# *Recklessness, Timidity, and Pascal's Mugger*

*Ryan Doody*

*March 24, 2022*

## *Pascal's Mugging*

Pascal's Mugger attempts to convince (the fictional) Pascal to hand over the contents of his wallet ($\$n$) by promising his a much much much greater reward ($\$N$) if (and only if) he does.

Under what conditions might such a mugging work? And, if successful, does this show that there's something wrong with Pascal's decision-making?

If Pascal is (epistemically) rational, he should have some positive credence ($p > 0$) that the Mugger is telling the truth.

Pascal's Mugging versus Expected Value

**P1** If it's rational to maximize expected value, and value might be boundless, then it's rational to be *reckless*.

**P2** If you're reckless, you are vulnerable to fall victim to Pascal's Mugging.

**P3** It's not rational to fall victim to Pascal's Mugging.

**P4** Value might be boundless.

---

**C** It's not rational to maximize expected value.

You're *reckless* when you're willing to risk arbitrarily great gains at arbitrarily long odds so long as the upside is sufficiently enormous.

Is this a good argument? The truth of **P1** is meant to follow from the definitions of its terms, in particular:

> **Recklessness:** For any finite payoff $n$ (no matter how good), and for any positive probability $p$ (no matter how small), there's a finite payoff $N$ such that getting $N$ with probability $p$ is better than getting $n$ for sure.

$$N \cdot p \succ n$$

Some people call this property *Fanaticism*.

There are really two ideas underlying it. First, there is an *axiological* idea about how good things could be. Second, there is a *decision-theoretic* idea about how to weigh risks.

**P2** is supported by the thought that Pascal's Mugger can take $\$n$ from Pascal by offering to give him $\$N$ tomorrow so long as Pascal thinks there's a $p$ chance the Mugger is honest.

Can **P2** be resisted? If so, how?

If you deny **P3**, please see me after class for an exciting, once-in-a-lifetime, financial opportunity!

If you are a *Total Utilitarian*, you appear to be committed to **P4**—but so are many other views as well. In any case, on what grounds should we be confident that there is some upper limit on how good things could possibly be?

So, should we reject Expected Value Theory?

## *Timidity*

Expected Value Theory with an unbounded value-function is *reckless*. Other views, are *timid*:

> **Timidity:** There's a finite payoff $x$ and a positive probability $p$ such that getting $x$ with probability $p$ is no worse than getting any other finite payoff (no matter how good) with a slightly smaller probability.

Beckstead & Thomas argue that all views are either *reckless*, *timid*, or *non-transitive*. Here's the argument.

> **Devil at the Deathbed.** God gives you a ticket for a year spent in heaven. Then, the Devil appears and offers you a deal: trade your ticket for one that's good for 10 years in heaven, with probability 0.999. You accept. And then the devil offers you a new deal: trade that ticket for one that's good for 100 years in heaven, with probability $0.999^2$. You accept. And then ...
>
> Eventually, after 50,000 trades, you find yourself with a ticket for $10^{50,000}$ years in heaven, with probability $0.999^{50,000} < 1/10^{21}$.
>
> When you die, predictably, you spend 0 years in heaven.

It would be *timid* to not accept each of the deals. If you do, by *transitivity*, you should prefer the last ticket to the first. But, then, you are *reckless*.

Is it bad to be *timid*?

> **Strange Dependence on Distant Space and Time.**
>
> The problem is especially simple to understand if we suppose that timidity arises from the boundedness of value. How far away we are from the relevant baseline and how close we are to the upper bound can depend on how things are in remote regions of spacetime. In many situations, though, this would be deeply implausible. For example, if we had some way of greatly improving the lives of the roughly one billion people who now suffer from extreme global poverty, we could not plausibly claim that this would have only trivial value, on the grounds that a lot of great things had happened in the distant past or were now happening in other galaxies, so that we had already come close to the upper limit on how good improving lives could be. (Beckstead & Thomas, p. 15).

What's the objection? Do you find it plausible?

There are various ways to be timid. Here are three:

1. Your value-function is bounded.
2. You are risk averse (in a particular way).
3. You discount very small probabilities.

Are some of these ways of being *timid* better, or worse, than the others?