

## Lucky to be rational\*

Adam Elga

Paper presented on June 6, 2008

at the Bellingham Summer Philosophy Conference

---

1

It can be unsettling to think that one's beliefs reflect the influence of arbitrary or irrelevant factors. Here are some examples of such factors in operation:

- Fred comes to realize that if his parents had settled in a more conservative neighborhood, he would have—on the basis of essentially the same evidence—arrived at political views quite different from his actual views. Furthermore, his parents chose between liberal and conservative neighborhoods by tossing a coin. (Sher 2001)
- Jill believes in God, but her twin brother Jack does not. They were raised together and exposed to the same evidence. Their parents confess to Jack, “When you were infants, we gave you both treatments that influenced your brain development. Jill got the ‘theist treatment’: it made arguments for the existence of God seem more persuasive to her. You got the ‘atheist treatment’, with the opposite effect. We tossed a coin to determine who got which treatment.”
- A scholar tosses a coin to determine whether to attend Harvard or Oxford. He attends Oxford, and ends up endorsing the analytic/synthetic distinction. But he finds out that if he had gone to Harvard, he would have rejected the analytic/synthetic distinction, even though he would have been exposed to the same relevant evidence. (Cohen 2000)

In each case, the subject learns that an irrelevant factor made a crucial difference to what he ends up believing. For example, Jack finds out

---

\*Thanks to audiences at Harvard, the University of Vermont, and USC in 2006, audiences at Oxford and NYU in 2007, and Joshua Schechter (who commented on the paper at the 2008 BSPC).

that his atheism can be traced back to a fateful coin toss. Furthermore, he counts the coin toss as irrelevant in the following sense: He doesn't think that the toss outcome is evidentially relevant to whether there is a God.

To repeat: In each case, a subject comes to realize that an irrelevant factor made a crucial difference to how he evaluates evidence. Should that realization reduce his confidence in his original judgment? Consider the scholar. When he finds out about the influence of his graduate school on his views, he might react in one of two ways. On the one hand, he might think, "Attending Oxford introduced a distorting *bias* into my thinking. It made me underestimate the force of Quine's arguments against the analytic/synthetic distinction, arguments that were taken more seriously at Harvard than at Oxford. Now that I have found out about this bias, I am less confident that the analytic/synthetic distinction is sound."

On the other hand, the scholar might think, "Attending Oxford had an important *enlightening*<sup>1</sup> effect on my thinking. It allowed me to see that Quine's arguments have little force. If I'd had the misfortune of going to Harvard, I would have been blind to weaknesses in the argument. How lucky I was to attend Oxford!"

Which reaction is right? More generally, when one learns that a belief reflects the influence of irrelevant factors, how much should that reduce the strength of the belief?

2

The answer is: It depends. It depends on whether the irrelevant factor put one at risk for failing to live up to one's own standards for reasoning. If so, then learning about the factor should reduce the strength of beliefs influenced by the factor. If not, not. That is my thesis.

In explaining and defending this thesis, I will focus on the case of the scholar (the one whose choice of graduate school was made by the toss of a coin). It will be clear that the defense generalizes to other cases.

Here goes. In the above story, the scholar's choice of graduate school influenced his eventual attitude toward the analytic/synthetic distinction. But the exact manner of that influence was not specified. It will be helpful to fill in the details in two different ways.

---

<sup>1</sup>I owe the "biasing versus enlightening" contrast to Pronin et al. (2002, 790).

3

First way. Suppose that the scholar finds out the following:

**Same Standards** Harvard and Oxford instill the very same fundamental epistemic standards in graduate students. But students at each school violate those standards by putting too much stock in the opinions of the local experts. For example, Harvard students are disposed to accept arguments put forward by Quine, even when those arguments have what their own standards would count as defects. And Oxford students have similar dispositions with respect to, say, Strawson (Cohen 2000).

Learning the above should make the scholar much more cautious in his endorsement of the analytic/synthetic distinction. He should think, “Attending Oxford likely made me misapply my own philosophical standards when evaluating Strawson’s defense of the analytic/synthetic distinction – to reason in ways that I think are no good. In particular, it likely made me more inclined to accept those arguments than I should have been. Given this, I should become less confident that the analytic/synthetic distinction is sound.”

I hope it is plausible that the risk of having violated one’s own standards is call for caution. But as further evidence of this, and to bring out some important points, it will be helpful to consider the medical condition known as *hypoxia* (oxygen deprivation).

Hypoxia impairs reasoning and judgment, which is bad enough. But what makes the condition really insidious is that it can be undetectable. In other words, when hypoxia impairs someone’s reasoning, it can do so in such a way that the impaired reasoning seems (to the hypoxic individual) to be perfectly fine (hyp 2004). It is a sad truth that airline pilots have crashed planes by happily and confidently making horrible judgment calls as a result of suffering from hypoxia.

Knowing all of the above, when a pilot finds out that she is at risk for hypoxia, she should become more cautious about her own reasoning, at least when it comes to tricky matters. For example, suppose that a pilot has done some tricky aeronautical calculations and ends up confident that her plane has enough fuel to make it to Hawaii. But then she notices that she is at a very high altitude—an altitude that puts her at risk for hypoxia. Once she notices this, she should become much less confident

that she has done the calculations correctly.<sup>2</sup>

Notice that when the pilot gets evidence that she is hypoxic, she gets evidence that she has been reasoning in ways that violate her own standards. In other words, she gets evidence that she has reasoned in ways that she herself thinks are no good. That is why the evidence of hypoxia should reduce her confidence in the results of that reasoning.<sup>3</sup>

Notice also that this conclusion is independent of whether the pilot has in fact done the calculations correctly. For example, suppose that the pilot has gotten lucky: She suffers no ill effects from the high altitude, and performs the fuel calculation flawlessly. Even so, the pilot would be unreasonable to be confident in the correctness of the calculations once she realizes that she is at risk for hypoxia.

All of this is common sense. Nevertheless, the advice of ratcheting back one's confidence when at risk for hypoxia is not easy to follow. The reason is that it can be difficult to distance oneself from reasoning that seems perfectly good. As a result, piloting magazines often contain articles with alarmist titles such as "Every Breath You Take: Danger Lurks at High-Altitude" (Haenggi 2005) and "Hypoxia: the killer that hides itself." Here is a representative warning:

The problem with hypoxia is that as it takes effect, the pilot does not get any warning. With carbon monoxide poisoning there is at least a feeling of illness. With hypoxia quite the opposite can occur, where the pilot feels *more* confident and relaxed than at normal oxygen levels. The effect of hypoxia masks itself and for that reason should be feared, planned-for and anticipated. (hyp 2004)

Now return to the case of the scholar, described above. All of the lessons from the hypoxia discussion carry directly over. The scholar should become less confident in his reasoning about the analytic/synthetic

---

<sup>2</sup>For simplicity let us assume that hypoxia does not interfere with the pilot's reasoning on the best way to accommodate the news that she is hypoxic.

<sup>3</sup>I should mention that nothing mysterious is meant by "standards for reasoning." To be committed to some standards for reasoning is just to have beliefs about what are proper or improper ways to reason. So to find out that one is at risk for having violated one's standards is to get evidence that there is a tension in one's beliefs. It is to get evidence that certain of one's beliefs — e.g., one's stance on the analytic/synthetic distinction — are in tension with other of one's beliefs — e.g., one's beliefs about good ways to get to conclusions about the analytic/synthetic distinction.

distinction for exactly the same reason that the pilot should become less confident in the reasoning involved in her fuel calculation.

Furthermore, if it became known that philosophy graduate school generally had the sorts of effects described above, then the American Philosophical Association should issue alarmist warnings to incoming graduate students:

The problem with graduate school is that as it takes effect, the student does not get any warning. With carbon monoxide poisoning there is at least a feeling of illness. With graduate school quite the opposite can occur, where the student feels *more* confident and relaxed than when not in school. The effect of graduate school masks itself and for that reason should be feared, planned-for and anticipated.

The bottom line is that when one finds out that an irrelevant factor has influenced one's belief, that can be evidence that the belief in question was formed by a process that violated one's own standards. For example, it can be evidence that one is less able to catch errors in reasoning when that reasoning is widely endorsed by one's intellectual buddies.

When one finds out that an irrelevant factor influenced one's belief, that *can* be evidence that the belief in question was formed by a process that violated one's own standards. But it need not be. What then?

4

Let us construct a case in which the scholar learns that his choice of graduate school influenced his take on the analytic/synthetic distinction *without* putting him at risk for violating his own standards for reasoning. (Equivalently: without putting him at risk for having his conclusions conflict with his idea of how those conclusions should be arrived at.) Suppose that instead of the "Same Standards" story above, the scholar learns the following about the influence of graduate school on incoming students:

**Different Standards** At Harvard and Oxford, students come to adopt different fundamental epistemic standards. Harvard students adopt the Harvard standards, and perfectly apply them. Oxford students do the same for the Oxford standards. As a result, students at each school tend to agree with the local experts because they share standards with those experts.

I claim that learning that his beliefs were influenced in this way should *not* make the scholar more cautious in his endorsement of the analytic/synthetic distinction. Instead he should think, “The arguments I accept are in perfect accord with my own standards. Or at least: The influence of graduate school gives me no special reason to think that I’m violating my own standards. So the influence of graduate school on my thinking gives me no special reason to doubt that thinking.”

I expect that the reader will balk at this claim. I admit that I have some explaining to do. Let me lay my cards on the table.

Focus on the unsettling feeling one gets when one notices that one’s way of evaluating evidence reflects various accidents of fate: accidents in one’s early development, in one’s later schooling, and even perhaps in the evolutionary history of one’s capacities. My goal in this paper is to diagnose that feeling, and to pin it down.

Here is my diagnosis. I think that the unsettling feeling arises from two sources. First, finding out about irrelevant influences on one’s thinking can sometimes be evidence of a failure to live up to one’s own standards for reasoning. That’s what goes on in the “Same Standards” version of the scholar case, described above.

Second, even when such a failure is not in question, finding out about irrelevant influences on one’s beliefs can *still* be unsettling. For example, consider the “Different Standards” version of the graduate school case, in which students at each school reason in perfect harmony with their own standards. Assume, for example, that students Harvard achieve one reflective equilibrium, and students at Oxford achieve a quite different equilibrium. Even under these conditions, if I were in the scholar’s position, I would find it unsettling to think that if I’d gone to Oxford, I’d have come to the opposite conclusion about the analytic/synthetic distinction.

Where does *that* unsettling feeling come from? I propose that it comes from the following sort of reasoning:

CHANCE OF ERROR Even though both the Harvard and Oxford students live up to their own standards, someone’s standards are *wrong*. Rationality requires more than just living up to one’s own standards. And the Harvard and Oxford stances on the analytic/synthetic distinction are so different that at most one of them can be justified. So I had at least a 50% chance of ending up with an unjustified stance on the analytic/synthetic distinction. Further-

more, I have no independent reason to think that I got lucky. So it would be unreasonable for me to be confident that I got lucky.

The above reasoning has some real force. It would be a hard project to handle the problem that it poses, and I won't tackle that project here. What I would like to argue is that the above problem is no more troubling<sup>4</sup> than a much more general skeptical worry. Here is the general worry:

**POSSIBILITY OF ERROR** Some possible states of belief are coherent and stable—they look fine “from the inside”—and yet are unjustified. Furthermore, I have no independent reason to think I'm not in such a state. So it would be unreasonable for me to be confident that I'm not in such a state.

Now, I have nothing here to say about how to resolve this general skeptical worry. And I do not claim here that coherentism is correct—for all I say, some belief states may be perfectly coherent and yet unjustified. What I'd like to argue is that **CHANCE OF ERROR** presents no *additional* worry over and above the very general skeptical problem posed by **POSSIBILITY OF ERROR**.

To see why this is so, let me restate the last few steps in each argument.

The last few steps of **CHANCE OF ERROR**:

- I had at least a 50% chance of ending up with an unjustified stance on the analytic/synthetic distinction.
- I have no **independent** reason to think that I got lucky.
- Therefore, it would be **unreasonable** for me to be confident that I got lucky.

The last few steps of **POSSIBILITY OF ERROR**:

- Some perfectly coherent stances on the analytic/synthetic distinction are unjustified.
- I have no **independent** reason to think that I avoided such a stance.

---

<sup>4</sup>Thanks here to Earl Connee

- Therefore, it would be **unreasonable** for me to be confident that I avoided such a stance.

Notice that in each case, the crucial inference is from

I have no **independent** reason to think such-and-such.

to

It would be **unreasonable for me to be confident that** such-and-such.

In order to block the general skeptical argument, one must block this step – the step from “no independent reason” to “no good reason.” Again, I have no proposal here on how to do that. But suppose that there is a way to block that step, and thereby to block POSSIBILITY OF ERROR. Then that way would also block CHANCE OF ERROR. That is why CHANCE OF ERROR poses no additional difficulty over and above POSSIBILITY OF ERROR.

To repeat: The crucial move in CHANCE OF ERROR is the inference from “no independent reason” to “no good reason.” That move has nothing in particular to do with irrelevant influences on one’s beliefs. It does not require any thoughts of the form: “If I had gone to a different graduate school, I would have come to the opposite conclusion.” Why, then, do such thoughts prompt the worry? I speculate that they do so because they make salient the possibility of global error.

Let me sum up the diagnosis: Getting the news that one’s beliefs were influenced by irrelevant factors prompts an unease deriving from two sources. First, the news may be evidence that one has violated one’s own standards for reasoning. Second, the news may make salient the possibility of global error, and hence make salient a general skeptical worry. But such news never adds to the force of the skeptical worry, only to its prominence. So increased caution is only in order in the first case, the case in which one gets evidence that one has violated one’s own standards for reasoning.

5

Zoom out. Here is what has been accomplished. The big question is: Should finding out that a belief reflects the influence of irrelevant



factors reduce the strength of that belief? We have considered two cases, illustrating two ways in which the irrelevant factor might influence the belief in question. In the “Same Standards” case, the factor influences one’s belief by making one violate one’s own standards for reasoning (in that particular case, by accepting too easily the arguments of the local experts). The hypoxia example shows that in such cases, one should reduce the strength of the relevant belief.

In the “Different Standards” case, the irrelevant factor influences one’s belief not by making one violate one’s standards, but rather by influencing which standards one comes to accept. In the cleanest version, one is sure that one’s reasoning lives up to one’s own standards, but is also sure that one’s standards reflect the influence of an irrelevant factor.

I claimed that in such cases, finding out about the influence of the irrelevant factor should *not* reduce one’s confidence in the belief in question. But I admitted that it can be unsettling to notice that an irrelevant factor influenced one’s standards. Finally, I offered the following explanation of that unsettling feeling: Thinking about such cases makes salient the possibility of global error, and hence makes salient a general skeptical worry about reasoning.

6

Despite the above arguments, you may still have a lingering sense that even in the “different standards” case, it is unreasonable to just insist: “Yes, my graduate school influenced my standards. So I sure was lucky to have the coin toss send me to the *good* school!”

To accentuate this sense, consider the following case.<sup>5</sup> You find out the following: There was a coin toss while you slept last night. If the coin landed heads, nothing special happened. If the coin landed tails, your brain was put into a vat, and fed just the experiences you would have otherwise had.

In response to this information, one might give the following “I got lucky” response: “Yes, the coin determined whether I would be envatted. So I sure was lucky to have the coin put me in the *good* (non-envatted) case. How lucky I am to have my senses operating veridically!”

But this response rings false. Surely in this chancy-envatting scenario

---

<sup>5</sup>The comparison was suggested to me independently by Gideon Rosen and Susanna Rinard.

one should be seriously uncertain about whether one has been envatted. What's the difference? Given that the "I got lucky" response is so feeble in the chancy-envatting case, how could it be plausible in the chancy-graduate-school case?

The answer is that in the case of perceptual appearances, there is always plenty of room to take the appearances as mere inputs, to step back from them and recalibrate them given positive evidence that they are misleading. One does this whenever one puts on tinted glasses, or pays attention to the "objects may be closer than they appear" sign on one's side-view mirror.

In contrast, it is at least sometimes permissible to stand behind one's all-things-considered judgments, even while being convinced that a coin flip determined the way one would make those judgments. For suppose not. Suppose that it is never reasonable to stand behind one's judgments while thinking that a coin toss determined how those judgments would be made. Then no reasonable person is ever confident that her current method for producing judgments was the result of a coin toss. And no reasonable person ever *learns* that her current method for producing judgments was the result of a coin toss.

But that is absurd! Surely it is possible to learn that one's method for producing judgments was the result of a coin toss. For example, think back to Jill, whose parents confess to her that they tossed a coin to determine whether she would get the 'theist treatment' or the 'atheist treatment'. *Must* Jill think her parents liars or deluded? Must she keep thinking this even after her parents show her the theist and atheist drugs and demonstrate their effects on third parties? When they play for her the video footage of them tossing the coin and administering the injection to her infant self? When she reads her name in a yellowing newspaper article with headline: PARENTS TOSS COIN TO DETERMINE SPIRITUAL LEANINGS OF OWN CHILD?

No.

She should, upon seeing all of the above, become very confident that her own present all-things-considered judgments about God are in large part a result of the coin toss.

Now it may be that her present judgments about God are different than the ones she had before she learned about the treatment. As I've argued, that should be so whenever the treatment subjects her to a risk of reasoning in a way that she herself thinks is no good. But at the end of the day, she may end up in some perfectly coherent state of mind.

She will end up with *some* state of opinion about whether there is a God. And she should end up confident that this opinion was in part the result of a coin toss.

(To make the dependence clear, assume that the theist and atheist treatments don't just influence one's *initial* spiritual inclinations. They also influence one's all-things-considered judgments about spiritual matters, even given evidence that one was subject to the treatment.)

So it is sometimes permissible to say, of one's all-things-considered epistemic judgments: "I got lucky." For denying this leads to the absurdity that no reasonable person ever learns that her judgments are the result of a chance process.

In contrast, there is no corresponding absurdity lurking in the perceptual case. When one learns that a chance process made the difference between envatting and no-envatting, one is *never* reasonable in just saying "I got lucky." One should instead treat one's perceptual experience as potentially misleading. And that is compatible with thinking that a chance process influenced whether one's perceptions are veridical or illusory.

Bottom line: sometimes—once doubts about incoherence have been set aside—it is reasonable to say: "I am lucky to be rational."

### References

- Hypoxia: the killer that hides itself. *Instrument Pilot*, January-February 2004. URL <http://www.pplir.org/journals/old/InstrumentPilot41.pdf>. Editorial.
- G. A. Cohen. *If you're an egalitarian, how come you're so rich?* Harvard University Press, Cambridge, MA, 2000.
- Mike Haenggi. Every breath you take: Danger lurks at high-altitude. AVweb.com, October 2005. URL <http://www.avweb.com/news/aeromed/190790-1.html>.
- E. Pronin, C. Puccio, and L. Ross. *Understanding misunderstanding: Social psychological perspectives*. Cambridge University Press, 2002.
- G. Sher. But I could be wrong. *Social Philosophy and Policy*, 18(2):64–78, 2001.