

FROM OUTSIDE OF ETHICS

One of an occasional series of reviews of books outside the bounds of moral, political, legal, and social philosophy that may nevertheless be of deep interest to people working in those fields.

Marušić, Berislav. *Evidence and Agency: Norms of Belief for Promising and Resolving*. Oxford: Oxford University Press, 2015. Pp. 256. \$65.00 (cloth).

I. THE PROBLEM OF DIFFICULT ACTION

Here's a plausible thought: we should believe what our evidence supports. Yet we often promise or resolve to do difficult things—exercise, quit smoking, be faithful—and we often trust others who do the same. Sincere promising and resolving seem to require belief. Yet we know that people who've similarly resolved often fail. Nevertheless, despite this (sometimes substantial) evidence that we'll likely fail, we persist in sincerely promising, resolving, and trusting others when they do the same.

Berislav Marušić presents this problem of difficult action as a challenge to the evidentialist view implicit in the plausible thought above. Evidentialism, he formidably argues, threatens to make this practice of promising and resolving, which is so central to our lives, irrational, if not unintelligible. He then offers an interesting and original alternative view inspired by existentialist ideas that contemporary, analytic philosophers tend to avoid, whether out of fear, distaste, or confusion. (The light Marušić sheds on these ideas, clarifying them so we can understand and assess them, is one of his many important contributions.) We can rationally believe against the evidence, Marušić argues, and to refuse to do so is to succumb to a kind of bad faith.

At least two thoughts should tempt us toward Marušić's position. First, it seems implausible that we can never rationally believe that, despite our spotty track record, this time we will follow through. Second, if belief is always irrational in such cases, we face an unfortunate dilemma: we must either never resolve or promise to do the difficult and improbable or embrace irrationality and self-deception.

The first option is unacceptable, maybe impossible: such resolutions and promises make up the fabric of our lives. We simply can't wholesale avoid them. The second option is unfortunate: it allows for sincere resolve and trust at the cost of irrationality or self-deception.

A major appeal of Marušić's view is that it promises to avoid these consequences—from the implausible to the unacceptable and unfortunate. Doing so, Marušić argues, requires that we focus on agency, rather than on evidence or likelihoods of failure. It requires deliberating about what to do, not predicting what we will do. Prediction misses the fact that our actions are, even in the most difficult cases, largely up to us. Ignoring this, Marušić argues, constitutes a denial and disrespect of our agency—a kind of bad faith. Thus, not only can belief against the evidence be rational, but also taking agency seriously requires it.

II. THE VIEW

Consider the following case:

Marathon. Suppose I'm deliberating about whether to run a marathon. This is hard—and not just during the race. It requires serious dedication and training. This may be no big deal for otherwise-athletic folks, but I have only maintained about 30 percent of my exercise resolutions. Plausibly, there's only a 30 percent chance that I'll maintain this one.

If I resolve to run the marathon, I should believe that I will maintain my resolution. But to believe this is to believe against my evidence. Nevertheless, Marušić argues, believing that I will maintain my resolution is a rational option for me. This is because running the marathon, though difficult and improbable, is ultimately up to me. Whether I succeed depends on whether I keep trying and training. Recognizing this, Marušić argues, allows me to rationally believe, despite my evidence, that I won't wimp out.

This much seems right: evidence suggests a chance of success, so maybe it is rational to believe that you'll possibly run the marathon—not that you will or won't. But to focus on the evidence at all is to treat something that is up to you as if it isn't, Marušić argues, and this is a mistake. To see this, consider a variant on the marathon example:

Lottery. I am a superb athlete who always fulfills her exercise resolutions. There is, however, a substantial chance that I won't run the marathon: a lottery determines who runs, and only 30 percent of entrants win.

The chance of failure is technically the same, but it seems that I should view it differently in Lottery. What is the relevant difference? It's up to me to keep training and trying, but it isn't up to me to win a lottery. This seems right, and it provides some evidence for Marušić's view that the deliberative perspective affects the significance of facts about likely success.

We must respect this difference, Marušić argues, but a well-informed observer whose only aim is prediction needn't. Consider that a bookie offering "bets on the athletic achievements of philosophers" (20) should do his job and focus on evidence. We, however, must respect our agency and that of our loved ones, Marušić argues. So, we shouldn't just follow the evidence wherever it leads.

Marušić's arguments, however, support a stronger position. The difference between Marathon and Lottery is one we can acknowledge whatever our relation to the deliberating agent. If respecting agency requires ignoring certain evidence, then the bookie also disrespects our agency. Perhaps disrespecting strangers' agency is less bad—but it's still disrespect. Marušić's view, then, is more general than he suggests. (This, I think, is a strength.)

Thus, Marušić holds that attending to evidence about our likely success constitutes overlooking or denying and, thereby, disrespecting agency. But we should be suspicious here. The disrespect claim seems largely metaphorical; the denial claim seems false. For suppose I have resolved to do something difficult that I know I am likely to fail at, like starting an exercise routine. Closely evaluating my evidence—including my chance of failure—seems crucial for strategizing

about how to proceed. Rather than denying or ignoring my agency, this seems a way of taking it seriously in all its imperfections.

Marušić agrees that it's important for us to have a realistic view of our agency: he doesn't want us to overestimate it any more than he wants us to ignore it (these are the two forms of Sartrean bad faith; 25). But how can we have a realistic view of ourselves and our situation while ignoring evidence about our likelihood of success? By focusing, Marušić argues, on the difficulty of what we've set out to do. Similarly, a realistic view of others' testimony, their promises and resolutions, requires that we attend to their trustworthiness. These crucial notions are meant to mark the path between the two bad options of under- and overestimating our agency. They also mark the trickiest and most difficult problem for Marušić's view (125), which I'll focus on in what remains. I'll aim to challenge Marušić's view in a way that underscores some fruitful paths that engaging with this original and insightful work might take us on.

III. WORRIES

My central question is, how should we understand 'difficult' and 'trustworthy' so that they give the right results about cases but don't collapse into the sorts of evidential considerations we must ignore? I have three related worries. The first is that there is no plausible practical conception of 'trustworthy'. The second is that the practical conception of 'difficult' doesn't get important cases right. The third is that if the relevant conceptions of 'difficult' and 'trustworthy' are just the practical flip side of theoretical considerations about likelihoods, then they'll give us the same results that focusing on likelihoods would. But then Marušić's view, which promised to permit rational belief against the evidence, doesn't do so: a focus on 'difficulty' and 'trustworthiness' doesn't help us avoid our problems—it just makes us take the scenic route. I'll explain each of these in turn.

A. *No Purely Practical Conception of 'Trustworthy'*

Your lover was unfaithful but is now contrite and repenting. As you're considering whether you can forgive the betrayal, they sincerely promise that it won't happen again. If you accept your lover's promise, you should believe that they will be faithful—but then you'd be believing against the evidence (most cheaters are repeaters).

Just as Marušić thinks that we can rationally believe against the evidence in the first person case, he thinks that we can rationally trust against the evidence in this second person case. Marušić's view isn't the silly one that belief or trust is always rational in such cases, however. You should only trust your lover, he argues, if you have reason to think they're trustworthy (185, 193 n. 31).

Trustworthiness thus plays a crucial role for Marušić: it allows him to say when trust against the evidence is rational and when it's not. Importantly, for Marušić we shouldn't understand trustworthiness in terms of reliability (200). The evidence pertaining to your lover's reliability is unflattering—yet you may rationally trust against it, if your lover is trustworthy. Marušić thus needs these concepts to be different.

Trustworthiness and reliability are different, Marušić argues, because "trust and reliance are different" (200). But this is too quick. Trust and reliance may differ without trustworthiness and reliability also differing. Trusting you may well

involve taking a different stance toward you than relying on you does. Nevertheless, what justifies my trust in you (your trustworthiness) may just be the same as what justifies my relying on you (your reliability).

Marušić provides another reason to distinguish trustworthiness from reliability. We investigate trustworthiness differently than we investigate reliability. In considering whether you're trustworthy, he says, I'm considering whether to accept your offer of an answer. In considering whether you are reliable, I'm instead seeking to answer the question for myself (200).

But what am I doing when I consider whether to accept your offer of an answer? And is seeking to answer the question myself the only, or even the best, way of evaluating your reliability? The answer to the first question is unclear; the answer to the second is "no." Consider the latter first.

Suppose you say that it's three o'clock. I could evaluate your reliability by checking the time myself: if you're right, that is evidence of your reliability. But I'd probably get more or better evidence by investigating your honesty, eyesight, or time-telling ability instead. Furthermore, these more onerous investigations are often necessary: in evaluating whether to accept an expert's offer of an answer, I'm often not an expert and thus not in a position to answer the question myself. Investigating your reliability is thus usually best done by investigating your character—moral, epistemic, or other.

Now back to the first question: how should I investigate the other matter, whether to accept your offer of an answer? Plausibly, I should proceed in exactly the same way: by investigating your honesty, and so on. But then I'm just investigating your reliability.

We are yet to see a good reason for thinking that trustworthiness and reliability are distinct as Marušić needs them to be. Furthermore, we have seen how natural and plausible it is to think that they are importantly connected: what justifies the one also justifies the other.

Even Marušić often comes close to running the two together: your trustworthiness is "of great importance in determining when it is rational to trust" you, and your "tainted track record"—your reliability?—matters in "settling" whether you'll follow through (200). This is why, he writes, your lover should seek to "make it credible" that they will keep their promise (200). But how do we make it credible that we will follow through except by providing evidence that we will?

This is the difficulty. Marušić needs trustworthiness to be distinct from reliability, or else we rely on the wrong sorts of reasons (thereby disrespecting our agency). But Marušić also needs trustworthiness to function very much like reliability, or else it cannot determine when distrust is appropriate. It can't always be rational to trust loved ones, just because they are loved ones: distrust must sometimes be rational. When? When our loved ones are not trustworthy. But investigating our loved ones' trustworthiness seems to require investigating their character and track record, which seems pretty close to investigating their reliability.

B. Practical Conception of 'Difficulty' Gets Things Wrong

Return to the problem of first person deliberation. The challenge for Marušić's view is to negotiate between two sorts of bad faith, which he illustrates thus: "The gambler who succumbs to the first kind of bad faith reasons that since he is un-

able to rationally predict that he would keep his resolution, he should not make it. The gambler who succumbs to the second . . . neglects the difficulty of keeping his resolution and simply chooses the option he prefers. Perhaps he thinks that all it takes to quit gambling is to make the resolution to quit, or he thinks that it will be easy to keep the resolution" (129). Avoiding both kinds of bad faith requires considering the difficulty of keeping our resolutions (129). Marušić initially defines a difficult action as one for which there is a significant chance of failure (87). This won't do. If 'difficult' means "likely to fail," then, evaluating the difficulty of keeping our resolution, we'd have to consider our chance of failure. But that is precisely what, according to Marušić, we shouldn't do. So this definition is inconsistent with Marušić's aims.

It is also wrong: an action can be difficult for me without being such that I'm likely to fail. Moving this heavy dresser would be difficult for me, but I have no doubt I'll succeed if I try—it's just that succeeding will take a lot of effort.

This suggests a better way of understanding difficult action: as requiring effort or perseverance. (Marušić calls this the "practical" sense of difficult, distinguishing it from the former "theoretical" sense [86]. But, as we've seen, we should reject the latter.) This allows us to separate how difficult something will be from whether we're likely to fail at it.

Marušić can embrace this suggestion and reformulate his central problem as that of difficult and improbable action. He can then clearly and consistently state his requirement that we consider the difficulty (effort) required for success, and not our chances of failure.

The question now is whether thinking only about how much effort it will take to follow through on our promises or resolutions gives us the right results. The toughest cases for Marušić involve very difficult and improbable actions, like quitting gambling or overcoming addiction. On the same spectrum are cases of overcoming deeply ingrained bad habits. Consider the following:

Procrastinate. Professor Procrastinate is invited to review a book. It would be best if he wrote the review on time; it would be terrible if he accepted but procrastinated too long. Unfortunately, he "has evidence from his past work habits that there is an excellent chance that he will procrastinate" (143).

Though Procrastinate is a procrastinator, his bad habit is just that: a habit. However ingrained this habit is, however difficult it is for Procrastinate to get his work in on time, it's still up to him to do so. This doesn't mean that he can succeed by mere exertion of willpower. It does, however, mean that breaking his habit requires, first and foremost, personal choice and persistence. Procrastinate must choose to stop procrastinating, and he must persist despite the difficulty of doing so.

These are tough cases because they make Marušić's opponent's view seem so plausible. Marušić insists that respecting and taking responsibility for our agency requires us to ignore evidence of our likely failure. But how does acknowledging such evidence entail failing to take responsibility for our agency? Isn't doing so, instead, a way of recognizing and accounting for our agency—our deeply flawed agency? And isn't that the responsible thing to do?

We can push this worry further by noting the need for strategic choice in such cases. By employing strategic choice, rather than willpower, I avoid temptation rather than resist it. For the gambler, this might involve paying to be banned from casinos (this is a thing) or entering a residential rehab program. For Procrastinate, it might involve using the Self-Control app to block the internet (also a thing). In so proceeding, we tie ourselves to the mast, as Odysseus did to avoid the sirens.

Strategic choice is clearly the wise option in such circumstances: exerting willpower is costly, tiring, and less likely to work. If we really want to break such habits, we should recognize that sheer exertion of willpower won't cut it. We are only likely to succeed if we are strategic.

But now we've switched to talk of chances and adopted exactly that third person perspective that, Marušić argues, denies our agency. Can we avoid this? Can considering only difficulty justify strategic choice? The worry is that if we don't think we're likely to fail, we lack sufficient reason to tie ourselves to the mast.

Marušić resists this: tying yourself to the mast makes it easier to do the difficult thing you want to do. That is reason enough to do it. This seems right: I know I will finish this piece on time, yet I use the Self-Control app because it makes concentrating easier. It also makes it more likely that I'll finish on time, but that needn't be my reason for using it.

Procrastinate, Marušić writes, must recognize that it's "difficult for him to finish his work on time" (145). He must also recognize the cost, to himself and others, of his procrastinations. In light of this, it would be foolish and irresponsible for him to go optimistically forward, confident he will succeed since doing so is, after all, up to him. (This is the second sort of bad faith.)

These considerations emerge out of Procrastinate's recognition that it is difficult for him to finish on time. They might, Marušić argues, add up so as to make it "improper for him to promise to review the book" (145), unless he does something to make his promise credible. Strategic choice can help him do this.

This is what we want from agents like Procrastinate: either commit and follow through, or don't commit in the first place. What we definitely don't want Procrastinate to do is commit, brace himself, and try really hard—using willpower alone—to finish on time.

This much is right: that strategic choice makes following through easier is a reason to do it. But I think that Marušić needs something stronger. The difficulty of meeting a deadline and the costliness of relying solely on willpower do motivate me to pull an Odysseus and block the internet—but they don't necessitate that decision. The only thing, I think, that could necessitate it is focusing on our chance of failure.

Recall my heavy dresser: it might be difficult for me to move it, but I know I can. It might be wise for me to wait for someone to help me, since doing it myself is hard and I could stub a toe or pull a muscle. But I want to move the dresser now! Maybe I just want to; maybe it's better to do it today. My reason for waiting could thus be outweighed by my reasons for doing it now.

Strategic choice is clearly the right option when you want to do something and you believe that you will probably fail at it. Without that belief, you might still opt for strategic choice to minimize effort. But you might reasonably not, if

doing so is costly (as rehab is), or if you value success through sheer force of will. In those cases, the fact that something is going to be more difficult, even very difficult, without strategic choice is not sufficient reason for strategic choice. For example, Procrastinate might reason, “I want to overcome my procrastination by pure force of will. I realize that’s much more difficult than doing it with the crutches of apps, writing buddies, and so on. But I want to exercise my freedom and my self-control. And I can do this. Succeeding is, after all, up to me.”

Similarly, the gambler might reason, “Quitting will be a lot easier if I avoid the casino and my gambling friends. But those are my friends! And I’m happy at the casino, even when not gambling. I don’t want to give that up. I guess I’ll have to work harder, so I can have both my friends and a new, gambling-free life.”

These agents express familiar positions and values. I nevertheless want to grab them by the shoulders, shake them, and shout, “Don’t do it this way, you will fail!” If Marušić is right, I shouldn’t do this: I’d be disrespecting their agency and inviting them to do the same.

In fact, Marušić grants that strategic choice requires taking an “observer’s view” (131). He also agrees that it’s crucial in such cases. He thus concludes that sometimes—but only sometimes—we should make predictions about our likely success. He doesn’t think that this undermines his argument, however, because it limits the need for prediction to only certain situations (132).

However, I think that Marušić again underestimates the force of his own argument. If he’s right that taking the observer’s perspective disrespects agency, then employing strategic choice, which requires taking that perspective, disrespects agency. That strategizing from this perspective is the wise option doesn’t make it less disrespectful. The stronger position, which taking Marušić seriously pushes us to, is that taking the observer’s perspective is always disrespectful wherever we are and whatever our situation.

Thus, if Marušić is right, then, as deliberating agents, we are allowed to look at the wreckage of our past resolutions and say, “This is going to be hard.” But looking at that same wreckage, we should never conclude that we are likely to fail. Even if that wreckage is just our terrible track record, and even though it is a short step from “I failed seven out of the past ten times” to “I have a 70 percent chance of failure,” we are not to make that step. Indeed, if someone were to ask us, “How likely are you to succeed?” we should answer, “You heard me! I’ve resolved. Such predictions are not my purview.” Procrastinate and Gambler can respond in just this way to my shoulder shaking. So, it seems, considerations of difficulty, understood in a purely practical guise, just cannot get the right verdict in these hardest of cases. And we can only let prediction in at the cost of precisely the sort of disrespect that Marušić is at pains to avoid.

C. *An Elaborate Detour?*

Suppose now that there are plausible, practical conceptions of ‘difficult’ and ‘trustworthy’ that get the right verdicts. A worry remains, that they’ll make Marušić’s view extensionally equivalent to his opponent’s. Of course, that two views get the same results doesn’t mean that they’re equally good—Marušić’s

could still be better. But if Marušić and the evidentialist agree on all the cases, then Marušić has failed in his promise of rational belief against the evidence.

One aspect of this worry emerges from our discussion of strategic choice. Difficulty considerations should make us approach resolutions more wisely, by leading us to “do something different this time,” like employ strategic choice. Marušić holds that in doing so we make it “more credible” that we will follow through, and “thereby” practically rational to resolve or promise (146).

There is a problem here, however. First, doing something different changes the relevant evidence, so we no longer believe against the evidence. Second, if these aren’t cases of believing against the evidence, then they cannot motivate Marušić’s view.

On the first point, by approaching his review differently, so that it’s easier to finish on time, Procrastinate makes it more likely that he will finish on time. He simultaneously makes his track record irrelevant—he’s doing something importantly different, after all. In believing he’ll succeed, Procrastinate isn’t believing against the evidence. He chose this way of doing things precisely because it makes him likely to succeed.

Straightforwardly following Marušić’s advice to avoid the evidence risks overestimating our agency. Considerations of difficulty and trustworthiness are meant to help us recognize the limits of our agency and be appropriately strategic. Straightforwardly taking this advice, however, makes our promises and resolutions more credible and, simultaneously, compatible with our evidence. This eliminates the only purported examples of cases in which we can rationally and plausibly believe against the evidence. It also makes the problem of difficult action a lot easier—it’s no longer the problem of difficult and improbable action.

If this is right, then the difference between Marušić and his opponents is less than it seemed. The unflattering thought is this: replacing our standard belief and evidence model with Marušić’s is like replacing it with a Rube Goldberg machine. On the standard model, the evidence comes in the front door and directly justifies our belief. On Marušić’s, the evidence must throw a paper airplane through the window, which pushes over some dominoes, which hit a marble, that falls into a bucket, that tips over, . . . , and then results in a belief—the same belief you would have formed if you had just looked directly at the evidence.

Marušić’s view is meant to allow us to rationally believe we can overcome difficulties and do unlikely things even when the odds, as well as the evidence, are stacked against us. Similarly, it’s meant to allow us to trust our loved ones even when the odds, as well as the evidence, are stacked against them. This is the most interesting and controversial bit of Marušić’s view. If this worry is right, then we have lost our grip on it.

IV. CONCLUSION

By presenting the problem of difficult (and improbable) action, Marušić raises what is perhaps the best challenge to the dominant, evidentialist model of rational belief. It can seem inescapable that we should believe what our evidence supports—yet if we must, then we cannot rationally and sincerely promise or resolve to do difficult things. Marušić’s own solution to this problem is insightful and

deeply original. It is also a pleasure to read: the chapters presenting Marušić's positive proposals are as close to a philosophical page-turner as you can get. Challenges remain, of course. Yet one senses, even as one resists the view, that one is learning something.

ACKNOWLEDGMENT: Many thanks to Lindsay Crawford, Maya Eddon, Catherine Elgin, Alan Hájek, Sophie Horowitz, Hilary Kornblith, Kathryn Lindeman, Elisa Mai, Berislav Marušić, Alejandro Pérez Carballo, Nishi Shah, Yuan Tian, and Kenny Walden for helpful discussions that shaped my thinking on these matters.

KATIA VAVOVA
Mount Holyoke College