biased seemings

belief and upbringing

Pascal's wager

challenge to the wager

Does what we should believe depend on why we believe it?

Our topic today is a cluster of issues about the relationship between **what** you believe and **why** you believe it.

The first concerns a rule of belief which we introduced last time:

Seems → Belief
If it seems to you that P is true, and you have no argument against P, you should believe P.

But this leads to a question: what if how things seem to us is based on bias of some sort?

The second begins with the fact that many of us have certain beliefs because of the way in which we were raised; we would have different beliefs if we had been raised differently. But shouldn't that make us doubt those beliefs?

The last concerns a question about the relationship between belief and action. We can concede that it makes sense for me to choose a certain action because it would make me happy. Does it also make sense to choose to have a certain belief because it would make me happy?

Seems → Belief

If it seems to you that P is true, and you have no argument against P, you should believe P.

Let's start with the first question. This is the problem of how we can trust the way things seem to us, if the way things seem to us can be affected by biases and beliefs which may well be false.

Let's look first at how things seem to us in our visual experiences. Some interesting studies have been done which seem to show that our background beliefs, expectations, and desires can have an effect on how things visually appear to us.

In one well-known study, white Americans were first shown a picture of either a white man's face or a Black man's face, and then shown a picture of either a tool or a gun. Under time pressure, they had to categorize what they were shown. Participants primed with a Black man's face mischaracterized tools as guns significantly more than those primed with a white man's face.

The best way to interpret this study is controversial. But what seems reasonably clear is that whether the participant saw a white face or a Black face affected whether it seemed to them that they were being shown a gun or a tool.

In one well-known study, white Americans were first shown a picture of either a white man's face or a Black man's face, and then shown a picture of either a tool or a gun. Under time pressure, they had to categorize what they were shown. Participants primed with a Black man's face mischaracterized tools as guns significantly more than those primed with a white man's face.

The best way to interpret this study is controversial. But what seems reasonably clear is that whether the participant saw a white face or a Black face affected whether it seemed to them that they were being shown a gun or a tool.

Similar results have been obtained in less politically charged contexts. In one case, people are given two beers, one of which has some balsamic vinegar in it, and asked to pick which one they liked better. A majority chose the one with balsamic vinegar in it. The experiment was then repeated with the change that participants were told in advance that one of the beers had some vinegar in it (but not which one). A majority chose the one with ought vinegar in it. Some infer that the expectation of a vinegar taste changed the way the liquid tasted to the subjects.

One reason why these cases are interesting is that they call into question Seems → Belief. If our background beliefs can affect the way things seem to us, then it is tempting to say that we should trust the seemings only if we should have the belief on which the seemings are based.

One reason why these cases are interesting is that they call into question Seems → Belief. If our background beliefs can affect the way things seem to us, then it is tempting to say that we should trust the seemings only if we should have the belief on which the seemings are based.

Here's an interesting case, from the contemporary philosopher Susanna Siegel:

Jill, for no particular reason, has the belief that Jack is angry. This is a belief which Jill should not have.

When Jill sees Jack, Jill's belief that Jack is angry at her makes Jack look angry to her — it causes it to seem to her that Jack looks angry.

On the basis of the fact that it visually seems to her the Jack is angry, Jill's belief that Jack is angry at her is strengthened.

At the start, Jill should not believe that Jack is angry. If Seems → Belief is true, it looks like at the end she **should** believe that Jack is angry. But can this be right? Does Jill really have a better reason for her belief at the end than at the start?

At the start, Jill should not believe that Jack is angry. If Seems → Belief is true, it looks like at the end she **should** believe that Jack is angry. But can this be right? Does Jill really have a better reason for her belief at the end than at the start?

We have been focusing on perceptual cases. But the moral of the above cases would seem to apply even more strongly to cases of non-perceptual seemings.

Consider the way in which your political beliefs can affect what claims seem true to you. This is an instance of the well-known phenomenon of **confirmation bias**.

Cases of confirmation bias are structurally the same as the Jack/Jill case: one begins with a belief (which might well be a belief one should not have), that belief causes other claims to seem true, and those other claims support the original belief.

If Seems → Belief is true, this kind of thing is perfectly ok. One's belief in P can be justified by one's belief in Q, even if one believes Q because Q seems true and Q seems true because one believes P.

But doesn't this seem like the kind of circular reasoning we would reject in other contexts?

But doesn't this seem like the kind of circular reasoning we would reject in other contexts?

It is worth thinking about how we might modify Seems → Belief in response to cases of this kind. Here's one suggestion:

Restricted Seems → Belief

If it seems to you that P is true, and you have no argument against P, <u>and the seeming is not caused by a belief you should not have</u>, you should believe P.

This rule restricts the seemings you should trust to the ones that are not caused by beliefs you should not have. This would block the result that Jill should believe that Jack is angry.

The problem, though, is that it is hard to know how one could employ this rule. After all, the problem with the cases under discussion is that one can't tell from the inside when a seeming is caused by one of one's beliefs.

This rule restricts the seemings you should trust to the ones that are not caused by beliefs you should not have. This would block the result that Jill should believe that Jack is angry.

The problem, though, is that it is hard to know how one could employ this rule. After all, the problem with the cases under discussion is that one can't tell from the inside when a seeming is caused by one of one's beliefs.

Here's an idea. Perhaps you should not trust seemings when **you have good reason to think that** the seeming is based on an unjustified belief, and hence good reason to think that the seeming is unreliable:

Restricted Seems → Belief 2.0

If it seems to you that P is true, and you have no argument against P, <u>and you have no good reason to think that the seeming is unreliable</u>, you should believe P.

What does this say about the case of Jack and Jill? It says that, if Jill is not aware that her belief played a role in Jack seeming angry, she should form the belief that he is angry. (After all, she had no way of knowing that the seeming was unreliable.)

Restricted Seems → Belief 2.0
If it seems to you that P is true, and you have no argument against P, <u>and you have no good reason to think that the seeming is unreliable</u>, you should believe P.

What does this say about the case of Jack and Jill? It says that, if Jill is not aware that her belief played a role in Jack seeming angry, she should form the belief that he is angry. (After all, she had no way of knowing that the seeming was unreliable.)

But suppose that Jill is told that beliefs about people can play a big role in determining how you perceive their emotions. This information would give her reason to think that Jack's seeming angry to her is unreliable — and in that case she should not reinforce her belief that Jack is angry.

In general, it seems like the best course of action for those who accept Seems → Belief is to restrict it in some way, and to educate themselves about the various situations in which background beliefs (or other mental states) are most likely to affect how things seem to them.

Let's turn to our second topic: the dependency of our beliefs on our upbringing.

We've all considered the thought that we would have different beliefs if we were raised in a different society, or by a different family. And it is a familiar idea that this can, and perhaps should, lead us to doubt those beliefs.

Here's an example:

> I was raised a Catholic, and still am. But I know that other people were raised in different religions, and that people tend to believe the religion in which they were raised. On reflection, I think that it is probably true that, if I had been raised a Muslim, I would probably still be a Muslim. But it is just an accident that I was born into a Catholic family. So it is an accident that I think that Catholicism is true. So, I should give up my belief in Catholicism.

I was raised a Catholic, and still am. But I know that other people were raised in different religions, and that people tend to believe the religion in which they were raised. On reflection, I think that it is probably true that, if I had been raised a Muslim, I would probably still be a Muslim. But it is just an accident that I was born into a Catholic family. So it is an accident that I think that Catholicism is true. So, I should give up my belief in Catholicism.

Of course, it is possible that you were raised Catholic but now think that you have independent arguments to support your Catholicism. But suppose that that is not the case; would it then make sense to discard your belief, on the grounds that were you raised differently you would not have that belief?

Can we formulate a possible rule of belief which would be a candidate to explain this? Here's a natural suggestion:

Social Dependency → No Belief

If you believe P, but would not have believed P if you had been raised in a different society, you should not believe P.

Can we formulate a possible rule of belief which would be a candidate to explain this? Here's a natural suggestion:

Social Dependency → No Belief
If you believe P, but would not have believed P if you had been raised in a different society, you should not believe P.

This has some plausibility to it. For if the only reason why you have some belief is that you were raised to believe it, shouldn't you then think that the belief is just an accident of your upbringing, and should be discarded?

But on closer examination this leads to some pretty implausible results. For the following all look reasonably plausible:

If you had been raised in the family of Genghis Khan, you would have thought that torture is permissible.

If you had been raised in ancient Greece, you would have thought that slavery is permissible.

If you had been raised in the middle ages, you would have thought that some animals came to life by spontaneous generation.

Social Dependency → No Belief
If you believe P, but would not have believed P if you had been raised in a different society, you should not believe P.

If you had been raised in the family of Genghis Khan, you would have thought that torture is permissible.

If you had been raised in ancient Greece, you would have thought that slavery is permissible.

If you had been raised in the middle ages, you would have thought that some animals came to life by spontaneous generation.

But should these facts cause you to abandon your belief in the wrongness of torture or slavery, or your belief in the truth of the theory of biogenesis? Surely not.

And that just seems to show that Social Dependency → No Belief is false.

But should these facts cause you to abandon your belief in the wrongness of torture or slavery, or your belief in the truth of the theory of biogenesis? Surely not.

And that just seems to show that Social Dependency → No Belief is false.

But that might seem to be a somewhat unsatisfactory stopping place. Can't reflection on the fact that your beliefs are just a product of certain kinds of socialization give you good reason to doubt those beliefs?

If you think that it can, then there are two options. The first would be to try to find a way to modify the above rule of belief into one which does not have consequences like the ones just listed.

But there is also another option. Here's an example:

> On reflection, I realize that the only reason why I believe that God exists is that my parents told me this. So I was trusting in my parents' reliability. But I now have good reason to think that my parents are entirely unreliable, and that I shouldn't trust anything they say. So, I now think that the reason why I believed that God exists was not a good reason. So, I should abandon this belief.

But there is also another option. Here's an example:

> On reflection, I realize that the only reason why I believe that God exists is that my parents told me this. So I was trusting in my parents' reliability. But I now have good reason to think that my parents are entirely unreliable, and that I shouldn't trust anything they say. So, I now think that the reason why I believed that God exists was not a good reason. So, I should abandon this belief.

This kind of reasoning seems perfectly legitimate. But it seems that it has nothing special to do with the dependence of your beliefs on your upbringing.

Someone carrying out the above line of reasoning seems to be relying on something like the following rule of belief:

Bad Reasons → No Belief
If the only reason why you believe P is that you believe Q, and you come to believe that Q is a bad reason to believe P, you should not believe P.

Someone carrying out the above line of reasoning seems to be relying on something like the following rule of belief:

Bad Reasons → No Belief
If the only reason why you believe P is that you believe Q, and you come to believe that Q is a bad reason to believe P, you should not believe P.

There are at least two ways in which you might come to find that your believing Q is a bad reason for your believing P.

First, you might discover that Q is false. For example, you might believe that South Bend has great weather only because you believe that South Bend is in California, and discover that the latter belief is false. That would be a good reason to give up your belief that South Bend has great weather.

Second, you might discover that Q does not make P likely to be true. For example, you might believe that South Bend has great weather only because you believe that South Bend is in Indiana, and discover that being in Indiana is not likely to make a city have great weather. That too would be a good reason to give up your belief that South Bend has great weather.

Social Dependency → No Belief
If you believe P, but would not have believed P if you had been raised in a different society, you should not believe P.

Bad Reasons → No Belief
If the only reason why you believe P is that you believe Q, and you come to believe that Q is a bad reason to believe P, you should not believe P.

If you find this plausible, that provides a kind of indirect reason for doubting Social Dependency → No Belief. On this view, discovering that your beliefs are due to your upbringing might well be an occasion for you examining the reasons for those beliefs.

And, when you do that, you might find that your reasons for holding those beliefs are bad. Then, plausibly, you should ditch them. But here the dependence of your beliefs on society is just the occasion for re-examination — it is not the **reason** why you should ditch them.

After all, there is no reason why you could not discover that your beliefs are due to your upbringing but, on examination, find that you have no reason to doubt beliefs that you came to have in this way.

Let's turn to our third topic. This is the question of whether we should ever form beliefs, not because we have reason to think that the belief is true, but for **practical** reasons. Our focus will be a famous argument from Blaise Pascal.

Pascal was a 17th century French philosopher, theologian, and mathematician; he made foundational contributions to, among other areas, the early development of the theory of probability.

Pascal was one of the first thinkers to systematically investigate the question of how we should make decisions under situations of uncertainty, where we don't know all of the relevant facts about the world, or the outcomes of our actions.

He thought that one such decision was the decision whether or not to believe in God:

Let us then examine this point, and say, "God is, or He is not." But to which side shall we incline? Reason can decide nothing here. There is an infinite chasm which separates us. A game is being played at the extremity of this infinite distance where heads or tails will turn up. What will you wager? According to reason, you can do neither the one thing nor the other; according to reason, you can defend neither of the propositions.

Pascal thought that God so far exceeds our comprehension that we have no way of using our reason to decide whether or not God exists.

But, Pascal thinks, this does not remove the necessity of choosing whether or not to believe in God.

But, Pascal thinks, this does not remove the necessity of choosing whether or not to believe in God.

> "Yes, but you must wager. It is not optional. You are embarked. Which will you choose? ... Let us weigh the gain and the loss in wagering that God is. ... If you gain, you gain all; if you lose, you lose nothing. Wager then without hesitation that He is."

"Yes, but you must wager. It is not optional. You are embarked. Which will you choose? ... Let us weigh the gain and the loss in wagering that God is. ... If you gain, you gain all; if you lose, you lose nothing. Wager then without hesitation that He is."

Pascal is here drawing an analogy between the choice whether or not to believe in God and the choice whether or not to make a bet.

Betting, after all, is another case in which we make decisions under uncertainty.

The study of how it is rational to act under certain kinds of uncertainty is now known as "decision theory." We can use some concepts from decision theory to get a bit more precise about how Pascal's argument here is supposed to work.

Let's look at one more quote to get a sense of Pascal's thinking:

> "It would be unwise of you, since you are obliged to play, not to risk your life to win three lives at a game in which there is an equal chance of winning and losing. But here there is an infinity of happy life to be won ... and what you are staking is finite. ... And thus, since you are obliged to play, you must be renouncing reason if you hoard your life rather than risk it for an infinite gain, which is just as likely to occur as a loss..."

Here Pascal is thinking of bets where you might win or lose something by playing, but where what you win is greater than what you lose.

Let's consider how we might reason about decisions of this sort

Let's consider how we might reason about decisions of this sort

I'm about to flip a coin, and offer you the following bet: if the coin comes up heads, then I will give you $5; if it comes up tails, you will owe me $3. You know that it is a fair coin. Should you take the bet?

Courses of action	Heads	Tails
take the bet	$5	-$3
don't take the bet	$0	$0

Courses of action	Heads	Tails
take the bet	$5	-$3
don’t take the bet	$0	$0

There is a ½ probability that the coin will come up heads, and a ½ probability that it will come up tails. In the first case I win $5, and in the second case I lose $3. So, in the long run, I’ll win $5 about half the time, and lose $3 about half the time. So, in the long run, I should expect the amount that I win per coin flip to be the average of these two amounts — a win of $1.

Here neither course of action dominates the other; but it still seems that you should clearly take the bet. Why?

There is a ½ probability that the coin will come up heads, and a ½ probability that it will come up tails. In the first case I win $5, and in the second case I lose $3. So, in the long run, I'll win $5 about half the time, and lose $3 about half the time. So, in the long run, I should expect the amount that I win per coin flip to be the average of these two amounts — a win of $1.

We can express this by saying that the **expected utility** of taking the bet is $1. It seems that one should take this bet because the expected utility of doing so is greater than the expected utility of not taking the bet.

To calculate the expected utility of an action, we assign each outcome of the action a certain probability, thought of as a number between 0 and 1, and a certain value (in the above case, the relevant value is just the money won). In the case of each possible outcome, we then multiply its probability by its value; the expected utility of the action will then be the sum of these results.

Let's see how this looks by returning to our simple bet.

Courses of action	Heads	Tails	Expected utility
take the bet	\$5	-\$3	.5 * \$5 + .5 * (-\$3) = \$1
don't take the bet	\$0	\$0	.5 * \$0 + .5 * \$0 = \$0
	Probability = 0.5	Probability = 0.5	

The higher expected utility of taking the bet seems to explain why this would be the right move.

The higher expected utility of taking the bet seems to explain why this would be the right move.

Reflection on this sort of example seems to make the following principle about rational action seem quite plausible:

The rule of expected utility
It is always rational to pursue the course of action with the highest expected utility.

This suggests the following rule of belief:

Expected Utility → Belief
If believing P has a higher expected utility than not believing P, you should believe P.

Let's return to the passage discussed above.

> "It would be unwise of you, since you are obliged to play, not to risk your life to win three lives at a game in which there is an equal chance of winning and losing. But here there is an infinity of happy life to be won ... and what you are staking is finite. ... And thus, since you are obliged to play, you must be renouncing reason if you hoard your life rather than risk it for an infinite gain, which is just as likely to occur as a loss..."

Our question is: how might Pascal argue that believing in God has higher expected utility than nonbelief?

First, he emphasizes that "there is an equal chance of gain and loss" — an equal chance that God exists, and that God does not exist. This means that we should assign each a probability of 1/2.

Second, he says that in this case the amount to be won is infinite. We can represent this by saying that the utility of belief in God if God exists is ∞.

Let’s suppose, plausibly, that if we believe in God, and God does not exist, this involves some loss of utility. This loss will be finite — let’s symbolize it by the word “loss”.

One might represent these assumptions as follows:

Courses of action	God exists	God does not exist	Expected utility
believe	∞	loss	∞
don’t believe	0	0	0
	Probability = 0.5	Probability = 0.5	

Courses of action	God exists	God does not exist	Expected utility
believe	∞	loss	∞
don’t believe	0	0	0
	Probability = 0.5	Probability = 0.5	

So it looks as though the expected utility of believing in God is infinite, whereas the expected utility of nonbelief is 0. If the rule of expected utility is correct, it follows that it is rational to believe in God - and it is not a very close call.

Let’s look at a few objections to the idea that the above chart accurately represents our choice of whether or not to believe in God.

Courses of action	God exists	God does not exist	Expected utility
believe	∞	loss	∞
don’t believe	0	0	0
	Probability = 0.5	Probability = 0.5	

Objection 1: the probability that God exists is not 1/2, but some much smaller number -- say, 1/100.

Courses of action	God exists	God does not exist	Expected utility
believe	∞	loss	∞
don't believe	0	0	0
	Probability = 0.01	Probability = 0.99	

Objection 1: the probability that God exists is not 1/2, but some much smaller number -- say, 1/100.

This is a real strength of Pascal's argument: it does not depend on any assumptions about the probability that God exists other than the assumption that it is nonzero. In other words, he is only assuming that we don't know for sure that God does not exist, which seems to many people - including many atheists - to be a reasonable assumption.

Courses of action	God exists	God does not exist	Expected utility
believe	∞	loss	∞
don't believe	0	0	0
	Probability = m	Probability = n	

Objection 2: Pascal is assuming that, if God exists, there is a 100% chance that believers will get infinite reward.

To accommodate this possibility, we would have to add another column to our chart, to represent the two possibilities imagined. Let's call these possibilities "Rewarding God" and "No reward God", and let's suppose that each has a nonzero probability of being true.

Courses of action	Rewarding God exists	No reward God exists	God does not exist	Expected utility
believe	∞	0	loss	∞
don't believe	0	0	0	0
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Objection 2: Pascal is assuming that, if God exists, there is a 100% chance that believers will get infinite reward.

As this chart makes clear, adding this complication has no effect on the result. Pascal needn't assume that God will certainly reward all believers; he need only assume that there is a nonzero chance that God will reward all believers.

Courses of action	Rewarding God exists	No reward God exists	God does not exist	Expected utility
believe	∞	0	loss	∞
don’t believe	0	0	0	0
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Objection 3: God might give eternal reward to believers and nonbelievers alike.

Let’s call the hypothesis that God will give eternal reward to all “Generous God.”

Courses of action	Rewarding God exists	Generous God exists	God does not exist	Expected utility
believe	∞	∞	loss	∞
don't believe	0	∞	0	∞
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Objection 3: God might give eternal reward to believers and nonbelievers alike.

Setting aside the possibility of No reward God, which we have seen to be irrelevant, taking account of the possibility of Generous God has a striking effect on the expected utilities of belief and nonbelief.

Objection 3: God might give eternal reward to believers and nonbelievers alike.

Setting aside the possibility of No reward God, which we have seen to be irrelevant, taking account of the possibility of Generous God has a striking effect on the expected utilities of belief and nonbelief.

Now, it appears, belief and nonbelief have the same infinite expected utility, which undercuts Pascal's argument for the rationality of belief in God.

However, Pascal seems to have a reasonable reply to this objection. It seems that the objection turns on the fact that any probability times an infinite utility will yield an infinite expected value. And that means that any two actions which have some chance of bring about an infinite reward will have the same expected utility.

But this is extremely counterintuitive. Suppose we think of a pair of lotteries, EASY and HARD. Each lottery has an infinite payoff, but EASY has a 1/3 chance of winning, whereas HARD has a 1/1,000,000 chance of winning. What is the expected utility of EASY vs. HARD? Which would you be more rational to buy a ticket for?

But this is extremely counterintuitive. Suppose we think of a pair of lotteries, EASY and HARD. Each lottery has an infinite payoff, but EASY has a 1/3 chance of winning, whereas HARD has a 1/1,000,000 chance of winning. What is the expected utility of EASY vs. HARD? Which would you be more rational to buy a ticket for?

How might we modify our rule of expected utility to explain this case? Would this help Pascal respond to the case of Generous God?

A natural suggestion is to say something like this: if two actions each have infinite expected utility, then (supposing that neither action has a very high chance of leading to a very bad outcome) it is rational to go with the action that has the higher probability of leading to the infinite reward. This sort of supplement to the rule of expected utility explains why it is smarter to buy a ticket in EASY than in HARD; and it also helps Pascal solve the problem of Generous God, since the believer receives an infinite reward if either Generous God or Rewarding God exists, whereas the nonbeliever only gets a reward in the first of these cases.

Courses of action	Rewarding God exists	Generous God exists	God does not exist	Expected utility
believe	∞	∞	loss	∞
don't believe	0	∞	0	∞
	Pr. = m	Pr. = n	Pr. = 1-m-n	

If we adopt this modified rule — which says that in cases where two outcomes each have an infinite expected utility, one should choose the action more likely to lead to one of these outcomes —then this argues for belief in the case of Generous God, so long as $m \neq 0$.

Courses of action	Rewarding God exists	Generous God exists	God does not exist	Expected utility
believe	∞	∞	loss	∞
don't believe	0	∞	0	∞
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Objection 4: God might give eternal reward to just those who do not believe.

It is conceivable that God would do the opposite of rewarding belief, and instead would reward only disbelief. Call this hypothesis 'Anti-Wager God.'

Courses of action	Rewarding God exists	Anti-Wager God exists	God does not exist	Expected utility
believe	∞	0	loss	∞
don’t believe	0	∞	0	∞
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Objection 4: God might give eternal reward to just those who do not believe.

It is conceivable that God would do the opposite of rewarding belief, and instead would reward only disbelief. Call this hypothesis ‘Anti-Wager God.’

Courses of action	Rewarding God exists	Anti-Wager God exists	God does not exist	Expected utility
believe	∞	0	loss	∞
don't believe	0	∞	0	∞
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Objection 4: God might give eternal reward to just those who do not believe.

It is no longer obvious that belief has a higher chance of reward than nonbelief: we need an argument that Rewarding God is more likely to exist than Anti-Wager God. This shows that Pascal's argument can't be completely free of commitments to the probabilities of certain theological claims.

Courses of action	Rewarding God exists	Anti-Wager God exists	God does not exist	Expected utility
believe	∞	0	loss	∞
don’t believe	0	∞	0	∞
	Pr. = m	Pr. = n	Pr. = 1-m-n	

Note also that this scenario is analogous to the hypothesis that God rewards only the adherents of certain specific religions, only one of which can be believed.

So far we have focused on objections which try to show that expected utility calculations do not deliver the result that it is rational to believe that God exists.

I want now to consider three quite different lines of reply to Pascal's argument, which do not involve trying to find a flaw in his calculations.

In cases with infinite utility, the rule of expected utility fails.

It is impossible to form beliefs on the basis of expected utility calculations.

It is irrational to form beliefs on the basis of expected utility calculations.

In cases with infinite utility, the rule of expected utility fails.

Consider the following bet:

The St. Petersburg

I am going to flip a fair coin until it comes up heads. If the first time it comes up heads is on the 1st toss, I will give you \$2. If the first time it comes up heads is on the second toss, I will give you \$4. If the first time it comes up heads is on the 3rd toss, I will give you \$8. And in general, if the first time the coin comes up heads is on the nth toss, I will give you $\$2^n$.

The St. Petersburg

I am going to flip a fair coin until it comes up heads. If the first time it comes up heads is on the 1st toss, I will give you \$2. If the first time it comes up heads is on the second toss, I will give you \$4. If the first time it comes up heads is on the 3rd toss, I will give you \$8. And in general, if the first time the coin comes up heads is on the nth toss, I will give you $\$2^n$.

Would you pay \$2 to take this bet? How about \$4?

Suppose now I raise the price to \$10,000. Should you be willing to pay that amount to play the game once?

What is the expected utility of playing the game?

The St. Petersburg

I am going to flip a fair coin until it comes up heads. If the first time it comes up heads is on the 1st toss, I will give you $2. If the first time it comes up heads is on the second toss, I will give you $4. If the first time it comes up heads is on the 3rd toss, I will give you $8. And in general, if the first time the coin comes up heads is on the nth toss, I will give you $\$2^n$.

What is the expected utility of playing the game?

We can think about this using the following table:

Outcome	First heads is on toss #1	First heads is on toss #2	First heads is on toss #3	First heads is on toss #4	First heads is on toss #5	.....
Probability	$2	$4	$8	$16	$32	.....
Payoff	1/2	1/4	1/8	1/16	1/32	.....

Outcome	First heads is on toss #1	First heads is on toss #2	First heads is on toss #3	First heads is on toss #4	First heads is on toss #5	.....
Probability	\$2	\$4	\$8	\$16	\$32	.....
Payoff	1/2	1/4	1/8	1/16	1/32	.....

The expected utility of playing = the sum of probability x payoff for each of the infinitely many possible outcomes. So, the expected utility of playing equals the sum of the infinite series

$$1+1+1+1+1+\ 1+1+1+1+1+\ 1+1+1+1+1+1+1+1+1+1+\ldots\ldots$$

But it follows from this result, plus the rule of expected utility, that you would be rational to pay any finite amount of money to have the chance to play this game once. But this seems clearly mistaken. What is going on here?

Does this show that the rule of expected utility can lead us astray? If so, in what sorts of cases does this happen? Does this result depend essentially on their being infinitely many possible outcomes?

But it follows from this result, plus the rule of expected utility, that you would be rational to pay any finite amount of money to have the chance to play this game once. But this seems clearly mistaken. What is going on here?

Does this show that the rule of expected utility can lead us astray? If so, in what sorts of cases does this happen? Does this result depend essentially on their being infinitely many possible outcomes?

Suppose that we set an upper bound of 100 coin flips on the game, so that if you get to the 100th flip you get $\$2^{100}$ (a very large number) no matter how the coin comes up. Then the expected utility of playing will be \$100. Would you pay \$99 to play this game?

Most would say not. One possibility is that this is explained by a combination of risk aversion and decreasing marginal utility. Could these also play a role in the evaluation of Pascal's wager?

It is impossible to form beliefs on the basis of expected utility calculations.

Suppose that I offer you $5 to raise your arm. Could you do it?

But now suppose I offered you $5 to believe that you are not now sitting down. Can you do that (without standing up)?

Cases like this suggest that it is impossible to form beliefs on the basis of expected utility calculations.

It is impossible to form beliefs on the basis of expected utility calculations.

Pascal considered this objection, and gave the following response:

"I am so made that I cannot believe. What do you want me to do then?"

"At least get it into your head that, if you are unable to believe, it is because of your passions, since reason tells you to believe and yet you cannot do so. Concentrate then not on convincing yourself by multiplying proofs of God's existence, but by diminishing your passions."

What does he have in mind here?

It is irrational to form beliefs on the basis of expected utility calculations.

Let's now turn to our last line of objection to Pascal.

Pascal's argument, as we have reconstructed it, relies on the following principle.

Expected Utility → Belief
If believing P has a higher expected utility than not believing P, you should believe P.

This principle seems plausible. But so does this one:

Low Probability → No Belief
If you think that P has a very low probability of being true, you should not believe P.

Expected Utility → Belief
If believing P has a higher expected utility than not believing P, you should believe P.

Low Probability → No Belief
If you think that P has a very low probability of being true, you should not believe P.

Pascal's reasoning shows that these rules can come into conflict, because sometimes believing something which you think has a very low probability of being true can have a higher expected utility than not believing it.

One important question for those who find Pascal's argument convincing is: how could this second principle be false?