ELSEVIER

Cognition 81 (2001) 105–111

COGNITION

www.elsevier.com/locate/cognit

Discussion

Single-case probabilities and content-neutral norms: a reply to Gigerenzer

Peter B.M. Vranas*

Department of Philosophy, The University of Michigan, 2215 Angell Hall, Ann Arbor, MI 48109, USA

Abstract

Gigerenzer (2001) argues that (a) statements about single-case probabilities are problematic because they specify no reference class, and that (b) norms should not be applied to specific situations in a 'content-blind' way. In reply I argue that (a) statements about single-case probabilities make sense and are unambiguous despite specifying no reference class, and that (b) although Gigerenzer is right that the application of norms to specific situations should take into account the content of the situations and thus should not be *content-blind*, Gigerenzer has not undermined the appropriateness of probabilistic and other norms which purport to have force in *every* situation and are in this sense *content-neutral*. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Heuristics and biases; Judgment under uncertainty; Norms of reasoning; Frequentism; Concepts of probability; Reference classes

1. Single-case probabilities and reference classes

In response to my statement that "I see no obvious way in which a subjectivist concept of probability is problematic" (Vranas, 2000, p. 182), Gigerenzer tells a story which supposedly "illustrates *one* problem with statements about the probability of single events: a reference class is not specified" (Gigerenzer, 2001, p. 94, this issue). The story concerns a psychiatrist who prescribes Prozac to his depressive patients. Those of his patients to whom he tells (1) become on average more anxious than those to whom he tells (2), because many of the former interpret (1) as (3) rather than (2):

* Fax: +1-734-763-8071.
E-mail address: vranas@umich.edu (P.B.M. Vranas).

0010-0277/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0010-0277(00)00134-7

(1) You have a 30–50% probability of developing a sexual problem.

(2) Out of every ten patients to whom I prescribe Prozac, three to five experience a sexual problem.

(3) You will experience a problem in 30–50% of your sexual encounters.

Now I grant that this story illustrates a *practical* problem, namely one of miscommunication, but how does it illustrate a *theoretical* (conceptual) problem with (statements about) single-case probabilities? Maybe Gigerenzer's point is that (1) is ambiguous or does not even make sense until a reference class is specified. But this is so only if the 'probability' in (1) is *frequentist*; I will argue that if the probability is *objectivist* or *subjectivist*, then we have statements about single-case probabilities which make sense and are unambiguous despite specifying no reference class.

Suppose first that the probability in (1) is objectivist, namely a 'chance' or 'propensity' (e.g. Giere, 1973; Mackie, 1973, pp. 179–187; Popper, 1959). Then (1) says something like (4):

(4) You have a mechanism which causes a sexual problem when triggered and which has a 30–50% chance of being triggered by Prozac.

To see why (4) requires no specification of a reference class, compare (4) with (5):

(5) Three to five out of every ten patients have a mechanism which causes a sexual problem when triggered and which is *always* triggered by Prozac; the remaining patients have no such mechanism.

If (5) is true then (4) is false, because then your chance of developing a sexual problem is 0 or 1, not 30–50%. Although (5) is about a class of patients, (4) is not: (4) is about a *particular* patient and is unambiguous (provided that the nature of the 'sexual problem' is clarified) despite specifying no reference class. It may be true that *normally* one would not assert (4) without having evidence like (2), which does specify a reference class.[1] But this is not *necessary*: I may believe that each patient has a different chance of developing a sexual problem, and in asserting (4) I may be just guessing that in your case the chance is 30–50%. So the point remains that, even if no reference class is specified, (4) *makes sense* and is unambiguous – or, more cautiously (since conceptual worries about chances or propensities do exist: Hájek, in press; Humphreys, 1985), Gigerenzer has given us no good reason to think otherwise.

Similar remarks apply if the probability in (1) is subjectivist, namely a degree of confidence (e.g. Kyburg & Smokler, 1980). Then (1) says something like (6):

(6) I am 30–50% confident that you will develop a sexual problem.

This is again a statement about a particular patient which makes sense and is unambiguous despite specifying no reference class (even if normally one would not assert (6) without having evidence like (2), which does specify a reference class). I conclude that the Prozac story illustrates no theoretical problem with (statements about) single-case probabilities – unless it is assumed that only frequentist probabilities are legitimate, an assumption which Gigerenzer explicitly disavows.

[1] Evidence like (2) does not by itself justify belief in (4) rather than (5): further evidence is needed. So I am *not* saying that the psychiatrist should tell his patients (4) rather than (2).

More generally, like many philosophers (cf. Carnap, 1945; Lewis, 1980/1986), I do not see the three main kinds of probability concepts – objectivist, subjectivist, and frequentist – as competing. They just are different kinds of concepts: chances or propensities (if such things exist) are not relative frequencies, nor are they degrees of confidence. So from the fact that *frequentist* probabilities cannot be defined without specifying a reference class it does not follow that *objectivist* or *subjectivist* single-case probabilities (which require no such specification) are problematic.

2. On the appropriateness of content-neutral norms

2.1. Content-neutrality versus content-blindness

In response to my statement that "it is incumbent on him to explain what is wrong with … content-neutral justifications of probabilistic [and other] norms" (Vranas, 2000, p. 186), Gigerenzer adduces three major considerations – polysemy, sampling, and reference classes – which supposedly demonstrate that "norms need to be *constructed* for a specific situation, not *imposed* upon in a content-blind way" (Gigerenzer, 2001, p. 93, this issue). I address these three considerations in the next three subsections, but first a clarification is needed. Contrary to what Gigerenzer says, I do not think I am using the term 'content-neutral' as he uses the term 'content-blind'. Let us say that a norm is *content-neutral* when it purports to have normative force in *every* context. The moral norm against murdering, for example, understood as the claim that one should *never* murder, is content-neutral. I take Gigerenzer to agree that probabilistic norms are content-neutral in this sense, although it is of course a further question whether these content-neutral norms are *appropriate* (*correct*), whether they *do* have normative force in every context. Now let us say that a norm is *applied in a content-blind way* to a particular context when it is applied without taking into account the specific features of the context. If one claims, for example, that I violated the norm against murdering because I killed someone, and one does not take into account whether the killing was *accidental*, then one applies the norm to my situation in a content-blind way. I agree of course with Gigerenzer's point that this is to be avoided, but this point addresses the *empirical* question of whether I *did* violate the norm; more relevant to the present, *normative* discussion is the question of whether I *should not* violate the norm (whether the norm is *appropriate*), and answering the empirical question does not amount to answering the normative one. I take then much of the disagreement between Gigerenzer and myself to be only apparent, because we are addressing different questions: some of Gigerenzer's examples concern content-blind *applications* of content-neutral norms to particular contexts, but I will argue that his examples fail to undermine the *appropriateness* of content-neutral norms.

2.2. Polysemy and the conjunction rule

The *conjunction rule* is the content-neutral epistemic norm according to which one's degree of confidence in the conjunction of two statements ought not to exceed

one's degree of confidence in (either) one of the conjuncts. Tversky and Kahneman (1983) found that, when people were given a description suggesting that Linda is a feminist but not a bank teller and were asked to rank certain statements according to their probability, most people ranked (a) 'Linda is a bank teller and is active in the feminist movement' as more probable than (b) 'Linda is a bank teller'. Gigerenzer argues that to conclude from these results that most participants violated the conjunction rule would be to apply the rule in a content-blind way: until we find out how the participants understood the polysemous term 'probability' we do not know whether their judgments of probability expressed (i.e. corresponded to) their degrees of confidence. I agree, but then so did Tversky and Kahneman (1983), who explicitly investigated the possibility that "the conjunction fallacy … may be viewed as a misunderstanding regarding the meaning of the word *probability*" (p. 303) but concluded that "the observed violations of the conjunction rule … are genuine fallacies, not just misunderstandings" (p. 304). Gigerenzer might reply by adducing evidence to the effect that most participants understood the term 'probability' non-mathematically (Hertwig & Gigerenzer, 1999). But regardless of who is ultimately right on the *empirical* issue of how most participants understood the term 'probability', the fact remains that Tversky and Kahneman did address this issue, so Gigerenzer is not entitled to accuse them of applying the conjunction rule in a content-blind way. Moreover, Gigerenzer has not addressed the *normative* issue of whether the conjunction rule has normative force in the context of Tversky and Kahneman's experiments. Even if most participants understood 'probability' non-mathematically and did not express their relative degrees of confidence in statements (a) and (b) when they ranked (a) as more 'probable' than (b), arguably their *(unexpressed) degrees of confidence* in these two statements (as opposed to their '*probability*' *rankings* of the two statements) still ought to satisfy the conjunction rule. So Gigerenzer has not undermined the appropriateness of the conjunction rule.[2]

2.3. Sampling and calibration

Suppose I am asked a series of questions and I indicate, after answering each question, my degree of confidence in the correctness of my answer. If only, for example, 70% of the answers for each of which I indicate a 90% degree of confidence are correct, then these answers exhibit *overconfidence*: they violate the norm of *calibration*, according to which the two percentages ought to be equal. Gigerenzer argues that overconfidence disappears in some cases in which people are asked *randomly* sampled questions but appears with questions which are *selectively*

[2] Even if some participants understood 'probability' non-mathematically and did not violate the conjunction rule when they ranked statement (a) as more 'probable' than (b), arguably they still committed a reasoning fallacy (distinct from the fallacy of violating the conjunction rule) if they were not *justified* in understanding 'probability' non-mathematically. Hertwig and Gigerenzer (1999) argue that they *were* justified, "because a mathematical interpretation would render the experimenter's description of Linda … irrelevant to the requested judgment" (p. 278). Not so, however: the description of Linda *was* relevant to some of the statements (e.g. 'Linda is active in the feminist movement') whose relative probabilities the participants were asked to evaluate.

sampled to look easier than they are.[3] For example, the question of whether New York or Rome is further south looks easier than it is because it has an apparently obvious but incorrect answer: 'Rome, since it is much warmer'. Gigerenzer claims that in such cases of selective sampling the norm of calibration has no force. Why not, however? To modify an example that I borrowed from Kahneman and Tversky (1996, p. 588) (Vranas, 2000, p. 185): would Gigerenzer be indifferent between getting a diagnosis of almost certain recovery from a calibrated and from an overconfident surgeon, even when the latter is overconfident *only* with respect to cases in which recovery *looks* certain but is in fact unlikely?[4] Gigerenzer has provided at most an *explanation*, not a *justification*, for overconfidence in cases of selective sampling, so Gigerenzer has not shown that the norm of calibration is inappropriate.

2.4. Reference classes

Gigerenzer considers a study (Slovic, Monahan, & MacGregor, 2000) in which one group of members of the American Academy of Psychiatry and Law were asked something like (7) and another group were asked something like (8):

(7) What is the probability that Mr Jones (an offender) will commit another violent act if given parole?

(8) How many among 100 people like Mr Jones will commit another violent act if given parole?

Gigerenzer notes that "the probability judgments were systematically higher than the frequency judgments" but argues that this discrepancy "need not be equated with a reasoning fallacy" because (similarly to the Prozac story) some people may have interpreted (7) as (9) rather than (8):

(9) If Mr Jones is given parole 100 times, how many times will he commit another violent act?

In reply I grant that the above discrepancy does not correspond to a reasoning fallacy, but who says that it does? What exactly is the content-neutral norm whose content-blind application or whose inappropriateness this example is supposed to illustrate? It cannot be a norm to the effect that one's degree of confidence in a statement ought to be equal to one's estimate of 'the' corresponding relative frequency: there is no such norm precisely because there is no unique corresponding reference class and thus in general no unique corresponding relative

[3] In response to my statement that "experimenters are typically careful to … stipulate random sampling" (Vranas, 2000, p. 187), Gigerenzer argues that until recently there was "no single [calibration] study that used random sampling" (Gigerenzer, 2001, p. 98, this issue). My statement, however, occurred in the context of my discussion not of calibration, but rather of experiments like the Engineer/Lawyer ones, in which experimenters are indeed typically careful to *stipulate* (though not to *use*) random sampling: "in the lawyer–engineer problem, subjects are typically told that the descriptions provided were selected at random" (Koehler, 1996, p. 8).

[4] Responding to an earlier version of this paper, Gigerenzer (pers. commun.) noted in effect that a surgeon who is calibrated with respect to a set of patients will in general be overconfident with respect to a subset and underconfident with respect to another subset. True, but this is so *regardless* of whether patients are sampled randomly or selectively, and need not undermine the norm that the surgeon ought to be (approximately) calibrated with respect to certain large sets of patients.

frequency. What about a norm to the effect that one's degree of confidence in a statement ought to be equal to one's estimate of the corresponding relative frequency *which one takes to be relevant*? Assuming that such a norm is well-defined, Gigerenzer has given us no reason to believe that it has no normative force in the context of the above study: arguably a participant who understands (7) as being about a degree of confidence and takes the relevant corresponding frequency to be the one in (9) ought to give the same answer to (7) and (9) – even if not to (7) and (8). I conclude that the above study, like Gigerenzer's other examples, provides no instance of a content-neutral norm which is not appropriate.

2.5. Conclusion

In the last three subsections I carried out the *negative* task of arguing that Gigerenzer has *not* established the existence of content-neutral norms which fail to be appropriate. In my earlier piece (Vranas, 2000, pp. 185–186) I carried out the *positive* task of providing three arguments for the appropriateness of (content-neutral) probabilistic norms. Given that Gigerenzer in his reply did not even address (let alone refute) those three arguments, I conclude that the balance of reasons favors the appropriateness of probabilistic norms. On the other hand, I agree with Gigerenzer that the application of a content-neutral appropriate norm to a particular problem can be a tricky matter and may require a thorough examination of the content of the problem.

Acknowledgements

I wish to thank Eugene Burnstein, Stephen Darwall, John Doris, Allan Gibbard, Gerd Gigerenzer, Vittorio Girotto, James Joyce, and Richard Nisbett for helpful comments.

References

Carnap, R. (1945). The two concepts of probability. *Philosophy and Phenomenological Research*, *5*, 513–532.

Giere, R. N. (1973). Objective single-case probabilities and the foundations of statistics. In P. Suppes, L. Henkin, A. Joja, & G. C. Moisil (Eds.), *Logic, methodology and philosophy of science* (Vol. IV, pp. 467–483). Amsterdam: North Holland.

Gigerenzer, G. (2001). Content-blind norms, no norms, or good norms? A reply to Vranas. *Cognition, 81*, 93–103.

Hájek, A. (in press). Conditional probability is the guide to life. In H. E. Kyburg Jr., & M. Thalos (Eds.), *Probability as a guide to life*. Open Court.

Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited: how intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, *12*, 275–305.

Humphreys, P. (1985). Why propensities cannot be probabilities. *Philosophical Review*, *94*, 557–570.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591.

Koehler, J. J. (1996). The base rate fallacy reconsidered: descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, *19*, 1–17.

Kyburg Jr., H. E., & Smokler, H. E. (Eds.) (1980). *Studies in subjective probability* (2nd ed.). Huntington, NY: Krieger.

Lewis, D. K. (1986). A subjectivist's guide to objective chance. In D. K. Lewis, *Philosophical papers* (Vol. 2, pp. 83–132). New York: Oxford University Press. (Original work published 1980)

Mackie, J. L. (1973). *Truth, probability and paradox: studies in philosophical logic*. Oxford: Clarendon.

Popper, K. R. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science*, *10*, 25–42.

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, *24*, 271–296.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.

Vranas, P. B. M. (2000). Gigerenzer's normative critique of Kahneman and Tversky. *Cognition*, *76*, 179–193.