- Sirota, M., Kostovičová, L., & Vallée-Tourangeau, F. (2015). How to train your Bayesian: a problem-representation transfer rather than a format-representation shift explains training effects. *Quarterly Journal of Experimental Psychology*, 68, 1–9. doi:10.1080/17470218.2014.972420
- Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, 91, 296–309. doi:10.1016/S0749-5978(03)00021-9
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, 89, 845–851. doi:10.1037/0022-3514.89.6.845
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315. doi:10.1037/0033-295X.90.4.293