

Open Peer Commentary

A statistical taxonomy and another “chance” for natural frequencies

DOI: 10.1017/S0140525X07001665

Adrien Barton,^{a,b} Shabnam Mousavi,^{a,c} and Jeffrey R. Stevens^a

^aCenter for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, 14195 Berlin, Germany; ^bInstitut d’Histoire et de Philosophie des Sciences et des Techniques, Paris I, and CNRS/ENS – UMR 8590, 75006 Paris, France; ^cDepartment of Statistics, The Pennsylvania State University, University Park, PA 16802.

barton@mpib-berlin.mpg.de <http://www.mpib-berlin.mpg.de/en/forschung/abc/index.htm> mousavi@mpib-berlin.mpg.de
 sxm70@psu.edu <http://www.stat.psu.edu/people/faculty/smousavi.html/jstevens@mpib-berlin.mpg.de>
<http://www.abc.mpib-berlin.mpg.de/users/jstevens/>

Abstract: The conclusions of Barbey & Sloman (B&S) crucially depend on evidence for different representations of statistical information. Unfortunately, a muddled distinction made among these representations calls into question the authors’ conclusions. We clarify some notions of statistical representations which are often confused in the literature. These clarifications, combined with new empirical evidence, do not support a dual-process model of judgment.

We disagree with Barbey & Sloman’s (B&S’s) claim that data on Bayesian reasoning support their dual-process model of human judgment. First, we clarify the dimensions along which statistical information can be expressed, and then point to how this common conceptual confusion can influence B&S’s interpretation of existing data. Second, we explain how new evidence contradicts this model.

1. Statistical representation. Statistical information can be represented in multiple ways along two orthogonal dimensions:

the number of events and the numerical format. First, the information may concern only one event (single-event probability) or a set of events (frequency). For example, the probability that a person has a positive test if she is ill is a single-event probability, in contrast to the frequency of people having a positive test among those who are ill. Second, the numerical information can be represented as percentages (20%); fractions (20/100); real numbers between 0 and 1 (0.2); or pairs of integers (“20 chances out of 100” for single-event probabilities, and “20 people out of 100” for frequencies).

Consider now a Bayesian task of computing the probability of a hypothesis H , given the data D , such as the probability of being ill, given the result of a test. In this context, there is yet another orthogonal dimension along which the statistical information can vary: the information can be expressed in a *conjunctive* or in a *normalized* format. The conjunctive format gives the relevant conjunctive information $P(H \& D)$ and $P(not-H \& D)$, or $P(H \& \bar{D})$ and $P(D)$. In this case the Bayesian computations are rather simple (see Eq. 1):

$$P(H | D) = \frac{P(H \& D)}{P(H \& D) + P(not-H \& D)} = \frac{P(H \& D)}{P(D)} \quad (1)$$

Alternatively, information can be expressed in a normalized format giving the normalized information $P(D | H)$ and $P(D | not-H)$, in addition to $P(H)$ – and not giving the relevant conjunctive information. The normalized format complicates computing the Bayesian results (see Eq. 2):

$$P(H | D) = \frac{P(D | H)P(H)}{P(D | H)P(H) + P(D | not-H)(1 - P(H))} \quad (2)$$

Because the number of events and conjunctive/normalized format dimensions are orthogonal, one can give statistical information in a Bayesian task in four possible ways, each of which can be represented as percentages, fractions, real numbers or pair of integers (see Table 1 for examples of Bayesian tasks, each represented in different numerical formats).

Confusion among these three orthogonal dimensions is common in the literature and poses particular problems in the

Table 1 (Barton et al.). *Taxonomy of statistical information: examples of Bayesian tasks*

	Normalized format	Conjunctive format
Single event probabilities	A 40-year-old woman who participates in routine screening has 10 out of 1,000 chances to have breast cancer. [$P(H)$]	The probability of breast cancer is 1% for a 40-year-old woman who participates in a routine screening. [$P(H)$]
	If such a woman has breast cancer, she has 800 out of 1,000 chances to have a positive mammography [$P(D H)$]	The probability that such a woman has a positive mammography and has breast cancer is 0.8% [$P(H \& D)$]
	If such a woman does not have breast cancer, she has 96 out of 1,000 chances to have a positive mammography. [$P(D not-H)$]	The probability that she has a positive mammography and does not have breast cancer is 9.5%. [$P(not-H \& D)$]
Frequencies	A proportion of 0.01 of women at age 40 who participate in routine screening have breast cancer. [$P(H)$]	(labeled <i>Natural frequencies</i> when represented as pairs of integers) 10 out of 1,000 women at age forty who participate in routine screening have breast cancer [$P(H)$]
	A proportion 0.8 of women with breast cancer will have positive mammographies. [$P(D H)$]	8 out of these 1,000 women have a positive mammography and have breast cancer [$P(H \& D)$]
	A proportion 0.096 of women without breast cancer will also have positive mammographies. [$P(D not-H)$]	95 out of these 1,000 women have a positive mammography and do not have breast cancer [$P(not-H \& D)$]

Note. Each of the four numerical formats can apply to the four probability/frequency and normalized/conjunctive combinations. Here, we arbitrarily assigned numerical formats for each cell.

B&S target article, because the authors draw false conclusions on this basis. First, they mention “natural frequency formats that were not partitioned into nested set relations” (B&S, sect. 2.4, para. 2). But non-partitioned frequency formats are simply frequencies expressed in a normalized format; therefore, natural frequencies *must* be partitioned into nested set relations. Consequently, B&S’s critiques of the so-called non-partitioned natural frequencies apply only to normalized frequencies (i.e., frequencies in a normalized format). Second, a sentence like “33% of all Americans will have been exposed to Flu” (target article, Table 5) concerns a whole population, not a single individual; therefore, it conveys frequencies, not single-event probabilities, contrary to the authors’ equating of percentages (referring to the numerical representation) and single-event probabilities (referring to the number of events). B&S misinterpret a key result of Brase (2002b), who showed that subjects perceived simple frequencies (i.e., frequencies represented as pairs of integers) as clearer, more understandable, and more impressive than single-event probabilities represented as non-integer numbers (Brase 2002b, pp. 388–89).

B&S contrast what they call “theoretical frameworks” (sect. 1.2, para. 2) based on natural frequency representations with the “nested set hypothesis” (sect. 2.10, para. 5). However, some of these contrasts appear a bit artificial. Consider the Gigerenzer and Hoffrage (1995) study, which predicted and showed that natural frequencies facilitate Bayesian inference. B&S claim this effect results from the clarification of the nested-sets structure of the problem. But Gigerenzer and Hoffrage (1995) had already made a more specific, related argument, stating that the facilitation of natural frequencies results from simplifying the Bayesian computations by giving the relevant conjunctive information. B&S’s idea that this facilitation results from the clarification of the nested-sets structure of the problem stands more in opposition to the use of an evolutionary argument (Cosmides & Tooby 1996) to predict the facilitating effect of natural frequencies, than in opposition to Gigerenzer and Hoffrage’s argument. However, some of the arguments used by B&S against this evolutionary stance are not valid. For example, they say that since both normalized and natural “frequencies exist in nature, it is unclear why only frequencies of the latter type are deemed important” (target article, Note 5). This is not true: Natural frequencies “exist in nature” in the sense that, in a natural sample of the population, counting the number of individuals belonging to the groups $H \cup D$ and $\text{not-}H \cup D$ yields natural frequencies (conjunctive information); whereas, counting in such a natural sample cannot result in normalized information that does not contain the conjunctive information. Therefore, normalized frequencies, which only give the normalized information, do not exist in nature in this sense. So, an evolutionary argument that predicts the facilitating effect of natural frequencies, but no such effect for normalized frequencies, can be defended against this charge that both normalized and natural frequencies exist in nature (even if one sees this evolutionary argument as speculative and feels uncomfortable about making precise predictions on this basis).

2. Facilitating effects. B&S propose a general dual-process model of judgment, which denies any facilitating effect of frequencies per se, because “facilitation is a product of general-purpose reasoning processes” (sect. 1.2.5, para. 1). As evidence against such an effect, B&S cite Girotto and Gonzalez’s (2001) facilitating effect when the information is given as “number of chances” in a conjunctive format, which is a way of expressing single-event probabilities. Recent work by Brase (2007), however, demonstrates that many people interpret such chances as natural frequencies, despite instructions to the contrary. Moreover, those who interpret chances as natural frequencies have higher rates of success than those who judge the information as single-event probabilities. This suggests that frequencies can have a facilitating effect in some circumstances,

in addition to the facilitating effect of computational simplicity. If this facilitating effect of frequencies is confirmed, it would make the dual-process model much more difficult to defend.

Conclusion. Though we do not dismiss the idea of a dual-process model outright, we think that B&S have not made a robust argument in support of such a model. The authors misinterpret data used to reject alternatives to the nested-set hypothesis. Further, the connection between the nested-set hypothesis and the dual-process model of judgment is not as crisp as one would like. Perhaps this is due to the rather vague nature of the dual-process model itself (cf. the criticisms of Gigerenzer & Regier 1996). The general project of building such a model sounds exciting, but we look forward to a more rigorous, clearly defined (and therefore falsifiable) dual-process model of judgment.

ACKNOWLEDGMENTS

We thank Nils Straubinger and Gerd Gigerenzer for their comments on the manuscript.

From base-rate to cumulative respect

DOI: 10.1017/S0140525X07001677

C. Philip Beaman and Rachel McCloy

Department of Psychology, University of Reading – Earley Gate, Whiteknights, Reading RG6 6AL, United Kingdom.

c.p.beaman@reading.ac.uk http://www.personal.rdg.ac.uk/~sxs98cpb/philip_beaman.htm r.a.mccloy@reading.ac.uk
http://www.psychology.rdg.ac.uk/people/lecturing/Dr_Rachel_McCloy.php

Abstract: The tendency to neglect base-rates in judgment under uncertainty may be “notorious,” as Barbey & Sloman (B&S) suggest, but it is neither inevitable (as they document; see also Koehler 1996) nor unique. Here we would like to point out another line of evidence connecting ecological rationality to dual processes, the failure of individuals to appropriately judge *cumulative* probability.

Recent data in studies by McCloy and colleagues (McCloy et al. 2007; McCloy et al., submitted) show that judgment of cumulative, disjunctive risk (i.e., the probability of avoiding an adverse event over a period of time during which one continually engages in a risky activity) benefits from presentation in a frequency, rather than a probability, format (McCloy et al., submitted). It does this in a similar manner to the way in which judgments of conditional probability avoid the base-rate neglect fallacy if presented in natural frequency format (Gigerenzer & Hoffrage 1995). Further, training in translation from probability to frequency formats shows similar improvements relative to baseline for both types of judgment (McCloy et al. 2007). However, the effects of both format and training are mimicked by presenting information in a partitive or “nested” set structure (in our studies, diagrammatically represented by probability trees rather than Euler circles). This suggests that similar processes may be involved in both problem types, and we applaud Barbey & Sloman (B&S) for attempting to break down the nature of those processes rather than remaining satisfied with a “natural frequency” label. However, we do not believe that B&S have (yet) produced a full and complete account of the means by which dual processes may produce rationality within certain given ecologies.

One worry is the assertion that people do not have an (evolved or otherwise) capacity to encode frequency information. This is a claim concerning a failure to observe a particular capability and might reflect failures in the observational technique employed as much as a failure in the capability itself. The data reviewed by B&S are based upon individuals’ inability to produce, on demand, explicit knowledge obtained of frequencies from episodes of incidental learning. This is known to be problematic