

Rationally Speaking #158: Dr. George Ainslie on, “Negotiating with your future self”

Julia: Welcome to Rationally Speaking, the podcast where we explore the borderlands between reason and nonsense. I'm your host, Julia Galef, and with me is today's guest, Dr. George Ainslie.

He is a research psychiatrist at the Veteran Affairs Medical Center, and also a professor of economics at the University of Cape Town, South Africa. He has published several books, and the focus of his research and of those books has been human motivation and willpower, and specifically, this struggle that all of us face from time to time, or on a regular basis, struggle between how we wish we behaved on the one hand, and how we actually behave on the other. That's going to be the jumping off point for today's episode.

George, welcome to the show.

Ainslie: I'm glad to be here with you.

Julia: Just to orient our listeners, I presume that everyone listening to this show can identify with this phenomenon in which, for example, it's Sunday, and you're thinking about what you're going to do tomorrow, and you decide okay, tomorrow, after work, I'm going to go straight to the gym, and yeah, maybe it'll be a little unpleasant, but it's totally worth it, because then I'll be getting in shape, and I'll be healthy and getting fit and so on.

Then, Monday rolls around, it's the end of the day, and suddenly, it doesn't feel so worth it anymore to go to the gym. Suddenly, it feels much more worth it to go home and kick back with some Netflix. What's happened here is there's this change in your preferences, where your preferences on Monday about gym versus Netflix are different from what your preferences had been on Sunday about that Monday choice. Also, much of the time, they're different from your preferences on Tuesday about what you wish you had chosen on Monday.

Even if you're one of those people who happens to love exercising, which I've never fully understood, I'm sure you can substitute your own example, right? Where you're choosing between working hard on this paper that's not due just yet, or you're choosing between spending your money now on some tempting impulse buy versus saving it for later when you might need it even more, etc. We just experience these flip-flopping of our preferences in which we behave differently than we wish we would, than our past self wished we would, and we behave differently than our future self will end up wishing we had chosen.

Ainslie: Working on the paper gets me where it hurts.

Julia: Right, there you go, there's something for everyone. This has been a puzzle that has confounded economists and other theorists, philosophers for generations. Let's start there. George, what is going on there?

Ainslie: First of all, I think we're dealing with a problem that's new to evolution. We sort of assume that we've come equipped by evolution with rationality and the ability to

look into the future, and actually no animals aside from us have any real capacity to plan for the future, even our closest relatives, the great apes, the chimpanzees, plan ahead no more than an hour or two. It's just not something that most anybody else is wired to do. Other animals, if they have to do something far-sighted, will have an instinct for it. They will have something that makes building a dam fun or hording nuts fun, and so they're in a different situation.

We, now that we're very imaginative and can think about the future, we have to put something together sort of by ourselves. It's something that we have kluged together, and not something that's really one of our inborn faculties.

Julia: Go on.

Ainslie: The problem with it is as, I think my contribution to this, is both animals, lower animals, non-human animals and we, are naturally equipped to value the future inversely proportioned to its delay. And that's not how a banker thinks, but it's how pigeons think and rats think and children think.

If you have people say spontaneously what they prefer, if you want a simple way to demonstrate that on yourself, I do this with audiences all the time, I ask, pretend it's a game show, and would you rather have \$100 today or \$200 three years from now, raise your hands.

A slight majority, or about equal, depending on the audience, say they'd rather have the \$100 now. Then I say well, let's say it's a cheaper game show, and you can't have the \$100 until six years from now or the \$200 until nine years from now, pick which one you want. And most everybody picks the \$200 nine years from now.

Of course, that's the same choice, seen at six years greater difference. And that comparison is really the basic experiment in intertemporal bargaining, that you start out with this demonstrable temporary preference, not based on your gut. It's not based on hot thinking or any of those explanations that have come up lately for it.

It's simply based on a, you start out with a preference for the smaller, earlier reward when it's close. That's something that we have to start with and deal with, and it never goes away. That's inborn.

Julia: Right. I think people often get confused between these two different ways that our preferences about the future can be structured. Where, on the one hand, people might say look, yes, we value the future less than the present, all else equal, but there is a rational justification for that, or there's multiple possible rational justifications for that, right? Like I may not be around six years from now, so if all is equal, I might as well eat my dessert now. Or, maybe my preferences in the future might be different, and so this good that is tempting me now might not even be very beneficial to me in six years, so I might as well get benefit from it now when I can.

Or, you know, interest rates, right? Like the same amount of money that I get now will accumulate interest over time and be worth more in 10 years than if you had

given me that amount of money in 10 years. That kind of discounting the future, I think it's more easy to reconcile that with a kind of normative rationality.

But the thing that you're describing is subtly but importantly different, right?

Ainslie: Yeah. Psychologists have now explored those objections in some detail, and it doesn't depend on something being right now, you can offset the whole thing by a month and you get the same result. \$100 in a month versus \$200 in three years and a month. You can deal with something besides money, you can control for the interest rate to change if possible. All of these things have been sort of put into play, and they don't make any difference.

Julia: Go on.

Ainslie: The phenomenon is that you're wired to change your mind. That what looks better in the more or less immediate future, a month from now, is not the same thing that looks good in three years. That is where we start.

Rationality, you have to sort of define it. Because here we've got the animal, it wants to do what it wants to do, maybe that's rational?

But you wind up becoming a victim of somebody else that is better at planning consistently. You get what's called money pumped. Somebody who can discount the way a banker does consistently, so that the choices stay in the same proportion to each other, he will buy your overcoat in the spring and sell it back to you in the fall, and wind up richer than you are, because he's been able to exploit your temporary preference for the smaller-sooner [reward], what I call hyperbolic discounting, by the way. I don't want to go into hyperbola and all that on the air, but I may use the term as a shortcut for what I'm talking about.

Julia: Right, great. Just to stick with the terminology for a minute, to contrast that with the simple preference for money now as opposed to money later, that would be exponential discounting, right?

Ainslie: Yeah, to be consistent, so that you know you may not value an overcoat in the fall as much as you do now, but to the relative value of the overcoats at the time offsets, stays the same over time so you can choose now or choose a year from now. Given the same time offsets, you'll make the same choice. Exponential discounting, what the bankers use, is the only formula that gives you that, that gives you consistent choice over time, as long as nothing else has changed.

Julia: Right, where as hyperbolic discounting is you're being inconsistent with yourself, essentially.

Ainslie: Yeah, you're overvaluing the immediate.

Julia: Right, great. It almost seems unnecessary to ask about empirical evidence for hyperbolic discounting, just because I experience it literally every day of my life, and I think many other people do too -- but nevertheless, I'd be interested to hear what

kinds of experiments have demonstrated its existence, especially if we have any evidence of hyperbolic discounting in animals, because I imagine that would bear on what kind of origin story we want to tell about it.

Ainslie: Sure. Obviously, you need empirical evidence, because the two are similar to each other. There's still some argument about whether we really have hyperbolic discounting underneath, or whether there's some way to construct it out of sets of exponential discounting estimates at different times. Or the favorite one is -- well, there's what's called beta delta discounting.

The economists really hate hyperbolic discounting because it means that what you choose at one moment doesn't determine what you're going to predict at the next moment, and that throws a lot of their mathematics [into chaos].

Their favorite objection is you get emotion or appetite aroused, and so you discount exponentially, fine, up until you've got an appetite, and then you've got a peak in your appetite, and that throws you into a temporary preference for this thing you've got an appetite for. We can show that happening, but it doesn't account for all of this, like the \$100 in a month versus three years and a month.

The basic experiment showing it is just like the one I gave you for the \$100. You do it with a lot of people. There were some excellent studies at first that even did it with different ages, young people and middle aged people and older people, and found that everybody's preference for that kind of money was hyperbolic, but the curve got a lot shallower as people got older.

Julia: They got better at overcoming the urge to-

Ainslie: Delay.

Julia: Yeah.

Ainslie: They had less of an urge. They got more patient, as they reported it. The problem is, of course, people are reporting something after they've done all their own self controlled thinking for years and years, and so that sort of distorts where you start from. Animals, I think your point is right, if it happens in animals, then it's really basic. I actually started out with animals. I started out with pigeons.

Julia: How did you do that?

Ainslie: ... I was with Vicker and Stein in Harvard Laboratories of experimental psychology sort of part-time. I was a medical student. I was doing this in my abundant spare time, which you can imagine wasn't that abundant.

I had these pigeons, and the first thing I did was see if they would do the Ulysses-facing-the-sirens experiment. I gave them a repeated trial. There was only one key in front of them. If they pecked the red key during an interval, let's say, they got one second's access to grain. If they didn't peck it, if they withheld it for the whole interval, they got three seconds' access to grain.

Ainslie: If they withhold the peck for three seconds, they get three seconds worth of grain, but if they're not withholding, they just get a little bit of grain, one second right away.

Ten seconds before they're going to get this choice, I give them a different key color, a green key. If they peck that key, they don't get the red key. They wind up having to wait and get the three seconds.

Some of them learn to peck the green key. This is, if you know your mythology, is Ulysses tying himself to the mast before he sails past the sirens.

Julia: Right. He knows he's going to experience temptation and he wants to get rid of the possibility that he might give in to that temptation.

Ainslie: Yeah. It would be nice to think that. I don't think the pigeon is that smart. I think we just have a set of trials rewarded from the viewpoint of ten seconds on, when he gets the green key, there's "what's it going to be like in ten seconds if I press the green key, or if I don't press the green key?" It's just more rewarding from that vantage point to press the green key. It doesn't require any insight at all.

Julia: Were the pigeons pressing the red key? Before you gave them access to the green key, would they press the red key all the time?

Ainslie: Regularly. Ninety-five percent of the time, yeah.

Julia: Okay.

Ainslie: I could make it shorter and shorter and get them to do it less, but it was ... I wanted them to do it most of the time. Only some of the pigeons learned to press the green key. Some of them never learn that.

Julia: Interesting.

Ainslie: They were at the extreme edge of evolution as far as time is perceived in the pigeon.

Julia: Right.

Ainslie: The point is they got it, and then that started the evidence. I did another experiment with my mentor, Dick Herrnstein, in which we did a regular amount versus delay with pigeons, but as if it were the one year versus three-year thing.

Julia: Yeah.

Ainslie: They got the phenomenon, that if you gave them those two seconds versus six seconds of grain, and was now versus in four seconds. If they chose when it was right away, they all peck right away. If right away they had twelve seconds delayed for the small and sixteen seconds for the large, same choice seeing it twelve seconds time difference... They always pick the larger [reward] later.

Julia: Wow.

Ainslie: You can see a nice curve going through the different choices between that. That was back in 1981, and we've had these data for quite a while, and people have run them back and forth and up and down since then. You find bankers tend not to report that.

Somebody I was having a fight with in print, a guy named Glenn Harrison, who was an excellent econometrician at Georgia State, had run similar experiments with Danes picked from a population sample. A random selection of all Danes, and had them come to this place. He gave them quite a bit of money. A hundred dollars, or even two or three hundred dollars, at an earlier time versus even more at a later time. He found they discounted just like a banker did. Perfect exponential discounting.

I objected, "Of course. We know we can do that. We've seen bankers do it." It's a skill one can learn. The point is, how do we get there?

We got into a dialogue, and now he and I have been doing experiments to try to find how people get there. We just completed an experiment at Georgia State, in which students -- who are a little less likely to act like bankers -- are given this choice of not quite as much money now, versus more sometime later in the term, a month or two months.

We see some of them picking smaller sooner, and some picking larger later. But if you give them a group of choices to make all at once, they're more apt to pick the group of larger later choices than the single choice of larger later.

We're getting much more specific about that, much more econometric about that.

Julia: Let's delve into what work is being done by the grouping of choices there.

Ainslie: You are then not choosing just the one. Your idea of self-control is kicking in, and saying, "This is not a one-time only thing. This is something I'm going to repeat." Our argument would be, we learn that by the time we're in college. We learn the difference between "okay, just this once" and "this is going to be my salary or my reward on a regular basis." That's kicking in.

Of course the interesting question is how we learn that, and I'd like to get into that at some point, because that's another concept I'm proud of and which is much harder to do research on.

Julia: Let's get into it now then.

Ainslie: Okay.

That is intertemporal bargaining.

If you realize fairly early on that I'm going to change my choice like that, and I do it

fairly often, how do you commit yourself to pick later what you prefer earlier on?

Of course there are physical ways to do it. There are ways that people offer you -- Christmas clubs, a famous way of having people give a certain amount every week, and they'll have enough to spend at Christmas. Even though it doesn't give any interest at all, it was a very popular thing, and that's why economists started to puzzle about that.

Julia: About why people-

Ainslie: About why people would get this rather inferior savings account, when all you got was lack of access to your own money.

Julia: Right.

Ainslie: The Ulysses model has caught on, and there are even websites where you can sign up for that. You can deposit some money, and if you don't do what you said you'd do, they'll send it to your most hated organization. The point is, that's often not available or not desirable. What do you do if you don't have that?

What you do, it's really elementary put in one way. Let's say you're on a diet and you're faced with a cookie. How do you not eat the cookie? You say, "If I eat this now, why won't I eat it next time?" In other words, you take this as a test case. This is just how it's going to be, "I'll have cookies and ice cream and whatever all the time, and unless I make a rule, unless I perceive this in some way that this is a test case for lots of similar cases in the future, I'm going to go on taking the cookie." If you perceive it like that, "If I eat the cookie, there goes the credibility of my diet," then people succeed.

This is something we all know. This is intuitively obvious, but it hasn't been put in mathematical form until recently. It hasn't really been seen as a game theoretic solution until recently. When you think about it, that's about the only way that you can all by yourself without being able to do some external thing that commits you, make sure you do what you're now resolved to do.

Julia: Right.

Ainslie: That's willpower, I would say, is the short answer.

Julia: Basically, with this concept of inter-temporal bargaining, you're envisioning a person, not as single entity consistent with himself over time, but as a series of people who inhabit the same body at different points in time?

Ainslie: It's easier to make this sound more complicated than it is. I'm not saying that these are dissociated people somehow, I'm just saying that you're going to want something different in the future, which is a familiar phenomenon.

Julia: Sure.

Ainslie: You only partly agree with what that person is going to want.

This is what the bargaining theorists call limited warfare. You both share the same goal. You don't want to get fat. As the cookie gets closer, you have some shorter term goals that diverge. Your long term interest is to not eat the cookie and your short term interest is to eat the cookie.

Each of their interests are just ways they learn to get rewards and they each have learned techniques -- so the short term interest is to, let's say, distract your attention. Or the long term interest is using this personal rule to say, "Well, this is an exception. This is my birthday. Somebody has offered me the cookie and it would be rude to refuse," or something like that, that means this is a one time event and it doesn't affect the credibility of my diet.

Julia: Right. The credibility is also a good that both I and my future self care about, right?

Ainslie: Yeah.

Julia: It's in both of our interests to be able to credibly plan for something and know that we'll follow through.

Ainslie: Yeah. Nobody wants to go on and gorge and get fat, so that's the limited warfare aspect of it. Again, there's been a lot of research about that by economists, mostly. This fits willpower simply into an inter-temporal bargaining model. It's also very much like the prisoner's dilemma.

Julia: This has come up in our podcast several times before, but I suspect there's still some listeners who either haven't heard it or don't remember it.

Do you want to explain it briefly?

Ainslie: Sure. Well, the one shot prisoner's dilemma, which has puzzled people so much, is that in its simplest form, two thieves get caught with burglar tools and the policeman interviews them separately and says, "Do you confess or not?" If neither one of them confesses, they can only be charged with having burglar tools, which is a lesser offense. Their interest is, jointly, to not confess. However, if one confesses and the other doesn't, then it's going to be to the advantage of the one who confesses and bad for the one who doesn't. The game theoretic solution is they should each confess, because the other guy has the same incentive to confess that you do, even though, if you trusted each other, you could not confess.

That's the dilemma, and there's been a lot of work on how to solve it in the one shot deal, where it's only one time you've been arrested.

The time that's relevant to self control is the repeated prisoner's dilemma. This is going to happen again and again. Of course, in that case, you have this working relationship with your burglary partner. You both know, well, don't confess. That's the way to do it, because if you did, you know the other guy would confess next time. If you don't confess this time, you can keep the trust going that neither one of

you will confess.

In a diet situation, defecting is to eat the cookie this time and cooperating is to not eat the cookie, on the assumption that that will help later selves to not eat the cookie.

The only thing different from the prisoner's dilemma is the later self can't literally get revenge on the earlier self.

Julia: Right. That's what I was going to say, yeah.

Ainslie: The fact that all selves will lose the prospect of keeping cooperating will have the same motivational impact. It would be as if you were being punished. The math works out the same.

Julia: Oh, right. So what's being punished is not me, the past self directly, but harm is being done to things that the past self cares about.

Ainslie: Yeah.

Julia: Right. Interesting. Do we have a model of why willpower sometimes succeeds and sometimes fail?

Ainslie: Sure. I mean, there's a couple of very common solutions. One is, of course, that the short term interest gets very good at proposing exceptions.

Julia: Right.

Ainslie: There's a wonderful passage in William James, which I just learned he copied from Haslip, I think. Anyway, without attribution in William James' work, he says that an old drunk told him that he would rather cross in front of a firing cannon to get alcohol than not to do it. He would find any excuse.

James goes on to list the possible excuses. That it's cold and he needs something to warm himself up, or it's just this once because it's a holiday, or even this will give me the resolution later. If I go on a real bad drink and get sick, it will give me the resolution to not do it in the future.

After a while, you don't have any credibility at all, your excuses are hollow. You can't use it.

That's really typical of a really serious addiction like alcoholism. I think this is why Alcoholics Anonymous says what you have to say is, "I'm helpless against alcohol." That doesn't mean that you really are helpless. It's really your will you're building up. But if you think of it as your will, you start coming up with these ways you bargain with your will and that's going to get you down the slippery slope.

If you say, "I'm helpless," it essentially says, "If I started acting like this is a regular case of willpower, I'm going to lose. I've got to treat this as something I just have to

hope for. All I have is my hope." Then if I take a drink, what it's going to do is spoil my hope, and that's the only kind of negotiation I can do. If you do it that way, then you get as much willpower as is possible.

Julia: What always trips me up about this is: I'm with you that having this good of my own credibility, my own ability to trust myself or follow through on plans is what ... When you put that at stake, then there's this incentive for my current self to cooperate with future selves, because I don't want to harm my credibility.

But there's this kind of self-fulfilling nature, or self-fulfilling problem, where -- what if I do slip up? Then the consequence of that is to say, "Well, I guess I'm not credible anymore. Now this is terrible. Now my prospects for the future are bleak." But if I say, "Well, okay, look. I screwed up that one time, but that doesn't mean my credibility is irrevocably damaged," then my prospects for the future aren't so bleak.

But if I know in advance that I'm going to be willing to let myself off the hook, then those stakes don't really work to keep me in line. I have an incentive, basically, to think of my credibility as being very at risk and very important -- until I break it, at which point I have an incentive to decide that actually my credibility wasn't really at stake at all and I get more shots.

Ainslie: Well, you've put your finger right on the hard problem, and it's the one that really differentiates the person who is skillful at using personal rules, or the therapist who is skillful at helping a person use personal rules, from somebody who isn't. You've got to do a sort of dance between those two.

People who deal with addiction are aptly familiar with this. People have pledged everything to stay sober, and then they don't stay sober. Then you have what's called the abstinence violation effect. You can see lots of articles on the abstinence violation effect. Suddenly, they all go to seed and back to skid row and takes a long time to get back.

AA says, "Well, that's going to happen. You're helpless. What did you expect? Nevertheless, you can get up and try it again. Saying that you're helpless again just shows that you're helpless." It doesn't give you permission to go and drink, but it gives you some kind of an expectation that all is not lost if you do drink. In fact, you can expect it at some point that you will drink. That seems to be a relatively good mini-max position to steer into.

Then you have lots of people who have gotten into this situation and don't know how to get out of it.

I think a lot of religious solutions are not about the story of man and the devil and things like that, but they are about how to sin and recover, not sin more, do something to help you sin less. They essentially name your stake in the bets as some supernatural being. I will swear by Saint so-and-so, or I will promise God, or something like that. If you break your word then Saint so-and-so is mad at you, or God is more distant from you.

Yet there are some ways that the religion puts in that you can get back in the good graces. You can do penance, you can just hope that God's grace will save you. One of the pioneers of this maneuver was John Calvin. This has been brought up by, what's his name, the great sociologist Max ...

Julia: Weber?

Ainslie: Weber, yeah. Max Weber pointed out that Calvin had done this by saying that you are saved or damned from the beginning and nothing you can do can affect that. You put the person in a situation where all he does is evidence about whether he is, in fact, saved or damned.

If he has done something bad, you can just hope that that's not too much of a [sign] that it doesn't really mean that he's damned. It seems to me that that kind of stake makes people better saving and controlling themselves and so on.

Julia: Right. This brings me back to the prisoner's dilemma. Because the reason that we fail at the prisoner's dilemma, when we fail, is if we're using basically a causal decision theory. Where we calculate that, whether or not we're in the world where our partner is cooperating or defecting, either way, we're better off -- we can cause a better outcome for ourselves -- by defecting.

But an alternative would be to use an evidential decision theory, where we make a choice about what to do based on what evidence that would give us about which world we're in. By choosing to cooperate with our partner, that is some evidence that -- if he's anything like us -- us cooperating is evidence that he would also cooperate.

There's this sort of pseudo-causal thing we can do there that allows us to both end up choosing to cooperate. That seems very much like what John Calvin was describing. My choices are evidence about whether I'm going to heaven or hell.

Ainslie: In the personal case, I'm not even sure that's pseudo-causal. I mean, it's a kind of bargaining you can do and you're not fooling anybody. It's a sound bargaining technique. The fact that you know this is a test, and so you make sure that you get the answer right, doesn't mean that knowing it spoils the test. It's not like the SAT where if you know the right answer it's not a valid test. Here you know the right answer and what keeps you from giving it is that it's hard, not that it's secret. If you do the hard thing it's a perfectly valid test.

Julia: Yeah, I guess I just mean pseudo-causal in the case of prisoner's dilemma because you can't actually communicate with your partner in the prisoner's dilemma, but you can sort of communicate with your future self. You can bargain.

Ainslie: Yeah.

Julia: Right. Well maybe this is a good point to bring up another philosophical thought experiment called Kavka's Toxin... Some eccentric rich person, which seems to be the starting point for many of these philosophical thought experiments, an eccentric

rich person says, "Okay, you can have this wonderful reward, this million dollars, if you can sincerely intend to drink this unpleasant toxin, which won't harm you but will cause you to suffer greatly for a day or so. At the moment that you sincerely intend to drink it, I will give you the million dollars and then I'll leave. Whether or not you drink the toxin, I will never even know."

You're in this situation where it is totally worth it to suffer for a day to get a million dollars, but you know that once you get the million dollars you will have no incentive to actually drink the toxin, the guy will be gone. Knowing that makes it sort of impossible to intend to drink it, because you know your future self won't have any incentive to drink it once he already has the million dollars.

This feels very similar to me to the question of whether or not you consider your credibility to have been harmed by violating your diet or violating your rule.

Ainslie: Oh, yeah. It's a wonderful paradigm of that. Interestingly, you could probably run the experiment now, because you could put somebody in a scanner and see if they sincerely intended it.

Julia: Right. That's right.

Ainslie: Then the question is, of course, how do you sincerely intend, knowing you can go and evade the drink? My answer, if you have the ability to commit yourself in your head, you can sincerely intend, but the only way to commit yourself is to put your credibility in some larger situation, something worth the suffering from the drink, at stake.

Like if you had been in this situation before where you sacrifice something for somebody else, or do something very hard that you then don't feel like doing the morning it comes up. If you've got a track record of saying, "This is my stake, my remembrance of doing all these things and my expectation of being able to do it again." If that's what you're staking against backing out on your intention, then you can probably sincerely intend and get away with it. And otherwise you probably can't.

I think that's maybe what Kavka was heading toward. He wasn't just talking about funny thought experiments to puzzle philosophers. I think this was really the inspiration for his experiment.

Julia: Right. Interesting. Do you have any thoughts on whether it's better for our willpower to think of our future selves as being us or as being other people? I feel, anecdotally, I've observed both effects in different people. Where some people are less inclined to be kind to their future self when they think it's just them, because they're like, "Well, I'm not hurting anyone else. I'm just hurting myself, so that's okay."

But if they think of their future self as another person, they might feel guilty for defecting on that other person. Guilty for saddling that person with all the work and getting off scot free themselves.

For other people it's sort of the reverse. Once they really identify with their future self and think, "Actually that's me. I'm the person who is going to suffer tomorrow, not this future person, this stranger," then they feel more motivated to do the work now instead of putting it off.

Ainslie: That's an interesting distinction, and I think the answer is that it depends on your personal style.

Julia: Yeah.

Ainslie: People have developed ways of controlling themselves. I know there's actually been some research on this in the scanner. If somebody is asked about what they expect to do in the future and imagine doing things in the future, with no contingencies on it, just ask that... whether the parts of the brain that really image this kind of activity light up more. That person who lights up more is going to be more apt to choose larger-later rewards in the standard experiment that we've been through, outside the scanner, than otherwise. That would seem to argue that it's your future self. If you make it relevant to yourself.

Julia: Right.

Ainslie: On the other hand, the religious idea that you are hurting your relationship with Saint X or God or my deceased mother or whoever you might think of, that might be more motivation than just thinking about your future self.

I think people develop different styles. I think the future self things is a fairly obsessional kind of thing. I think obsessive compulsive people, of whom I'm one, are more apt to come up with that, and more apt to suffer from its side effects.

It's not without side effects. You can become overly concerned with rules and rather inauthentic, as the existentialists would say, by emphasizing this too much as a way of controlling yourself. Other people who are much more apt to say, "I will make other people important to me and I will value their opinions more and build up a reputation with them and build up the importance of their welfare in your own planning," they're more apt to rely on, "Well, somebody would be hurt if they knew that."

Julia: Right.

Ainslie: I think people vary, and probably it's something we should ask right off as therapists about what kind of way people think. It probably makes a big difference.

Julia: Yeah, interesting. Do you have any thoughts of other potential risks or downsides for using this inter-temporal bargaining model to think about one's own willpower?

Ainslie: Well, the main risk really is that you become rather legalistic with yourself and you spoil the fun. It's harder to live in the moment if you're always assessing the moment for what this tells you about the future, what kind of a test this is. People can really get fairly deep into that. You're also risking creating an unconscious, the

way Freud imagined the unconscious. If you see yourself violating, if you almost see yourself violating your rule, there are ways to kind of not do it. To look away at the last minute, or to define it differently. Denial, those familiar Freudian defense mechanisms. That will weaken your self control and will build up this whole body of things that you've got to turn away before remembering, which is a dynamic unconscious.

Also, you are, as you put it before, the more you bet on being able to succeed, the bigger the loss if you don't succeed. You create the risk of an abstinence violation effect. Sometimes people will, once that's happened a couple times, they'll say, "All right, there's some things I just can't control myself for. I just can't speak after dinner because I've chickened out a couple times, and if I really try to bet something big against that, I will impair my larger willpower. I will just say, that's something I just can't use my will on..." You create these circumscribed areas of avoidance or symptom because you have set up these bets with yourself and failed.

Yeah, there's a lot of hazard to it. It's, on the one hand, the most accessible and universal way to control ourselves, and on the other hand, maybe the most costly if we go too far. I think most therapists, most therapies, if you look at the books, ways of avoiding the temptation, they're ways of dealing with your conscience and your beating yourself around the head because you've lost. You have things like "muster-bating," was one of the more colorful terms for it, in cognitive behavior therapy, or cognitive maps, which I think Carl Rogers used that. Of course, Freud's punitive super ego, these are the meat and drink of interpersonal therapists. They're all from having tried this maneuver and gotten in trouble with it.

Julia: Before we wrap up this section of the podcast, I'd love to hear if you have any advice for people trying to navigate this tricky balance that we've been talking about, and get as much of the goods as they can out of this framework of thinking about willpower, and minimize the bads that we've been talking about. Do you have any suggestions?

Ainslie: Yeah, I think my main one is to make friends with your future self.

Julia: I like it.

Ainslie: It's going to be you, after all. Don't get too technical. Don't use a gimmick to try to bind yourself so you absolutely can't get out of it. Your future self will do something just as bad on the other side, believe me. It's good to keep in mind that you are dealing with a community of people you care about, that is you, or maybe even more people you care about.

Really, it's got to be kind of informal, and you can't carry it to an extreme, or it's going to start getting distorted. It's not really a big league difference, and you know the rules to begin with. My advice is just don't get carried away with them.

Julia: Great. Okay, excellent. Well, this concludes this section of the podcast, so we'll move on now to the Rationally Speaking Pick.

[interlude]

Julia: Welcome back. Every episode, we invite our guest to introduce the Rationally Speaking Pick of the Episode. That's a book, or website, or movie, or something that has influenced his or her thinking in an interesting way. Dr. Ainslie, what is your pick for today's episode?

Ainslie: Well, way back in the beginning, I read Tom Schelling's book. Thomas Schelling. He's a Nobel Laureate now.

Julia: Oh, yeah.

Ainslie: His book, called the Strategy of Conflict, which basically invented Game Theory, and it's a very easy-to-read, down-to-earth book about the situations you get into in bargaining with somebody. He used it with his classes, and so they're all very accessible examples, and not too mathematical. At the other end of the curve ...

Julia: Do you think we can apply Schelling's Game Theory to bargaining with ourselves?

Ainslie: Oh, with ourselves and with other people. He really is a very canny theorist about bargaining.

Julia: Excellent.

Ainslie: People still use him a lot. The other thing is Don Ross, who is an expert both in philosophy, psychology, and economics. He's actually got an appointment in economics at Cape Town, and he's just written a book on the philosophy of economics. He is a polymath, and he goes deeply into all these things. That's been the most influential book lately on me.

Julia: Excellent.

Ainslie: That's those, and if I have the right, I'd like to suggest a movie.

Julia: Sure, go ahead. People don't often recommend movies. I wish they did.

Ainslie: That is Inside Out.

Julia: Oh, I love Inside Out.

Ainslie: Inside Out is exactly my idea about interests. A lot of the movies about inside your head and how this works have this little homunculus there in a booth who is controlling your gut, and your sex, and your muscles here and there, just like levers. That's not how, I think, this works. How it works is that you have these different interests, all of them have motives, and who have to bargain with each other for acceptance.

Julia: Right.

Ainslie: That's how the people Inside Out work. Now, he might have added Cookie Monster. He doesn't have a self that just wants to consume things, but aside from that, he's a wonderful illustration of just what I mean by the interests getting along with each other, the marketplace of interests that bargain with each other. Go to that. It's also a wonderful movie.

Julia: Excellent.

Ainslie: I've seen it a couple times, and I'll see it again.

Julia: Well, there is going to be an Inside Out 2, I'm almost certain.

Ainslie: Oh, is there?

Julia: So if you get in touch with the filmmakers fast enough, you might be able to get a Cookie Monster in there.

Ainslie: Oh, I will write him, and give him my suggestions. He apparently got his original ideas from a psychologist of emotion named Ekman.

Julia: Oh, yeah. Interesting.

Ainslie: He used all of his main emotions except for surprise. He figured you couldn't have an ongoing character called Surprise, but the rest of them are right out of Ekman, Paul Ekman. That's a book you might like too, if anybody wanted it.

Julia: Oh, what's that called?

Ainslie: Paul Ekman, E-K-M-A-N, has written a couple of books about emotion, and they're very good.

Julia: Excellent. Well, we're about out of time, so thank you so much Dr. Ainslie for being on the show.

Ainslie: Well, and thank you. It's been wonderful talking with you. You're a very good interviewer, and a very good asker of questions.

Julia: Well, thank you. I strive to be.

All right. Well, this concludes another episode of Rationally Speaking. Join us next time for more explorations on the borderlands between reason and nonsense.