

Don't steer with guilt

12 JULY 2015

I've spoken at length about shifting guilt or dispelling guilt. What I haven't talked about, yet, is guilt itself.

So let's talk about guilt.

Guilt is one of those strange tools that works by *not* occurring. You place guilt on the branches of possibility that you don't want to happen, and then, if all goes well, those futures don't occur. Guilt is supposed to steer the future towards non-guilty futures; it's never supposed to be instantiated in reality.

Guilt works by the same mechanism as threats: imagine the tribesperson who precommits to breaking the legs of anyone who steals their food. If this precommitment works, then it never needs to be carried out: violence is a dangerous business, and the tribesperson would much rather that they never need to break legs at all. The threat is something that the tribesperson places on possibilities that they disprefer, in attempts to ensure that they never come to be.

Imagine, by contrast, the tribesperson who threatens to breaking the legs of anyone who looks at them funny: they might find themselves attempting violence every single day, and this likely makes their life unpleasant, to say the least. In this case, I would argue that they're using their threats poorly. I would say that, if you keep finding yourself carry-

ing out a threat, then you really need to consider whether or not your threats are really capable of steering the future in the way you hoped.

Guilt is the same way: *if you find yourself regularly experiencing guilt, then you're using guilt incorrectly.*

Guilt works only when you wield it in such a way that it *doesn't happen.*

Guilt is costly when deployed. Once activated, it's usually strongly demotivating, and can easily lead to failure spirals or vicious cycles of depression.

As far as I can tell, the way that guilt-motivated people tend to operate is by working fervently in attempts to avoid the scourge of guilt. This may be effective when it works, but as soon as it starts to fail, the failure often cascades into a full-blown failure spiral (you're guilty that you're not working, which makes you feel bad, which makes it hard to work, which makes you guiltier, which you feel worse, which makes it harder to work, ...). As a result, guilt motivation often results in a boom/bust productivity/depression cycle that, as far as I can tell, results in people feeling quite bad about themselves and being much less effective than they would be if they could maintain a steady pace.

Some might argue that the boom is worth the bust, that the productivity is worth the depression. This seems straight up false to me (and I have some relevant experience): the fran-

tic productivity fueled by fear of guilt doesn't seem more effective (and often seems *less* effective) than intrinsically motivated productivity, and that's *before* we count the losses from periodic failure spirals. As far as I can tell, intrinsic motivation is just straight up more effective.

(This is something you have to accept before I can help you remove your guilt: it's much harder to remove guilt if you don't want to.)

Guilt is very costly when activated, so if it's getting activated regularly, then you're placing it on the wrong branches of possibility.

You might protest, "but then what do I *do* in the unsatisfying branches of reality? I need to find *some* way to prevent me from chasing short-term satisfaction at the expense of long-term benefits." If you regularly find yourself binging netflix TV shows, and you would rather not find yourself regularly binging netflix TV shows, then shouldn't you feel guilty whenever you do?

No! If the situation occurs regularly, then guilt is not the tool to use! You're welcome to feel guilty if you ever kidnap a baby or punch a homeless person, and you can tell that the guilt is working in those cases because you *never do those things*. But if you repeatedly find yourself in a situation that

you disprefer, then guilt is just not the tool to use. That's not where it's useful.

If you want to figure out how to avoid a certain recurring situation, then there's a different tool that *is* appropriate, that's much more effective at figuring out how to steer the future towards better places: Science!

When you find yourself binging netflix, don't heap loads of guilt on yourself post-binge. That sort of thing clearly doesn't prevent the binge. Instead, say to yourself, "huh, I appear to netflix-binge under certain conditions, despite the fact that I'd rather not. I wonder which conditions, specifically, led to that binge! What were the triggers? How could they have been avoided? What methods might help me avoid binging in the future?"

And then treat it like an experiment! Write up your hypotheses. Experiment with many different ways to fix your glitches. Write postmortems when you fail. If you attempt a fix and then find yourself binging *again*, then don't heap loads of guilt on yourself! *That still doesn't help*. Instead, say "Aha! So *that* attempted fix didn't work. I wonder if I can figure out why?" Cross a hypothesis or two off your list. Refine your models. Expand your hypothesis space. Gather more data.

Do science to it.

Don't bemoan individual failures. That's finite-task thinking. Instead, acknowledge that there's an unlimited number of changes you'd like to make to your behavior, and that some of them are more important than others, and that some of them are more costly than others, and that they all take time to fix. See the infinite stream of self-improvement that lies before you, add it to all the other streams you're optimizing, and then simply navigate the streams as quickly as you are able.

Don't feel terrible whenever you do something you wish you hadn't! That is a poor mechanism by which to steer the future. Instead, when you do something you wish you hadn't, identify the *pattern of behavior* that led to this, and add addressing *that* to your todo list. Then weigh the time you're losing against the time it would take to change the pattern, and weigh that against the other priorities that are vying for your attention, and then do what needs doing.

Sometimes you'll ignore a pattern of failure. Maybe the failures are relatively cheap and the pattern is hard to change, and fixing the pattern simply isn't worth your attention. In this case, when the failure occurs, there is no need to feel guilty: the failures are the price you pay for time spent not fixing them. You can't simply teleport to a new pattern of behavior, and so if you lack the time to change the pattern, then the occasional failure is a fair price. Trust yourself to fix the pattern if the costs ever get too high, trust yourself to understand that investing in yourself is important, and if fixing the pattern *still* isn't at the top of your todo list, then

don't worry about the individual failures. You have bigger things on your plate.

Other times, you'll decide that the pattern needs changing. Five minutes per day is thirty hours per year, and investing in yourself pays dividends. In this case, treat addressing the pattern of failure like a science project. Every new individual failure is data point about what doesn't work. Every avoided failure is a data point about what does. Heaping guilt on yourself whenever you hit a new failure would be nonsense — fixing the *pattern* is a science experiment, and individual successes or failures are your data points.

Most people use their individual failures as a signal to themselves that it's time to feel terrible. It is much more effective, I think, to use your individual failures as a chance to update your tactics.

This, in my experience, is the head-on cure for guilt: Don't treat the individual failures like a burden; treat changing the pattern like a science experiment.