Between Probability and Certainty

What Justifies Belief

Martin Smith

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2016

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2015943402

ISBN 978–0–19–875533–3

Printed in Great Britain by
Clays Ltd, St Ives plc

Introduction

The Risk Minimisation Conception of Justification

Some philosophers have claimed that cases involving lotteries provide vivid counterexamples to the traditional analysis of knowledge as justified, true belief (see Hawthorne 2003: 9, Pritchard 2007: 4). They reason along the following lines: suppose I hold a single ticket in a fair lottery of one million tickets. Suppose I am convinced, purely on the basis of the odds involved, that my ticket won't win. Do I *know* that my ticket won't win? Intuitively, I don't know any such thing, even if it happens to be true. Presumably, though, I have plenty of *justification for believing* that my ticket won't win—after all, given my evidence, this proposition has a 99.9999% chance of being true. How much more justification could one want? If I'm not justified in believing that my ticket won't win, then surely none of us is justified in believing much at all. Here is a case, then, in which a justified, true belief fails to qualify as knowledge.

This argument seems straightforward enough, and yet there are reasons for being uneasy about it. On reflection, lottery cases seem somehow *different* from the standard Gettier cases that are used to refute the traditional analysis of knowledge. Consider the following: I wander into a room, undergo a visual experience as of a red wall and come to believe that the wall is red. In actual fact the wall *is* red but, unbeknownst to me, it is bathed in strong red light emanating from a hidden source, such that it would have looked exactly the same to me even if it had been white. Intuitively, I do not know, in this case, that the wall is red, in spite of the fact that my belief is both justified and true.

We can observe a number of apparent differences between these two cases. In particular, while my belief in the Gettier case fails to actually qualify

as knowledge, it nevertheless seems to be a good or promising *candidate* for knowledge—and would have been knowledge if only conditions in the world had been more obliging. My belief in the lottery case, however, doesn't seem to be the sort of belief that could *ever* qualify as knowledge. In the Gettier case, the problem seems to lie with the *world*—and funny, abnormal goings on therein. In the lottery case, the problem seems to lie with *me* and the way in which I form my belief (see Ryan 1996: 136, 137).

The Hawthorne/Pritchard argument betrays a commitment to a certain, quite pervasive, way of thinking about epistemic justification. The picture is something like this: for any proposition P we can always ask how *likely* it is that P is true, given present evidence. The more likely it is that P is true, the more justification one has for believing that it is. The less likely it is that P is true, the less justification one has for believing that it is. One has justification *simpliciter* for believing P when the likelihood of P is sufficiently high and the risk of ~P is correspondingly low. Call this the *risk minimisation* conception of justification.

This general sort of picture can be identified in the work of a very broad range of epistemologists. Sometimes it is made more or less explicit (see Russell 1948: chap. VI, Chisholm 1957: 28, Derksen 1978, Alston 1988, Moser 1988, Plantinga 1993: chap. 9, Fumerton 1995: 18–19, Lewis 1996: 551, Swinburne 2001: chap. 3, Swinburne 2011, Conee and Feldman 2004: n. 32, Pryor 2004: 350–1, Pryor 2005: 181, BonJour 2010, Goldman 2011: s. 16.7). More often it is left implicit, as in the above reasoning. I don't know of any detailed arguments in favour of this picture, though the following thought is admittedly quite compelling: most epistemologists are *fallibilists* of one kind or another and hold that a belief can be justified even if one doesn't have evidence that makes it certain—even if one hasn't completely eliminated all risk of error. But if justification can fall short of evidential certainty, then what else could it possibly *be* if not evidential probability or likelihood? If justification does not require the complete elimination of error risk, then what else could it possibly require if not its minimisation?

When all is said and done, I'm unsure whether I can offer adequate answers to these questions—but I will attempt, in this book, to come to a rather different way of thinking about justification. Some of the views that I'll defend—or at least take seriously—might strike some as obviously wrong. An example might be the view that we can, sometimes, be justified in believing things that are very unlikely to be true, given our evidence. I think that this may in the end be correct—and by 'justification', I don't

simply mean justification that is 'practical' or 'prudential', but justification that is *genuinely epistemic*. But I will begin with ideas that are, I hope, less controversial.

It's clear that there is *something* that sets purely statistical evidence apart from evidence of other kinds. In the case described above, my evidence for the proposition that my ticket will lose the lottery is a clear example of evidence that is purely statistical in character. Here is another (based on an example due to Cohen 1977: §24): Suppose that 100 people attended a concert but only one ticket was ever sold. As such, only one person at the concert attended legitimately and the other 99 were gatecrashers. Suppose we know that Joe was one of the people who attended the concert but have no further information about him. Is Joe a gatecrasher? In one sense it has to be admitted that our evidence in favour of this proposition is very strong indeed. And yet, when pressed, most of us would, I think, be reluctant to give this evidence much weight.

Should I draw the conclusion that Joe gatecrashed the concert and treat him accordingly? Should I go about asserting that Joe is a gatecrasher—should I, for instance, inform his friends, his family, his employer? Should Joe be taken to court and appropriate punishment applied to him? Most of us would be very apprehensive about taking such steps just on the grounds that Joe attended the concert and that 99 out of 100 attendees were gatecrashers. And this is not just a philosopher's intuition—under prevailing legal practice, in a broad range of jurisdictions, statistical evidence of this kind would not be deemed sufficient for a positive finding of fact to the effect that Joe gatecrashed the concert (for some relevant references see Kaye 1982: s. I, Allensworth 2009: s. IIB). From the perspective of the risk minimisation conception, though, this apprehensiveness is puzzling. After all, there's no question that our evidence makes it *very likely* that Joe gatecrashed the concert. By believing, asserting, and acting upon this proposition, we would only be running a very small risk of error.

Perhaps when it comes to believing and asserting that Joe gatecrashed and applying appropriate sanctions to him, we wouldn't be willing to tolerate *any* risk of error, no matter how small. But this really seems not to be the case—for we would be perfectly willing to do such things on the strength of *other kinds* of non-conclusive evidence. Suppose that, instead of our having statistical evidence in favour of Joe being a gatecrasher, we have some eyewitness testimony to that effect. Suppose a witness testifies that she clearly saw Joe scaling the fence at the concert or some such.

As long as we had no reason to doubt the reliability of this testimony, we would usually be willing to take it at face value—and to repeat it and act upon it. Most of us would be quite comfortable with Joe being appropriately punished on the strength of evidence such as this. And, under prevailing legal practice, testimonial evidence of this kind, provided it is not contradicted or otherwise called into question, could constitute sufficient grounds for a finding to the effect that Joe gatecrashed the concert.

We are all perfectly aware, though, that testimony is fallible. Just because an eyewitness testifies that Joe gatecrashed the concert, this doesn't *make it certain* that he did—he may still be innocent. Witnesses are sometimes mistaken and they sometimes lie. The eyewitness testimony undoubtedly makes it *likely* that Joe gatecrashed but, plausibly, it doesn't make it quite as likely as 99%—and yet this is precisely how likely the proposition is, given the statistical evidence about which we seemed so apprehensive. By believing that Joe gatecrashed on the basis of the testimonial evidence, we would actually be running a *higher* risk of error than we would in believing this on the basis of the statistical evidence. As such, the risk minimisation conception straightforwardly predicts that the latter belief should be more justified than the former. Yet this would seem to be the very opposite of the truth.

Maybe there is nothing here that should seriously trouble a risk minimisation theorist. Maybe our judgements about the case reflect an unreasonable bias against statistical evidence and we should train ourselves to give them up. But if we were to try and devise a theory of justification that really did do justice to our judgements here, what would it look like? Apart from anything else, it would need to be a view on which justification somehow demanded *more* than probability, but *less* than certainty—a view on which no amount of purely statistical evidence could make for justification, even though something like testimonial evidence somehow could.

As I've mentioned, the testimony to the effect that Joe gatecrashed the concert still leaves open that possibility that he is innocent. This was an open possibility before we received the testimony and it remains so afterwards. What the testimony does, though, is force us to *reconceptualise* this possibility—to view it in a different sort of light. Once we've received the witness testimony, there is *something that it would take* in order for Joe to still be innocent—it would take a deceit, it would take a misperception, it would take some sort of departure from *normal circumstances*.

In contrast, if my only evidence against Joe is that he attended the concert and that 99 out of 100 attendees were gatecrashers then, while it may be unlikely that Joe is innocent, it wouldn't really take any departure from normal circumstances for this to be true. While I might, in a sense, be surprised to learn that Joe is innocent, in another sense there shouldn't be anything particularly surprising about this, given my evidence. I know that one of the attendees was innocent—and it is no more surprising that this should turn out to be Joe than anyone else. If I believe that Joe gatecrashed, on the basis of the statistical evidence, then my belief could turn out to be false without anything abnormal having transpired. If I believe that Joe gatecrashed on the basis of testimony, then it would take some abnormal circumstance to part my belief from the truth. Viewed in this way, the testimonial evidence really does seem to offer something more than the statistical evidence does.

These observations are in no way peculiar to testimonial evidence per se. Suppose *I* am an eyewitness to Joe's gatecrashing—suppose I get a clear look at him scaling the fence. Like testimonial evidence, direct perceptual evidence does more than just 'load the dice' in favour of a proposition. My perceptual experience may make it less likely that Joe is innocent, but it also makes this possibility *demand something more*—hallucination, perceptual malfunction, disguise, etc.—it makes it demand a departure from normalcy. Purely statistical evidence is distinctive for *not* having this effect.

These remarks do not, of course, amount to some new theory of justification—not yet. At present, they are perhaps little more than suggestive slogans—and they contrast sharply with my formulation of the risk minimisation conception as a relatively clear and precise thesis. My primary aim, in this book, is to try and make something more of these ideas—to build them into something that would count as a viable alternative to the risk minimisation conception.

While my way of thinking about justification may be unfamiliar, it is connected to a more familiar way of thinking about *knowledge*. Many epistemologists have been attracted to the idea that, in order for a belief to qualify as knowledge it must be *safe* from error—it is necessary that the belief could not easily have been wrong. This is often spelled out in terms of possible worlds: in order for a belief to be knowledge it must be true in all close or similar possible worlds at which it is held. In order for a belief to be justified, on my view, there is also a set of possible worlds

throughout which the belief must be true wherever it is held. These are not worlds that are 'close' or 'similar' as such, but worlds that are *normal*.

One might wonder whether those who adopt a safety condition upon knowledge have some additional incentive to adopt a structurally similar condition on justification, such as the one I defend here. There may be some truth to this—but we should proceed with caution. What is clear, though it has not been widely observed, is that combining a safety condition upon knowledge with the risk minimisation conception of justification makes for a kind of *ill-fit* between the two norms. As I will be arguing, there are in fact a number of familiar ideas about knowledge that sit very uneasily alongside the risk minimisation conception of justification.

I mentioned above that the risk minimisation conception of justification is left implicit in the work of a number of epistemologists. But the acceptance of this picture is, by and large, more self-conscious in the so-called 'formal epistemology' tradition where it has been brought to the fore via the paradoxes of rational acceptability. The risk minimisation conception, as it stands, is incompatible with the principle that justification is closed under multiple premise deductive consequence—the principle according to which, if one has justification for believing each of a set of premises, and these premises together deductively entail a conclusion, then one has justification for believing the conclusion. The lottery and preface paradoxes both serve to make this incompatibility vivid.

In spite of its incompatibility with the letter of the risk minimisation conception, the closure principle is one that some epistemologists hold dear. As such, there have been a number of attempts to refine or modify the risk minimisation conception in the hope of circumventing the lottery and preface paradoxes without having to abandon multiple premise closure. Such attempts turn out, however, to be beset by purely formal difficulties. A series of impossibility results, as I shall discuss, come close to demonstrating that nothing recognisable as a refinement of the risk minimisation picture can be consistently combined with multiple premise closure.

My own theory of justification, in contrast, *is* consistent with multiple premise closure, and offers a different sort of treatment of the lottery and preface paradoxes. It's important to bear in mind that these paradoxes are not puzzles waiting to be 'solved' in such a way as to preserve *all* of our pretheoretic impressions. Any treatment of the paradoxes will have

to trade off advantages against disadvantages—and my own theory of justification simply strikes this bargain in a different way to the risk minimisation conception. The lottery and preface paradoxes will both be discussed in detail, along with a number of related puzzles.

In this book, I will be presenting a range of arguments against the risk minimisation conception of justification—arguments of which this introduction has offered a brief preview. But it's not on the strength of arguments that the risk minimisation conception ever became the dominant view and, I suspect, it won't be dislodged by arguments either—or, at least, not by arguments *alone*. The risk minimisation conception is part of a complex of ideas that really do have a remarkable coherence and power—and this may go a long way towards explaining its enduring popularity. It's for this reason that the provision of some kind of alternative is of the utmost importance. My primary aim here is to develop this alternative.