The Two Envelope "Paradox"

By Frank Jackson, Peter Menzies, and Graham Oppy

A person is offered a choice between two envelopes, A and B. She is told that one envelope contains twice as much money as the other but has no information as to which one that is. She chooses A, say. Before opening A she asks herself whether she ought to have taken B instead. There is a line of reasoning which suggests that she should. Suppose that the amount of money in A is \$x. Then B either contains \$2x or \$0.5x. Each possibility is equally likely, hence the expected value of taking B is 0.5.\$2x + 0.5.\$0.5x = \$1.25x, a gain of \$0.25x. This conclusion cannot be right. The mere choosing of A cannot give her a reason to say that she ought to have picked up B instead. For the situation is symmetrical as between A and B, at least until one of them is opened. Moreover, had she chosen B initially the same reasoning would suggest that she ought to have chosen A instead. No matter which choice she makes the reasoning leads to the conclusion that she made the wrong choice!

It is easy to say how she ought to have done the expected value calculation so as to avoid the absurd (clearly mistaken) result. She should have reasoned that the following two situations are equally likely: A contains \$x, B contains \$2x; and A contains \$2x, B contains \$x. The expected value of taking A is then \$1.5x, and the expected value of taking B is also \$1.5x, and hence there is no reason for her to judge that she should have chosen B instead of A. The problem is to say what is wrong with the first way of doing the calculation.

The important difference between the two calculations is that in the first, fallacious calculation, but not in the second, correct way of calculating the relevant expected values, 'x' ranges over the amount of money in some particular envelope — A, as we supposed. (In the second calculation 'x' ranges over the amount of money that has a 0.5 chance of being in A and a 0.5 chance of being in B.) This means that the first way of doing the calculation involves supposing that for any value of x, if $x is the amount of money in some particular envelope, it is equally likely that $2x or $0.5x is the amount in the other envelope. This supposition is not part of what is given in setting up the puzzle. Moreover, it is *false*, at least for any rational, minimally informed person. Such a person will have a prior probability distribution concerning the *total* amount of money in the two envelopes which means it is *not* equally likely that if $x is the amount of money in one envelope, $2x or $0.5x is the amount in the other envelope, for all x. Suppose, for example, that she thinks that there is almost no chance that the total amount of money in the two envelopes will be greater than $1000. Then, if she considers the possible case in which the amount of money in one envelope is $600, she will not suppose that the probability that the other envelope contains $300 is the same as the probability that the other envelope contains $1200, for that would take the total over $1000. Conversely, for a small enough value of x, she will give almost no chance to the other envelope containing $0.5x, for that would take the contents of the other envelope below the smallest available item of currency.

Why will a rational, minimally informed person have such a prior probability distribution concerning the total amount of money in the two envelopes? Because it is common knowledge that there is a finite amount of money in the world and that there is a smallest unit in any currency. Of course, for *some* values of x the supposition of equal likelihood may well be true. Suppose it is true for x = 20. Then if she opens A, say, and finds that it contains $20, she knows that B either contains $10 or contains $40. As *ex hypothesi* in this case each possibility is equally likely, the expected value of the contents of B is $5 + $20, and accordingly she ought to decline the contents of A and choose B instead.

James Cargile [1] has recently discussed a variant of the two envelope paradox, but does not endorse the solution just given.[1] In the version he considers, our subject is given the following information. An unknown amount of money $x was placed in one envelope (marked with a red spot, say). A coin was tossed. If it came down heads, $2x was placed in the other envelope; if it came down tails, $0.5x was placed in the other envelope. She is then to choose one of the envelopes. As Cargile notes, in this case the person ought to choose the unmarked envelope because it 'corresponds causally to accepting a bet on a fair coin at payoff of double or half'[2]. But, as he observes, surely we correctly reach this conclusion by an application of the line of reasoning which led to the absurd result in the original version of the puzzle, and which we declared to be fallacious in the original version.

There is, however, a crucial difference between the two cases. In the variant case, for any value of x, if $x is the amount of money in the *marked* envelope, it *is* equally likely that $2x or $0.5x is the amount in the other envelope. By contrast with the original case, the prior probability distribution of the subject over the total amount of money likely to be in the envelopes does not undermine this. Suppose, as before, that she thinks that the total amount of money in the two envelopes is very unlikely to exceed $1000. What happens in the variant case if she opens the marked envelope and finds $600? Because she knows that the amount in the unmarked envelope was assigned by halving or doubling a sum of money she *now* knows to be $600, she will be forced to revise

---

[1] There is textual evidence that the solution we have offered was put to him. We presume that the reason he did not accept it relates to his variant on the paradox we now discuss.

[2] [1], pp. 212 - 3.

upwards the probability (from near to zero to near to a half) she now gives to the total being $1800. (She will also give close to a half to the total being $900, of course.) Whereas in the original case, finding $600 in the envelope in her hand would have the effect of forcing up dramatically the probability she gives to the envelope in her hand being the one with the most money, finding $600 in the marked envelope in the variant case instead forces up the probability she gives to the total amount of money being over $1000.

Moreover, the conclusion in the variant case does not violate symmetry considerations in the way that the conclusion in the original case does. In the original case there is no relevant difference between A and B, and yet we are led to the conclusion that merely choosing without opening A means that our subject should have chosen B. And we noted that if the reasoning worked we have an *a priori* argument that no matter which of A and B she chooses, she chooses wrongly. The situation is *not* symmetrical in the revised case. One envelope is marked and the other is not, and this matters. For the amount of money in the unmarked envelope was determined by taking the money to be put in the marked one, and doubling or halving it according to the toss of a coin.

There is a final matter that deserves comment. Our diagnosis of where the reasoning in the original case goes astray depended on the fact that the subject knows that the amount of money is bounded top and bottom. This assumption could be dispensed with in fantastical cases. We can imagine a case where the possible payoffs in envelopes A and B are unbounded top and bottom, and where the subject gives equal (infinitesimal) probability to each and every possible total distributed between the two envelopes. The symmetry of the situation would, however, be unaltered by this change. It would still be wrong to prefer one envelope over the other. However, we cannot offer the same diagnosis of the error in the expected value calculation to the opposite conclusion. There is all the same an error. As Richard Jeffrey [2], chapter 10, has pointed out in connection with other puzzle cases involving infinite domains, we can reasonably insist that "the standard method" for probabilistic and expected value reasoning ought not to be applied in such cases. Finitude assumptions are built into the very foundations that justify such reasoning.[3]

Australian National University

Canberra ACT 0200 *26 Feb 1993*

REFERENCES

[1] James Cargile, 'On a Problem about Probability and Decision', *Analysis* 52 (1992) 211 - 216.

[2] Richard Jeffrey, *The Logic of Decision* (London: University of Chicago Press, 2nd ed. 1983).

[3] We have benefited from discussions with Denis Robinson and Lloyd Humberstone, and many, many others.