Stop trying to try and try

22 NOVEMBER 2015

Imagine a graduate student of mathematics as they interact with a professor, attempting to understand something in the professor's area of expertise. They're working hard to wrap their head around the basic formalism. They're in "learning mode" — they're a student in the presence of a master, expected to try to understand the math but not necessarily expected to succeed. Even if they're doing quite well, they're still reminded of how math is big and they are small; they encounter wide swaths of knowledge that they do not yet have, and often feel humbled. They use their tools tentatively, aware that they may be using them inappropriately, and wonder when they'll become a master.

Now imagine the same student tutoring an undergraduate in linear algebra, a topic they know quite well. Now they're in "teaching mode." Math is still large; the graduate student is still small; but the context is very different. The focus is no longer drawn repeatedly to all the things they don't know yet — but it's not drawn to all the things they *do* know, either. The focus simply isn't on them, or their abilities. It's on the undergrad. The grad student, in the back of their mind is not thinking "wow math is so large I don't know enough yet I'm not sure I'll ever know enough", and they're *also* not thinking "wow I know so much this is great!" — they're

thinking about how to help the undergrad understand a complex concept.

I think that many people who are in learning mode expect that mastery feels like learning mode, except that instead of feeling like they know very little, they feel like they know quite a bit. By contrast, I think mastery looks much more like teaching mode — it looks like someone operating in a context where their knowledge and their skills are not the focus, but are just unconscious assumptions in the background.

Consider the grad student in teaching mode. Their approach to answering questions in teaching mode is very different than their approach in learning mode. That's not because all the questions they encounter in teaching-mode are simple — if you've ever been a tutor you know that tutors are commonly asked questions they can't answer in the moment. Rather, they approach questions differently because context is different. When the professor asks them questions, they're Expected To Do Their Best; when the undergrad asks them questions, they're just expected to answer.

In the first case, they're expected to try; in the second case, they're assumed capable, an assumption that fades into the background.

I describe this model because I think there is an analog of these two modes when it comes to "trying" to achieve any task — and today, I'm going to talk about trying.

My advice is simple: notice when you're expected to try, and consider reframing. It's much harder to solve a problem when you're Expected To Do Your Best than it is to solve a problem when you're immersed in various subtasks, with the assumption that you're going to solve the problem buried implicitly and unconsciously in the context.

For example, consider exercise. Many people find it much easier to exercise in a context where the exercise is in the background rather than the foreground. Imagine someone who plays recreational soccer, sprinting up and down the soccer field up till the brink of exhaustion. Now imagine them not playing soccer, but just trying to sprint up and down the field up to the brink of exhaustion. They probably push themselves a lot less in the latter case. If "sprint up and down the field a lot" is the main goal, then at each possible stopping point, part of them starts trying to convince the rest that they've exercised enough for the day, and they must spend willpower to continue. In a soccer match, by contrast, the focus is elsewhere. They aren't constantly pinging themselves with explanations of how they've done enough sprinting for today. They aren't generating reasons why it's OK to stop here. They're trying to *score a goal.* Getting exercise is a background assumption, not a conscious choice.

Switching contexts such that your actual goal is in the background rather than the foreground — such that pursuing it

is not a conscious choice that you need to reaffirm every time you find a stopping point — is a powerful tool.

This is not novel advice, of course, but it is perhaps a generalization over a few different common types of advice. As another example, consider two people trying to become friends on purpose (perhaps for romantic reasons). I conjecture that it's much harder for people to become friends on purpose than to become friends accidentally while pursuing some other endeavor.

If they're trying to become friends on purpose, then they're constantly asking themselves, "are we friends yet?", and like the grad student asking themselves "do I understand all of mathematics yet?", the answer will never be an unresounding "yes". They would do better to switch to a context where they're not constantly checking whether they're friends yet, and are instead just *being friends*.

This model suggests that it's much more effective to alter the context such that neither party is regularly checking the depth of the friendship, but such that a strengthening bond is the implicit background assumption. (This suggests one reason why online dating feels more socially awkward than going on a date with someone you met in some other context.)

For a third (somewhat silly) example, imagine that I woke up one morning and said "I'll try to run MIRI well today." (MIRI, the Machine Intelligence Research Institute, is an or-

ganization I run.) If I did this, I'd be in trouble. How does one run a research institute? What would my next actions be? Things that seem plausibly like what people-who-run-institutes-well would do? Things that seem defensible to the board of directors? I have no idea how to "try to run MIRI."

Now imagine instead that I woke up and said "I'm going to glance at my MIRI priority list, update it if today happens to be Monday, and then identify MIRI's biggest bottleneck and work on it directly." Now I'm in business, and might do something useful with my day.

Notice the difference. In the second case, I'm not asking myself whether I can run a research institute. I'm not asking myself *how* to run a research institute (though "study the strategies of people who ran other successful institutes" does occasionally get to the top of my priority list). I'm assuming myself capable — not consciously, but as a background assumption. I'm not assuming *success* — either I can run a research institute or I can't, the jury's still out on that one — but my capability is not the focus of my attention. I fret about much more practical things, like the tone to strike in a fundraiser announcement post, or how to prioritize paper-writing versus novel research. I'm never "trying to run MIRI;" I'm just working on the next top-priority task.

This, I think, is one of the main distinctions between "trying to try" and "actually trying".

Trying to try to run MIRI would *feel like* just trying to run MIRI — it would feel like thinking about what it takes to run an institute and reading books about running institutes and worrying whether the board of directors thought I was doing a good job and so on. From the inside, I'd probably think I was trying very hard to actually run an institute.

Actually trying to run MIRI feels very different from the inside. It doesn't feel like trying to make an institute run, it feels like trying to get all the most important emails handled while not letting administrative duties suck up my day. It feels like struggling to prioritize three important tasks that can't all be done. Actually trying to run MIRI does not feel like trying to run MIRI, it feels like a never-ending stream of smaller tasks.

I think many people imagine the difference between trying to try and actually trying involves something like Additional Effort or Additional Willpower. It's easy to imagine someone trying to try to (say) cure aging. Maybe they flounder around a bit and talk about how they want to join a biology startup, or start a biology startup, or get a biology degree, all while really deeply wanting to find some way to cure aging. It's also easy to imagine that the person "actually trying" to cure aging is doing something similar, but with more determination and a bit of pixie dust that makes things work out. The actually-tryer does the same things, but for them, the

startup works through dint of sheer willpower; or they get a biology degree while winning so many accolades that they get to set up their own laboratory.

This isn't how I imagine "actually trying." It's not trying-to-try with extra gusto. Actually trying looks like solving small subproblems, with the more ambitious target no longer the focus of attention, but rather a background task. Actually trying to cure aging doesn't look like a person getting a biology degree with *especially grim determination*, it looks like Aubrey de Grey wading through a mountain of mundane tasks while scraping together enough money to keep SENS running.

(SENS is currently fundraising, by the way.)

If you want to solve hard problems, stop trying to solve the hard problem directly. Change the context such that that's a background assumption: all your actions are going to be pointed roughly in the direction of solving-the-problem; what next? What's the next thing that needs doing? Work on *that*.

This is perhaps simple advice, but I myself have found it useful in the past. Many years ago, when I was in high school, a friend of mine came back from college having joined a fencing team. He wanted to show me some of the basics, so he tossed me a sabre, and we had at each other. We crossed swords a few times, and he said something

along the lines of "Nate, the goal isn't to hit my sword, the goal is to hit *me*."

It's an obvious thought, a simple thought, and a thought I had failed to think. After that, I wasn't trying to fence, I was trying to *hit him*.

Or consider the scene in The Matrix where Morpheus tells Neo "Come on, stop trying to hit me and hit me!" — at which point Neo's blows grow more intense, until he gets a fist past Morpheus' defenses. I suspect that many people watching that scene imagine Neo turning on the "try harder," pouring more effort into his punches and harnessing his frustration. When I watch the scene, I imagine a little bit of that, but mostly I imagine a similar mental shift to my "don't bang swords together; strike the enemy" mental shift — I imagine Neo had mostly been throwing out a bunch of martial arts moves that had recently been uploaded into his brain, in attempts to see if any of them worked against Morpheus, and that when Morpheus said "stop trying to hit me and hit me" Neo thought "oh yeah, I'm not supposed to be deploying martial arts moves and monitoring whether I'm fighting well enough, I'm supposed to be *hitting Morpheus*," and that his brain shifted from the "expected to try" gear to the "competence assumed" gear.

I think many people solve problems more effectively in the "competence assumed" gear," when they're not fretting about whether they can solve problems because they're too busy fretting about very specific actionable subproblems.

So if you want to tackle big problems, my advice is this: If you ever find yourself saying "I'm currently trying to solve [problem]", be wary. This is doubly true if you're Expected To Do Your Best.

If you find yourself saying "well I'm trying to solve aging, but it's a big problem, so I'll likely fail," then stop in your tracks. Not because of the underconfidence — aging *is* a big problem and you *will* likely fail to solve it — but because you're sprinting up and down the field when you'd be better off playing a game of soccer.

If you approach a big problem with Intent To Try, then at every plausible stopping point part of you will be trying to convince you that you've done enough. And thus, at every plausible stopping point, you'll need to spend willpower to continue. Find a soccer game instead — some way to focus your attention on useful object-level tasks, with the pursuit of the important goal turned into an implicit unconscious background assumption so deeply ingrained in your plan that you can hardly see it any more.

As for how you make or find the soccer games, that's a discussion for another day. For now, my generic suggestion is to (a) generalize from the above examples and (b) imagine someone who's "playing soccer" with respect to your task or problem, and ask yourself what they might be doing. The

key is to make the pursuit of your goal implicit, and spend your focus on the subproblems.