Journal of Ethics and Social Philosophy
VOL. 27, NO. 3 · MAY 2024
https://doi.org/10.26556/jesp.v27i3.3048

VALUE CAPTURE

C. Thi Nguyen

HERE IS A STORY about how metrics can change people. A relative of mine had been planning a long European vacation with some old friends, John and Shelley. My relative had been looking forward to seeing the sights with her friends—touring museums, seeing operas, having long dinners. But, she says, the entire vacation was dominated by John and Shelley's relationship with their Fitbits. John and Shelley would not go to the opera with her: not enough steps. They would cancel dinner dates because they had not met their daily step goals yet. My guess is that John and Shelley never consciously decided that step counts were more important than, say, art or friendship. The Fitbit just spoke more loudly in their internal deliberation, and there was no Artbit or Friendbit to compete. The clarity of those metrics just swamped quieter considerations.

And even if fitness was your main goal, the Fitbit can exert a narrowing influence. Exercise can be valuable in all sorts of ways that are not measured by a Fitbit. A Fitbit does not capture the ecstasy of complex, skillful motion. It does not capture the camaraderie of team sports, the meditative calm of paddling a canoe across a quiet lake, or the aesthetic loveliness of a delicate rock-climbing move. A Fitbit measures exactly one thing: steps. That limitation arises from its particular institutional and technological embeddedness. Fitbits are constrained by what mass-produced devices can easily measure and aggregate, given present-day technologies and institutional arrangements. We know how to make a watch that automatically measures steps, but not how to make a watch that automatically tracks your spiritual renewal.

Of course, you do not have to value what the Fitbit measures. You could just use a Fitbit as a source of data. But the Fitbit tempts us to do more. The Fitbit presents its output, not just as mere information but as an evaluation: a score. And when you buy into the Fitbit's preferred motivational scheme—when you adopt its scores as your values—you get all kinds of rewards. You gain the motivational benefits of having clear feedback about how well you are doing, of competing along a well-defined scale. All you have to do is give up on having fine and detailed control over your own values. Here is one way to put it: when you buy into a Fitbit's preferred value system, you are *outsourcing the process of value deliberation.*

Fitbit is just one example of a larger phenomenon that we can call *value capture*. Value capture happens when your environment presents you with simplified versions of your values, and those simple versions come to dominate your practical reasoning. Value capture offers you a quick shortcut—an opportunity to take on *prefabricated* values. You do not have to go through the painful process of value deliberation if you can get your values off the shelf.

I want to focus on one particularly clear, and quite common, form of value capture: when an institution presents you with some metric, and then you internalize that metric. You start exercising for your health, but you come to care about losing weight or optimizing your body mass index. Or you go on Twitter to connect to people and have fun but come to care more about maximizing your like, retweet, and follower counts. Or you go into philosophy graduate school for a love of wisdom but come out aimed at getting fancy grants, publications into highly ranked journals, and placement at a highly ranked institution. As anthropologist Sally Engle Merry puts it, the culture of indicators and metrics is "a form of governance that engages a person in governing himself or herself in terms of standards set by others."[1] I will focus for much of this paper on such institutional value capture. Metrics are the starkest case of value capture, and we are fortunate to have a rich empirical literature studying the social effect of metrics. But metrics are just a starting point; there are many other forms of value capture worth investigating.

Many of us feel an intuitive horror when contemplating cases of institutional value capture. But it is rather difficult to say in a principled way exactly why value capture is so horrifying. For one thing, value capture is often consensual. People buy Fitbits precisely because they know that those step counts will motivate them; they *want* to be captured because the motivational bump seems worthwhile. Such gamified technologies are frequently sold as a way to overcome weakness of the will and seem to succeed at doing so. The point of a Fitbit is to motivate you to walk more, and it does seem to work.

Why might this strike some of us as horrifying rather than as simply a useful and empowering tool? I will suggest that there is a problem with the *nature of the values* on offer. The problem with internalizing institutional metrics is not simply that we are getting our values from the outside. It is that such metrics are subject to the demand for a certain kind of stability and institutional usability. These institutional demands push our metrics away from the subtle, the dynamic, the sensitive—and toward what can easily be measured at scale, propagated across institutional units, and recorded in institutional memory. When we take on such metrics as our values—when we internalize them—we

1 Merry, *Seductions of Quantification*, 33.

are imposing a narrowed filter on our values. We are letting the logic of institutions play a determining role in the articulation of our values.

Institutional value capture offers us a delightful reward. Once we have permitted ourselves to be value captured, our values become clear, coherent, and shared. Now we can be easily understood—unambiguously, almost effortlessly. But such clarity requires a degree of stabilization. Such clear, stabilized values arise from and are deeply embedded in external institutions and institutional processes. That stabilization has some benefits and some costs. Sometimes those costs may be worth paying—but we should at first get clear about what, exactly, they are.

In value capture, we *outsource the process of value deliberation*. And, as with other forms of outsourcing, there is a trade-off. You get the outsourced objects quickly and easily, and they fit neatly into a larger network of other standardized and modular parts. Somebody else has formulated our values for us and done the work of embedding them in readymade systems of measurement and technologies of motivation. When we adopt those values, we gain access to readymade methods for justification. It is easy to justify yourself in the language of metrics because metrics are easy to understand. They have, in fact, been engineered to be so. The cost of value capture is that we give up on the process of finely tuning our values to our own context: our personalities, our peculiar culture, our particular corner of the world. Outsourced values are not custom-tailored. In value capture, you are taking on prefabricated values.

1. A CASE STUDY: THE LAW SCHOOL RANKINGS

The social draw of quantification has been the subject of some extremely useful recent empirical studies from anthropologists, historians, and sociologists. My favorite is *Engines of Anxiety,* a study of the cultural effects of the *US News and World Report* (USNWR) law school rankings by sociologists Wendy Espeland and Michael Sauder.[2]

Before USNWR, they say, there were no law school rankings. Students often picked law schools through a complex process of evaluation, deliberation, and self-reflection. They got to know a school by reading about its mission, by talking to people, or by visiting. Importantly, different law schools pursued different missions. Some were tuned to academic legal research, others to the corporate world. Some law schools were devoted to social activism—toward supporting the local community or serving underrepresented populations. The process of choosing a law school often triggered a certain degree of soul-searching in the

2 Espeland and Sauder, *Engines of Anxiety*.

students. The complex value plurality involved in the choice pushed students to reflect on what they wanted from their own legal educations and legal careers.

The rankings displaced all that. Espeland and Sauder studied online discussions between prospective law students. They found that once *USNWR* started publishing its rankings, those rankings came to dominate the choice process for most students. And the same is true for nonstudents: the public perception of law schools immediately reoriented itself along *USNWR*'s rankings. Espeland and Sauder say that the rankings drove value plurality out of the legal educational system. Many schools used to genuinely pursue their different missions. And many of those missions involved pursuing values that are not tracked by *USNWR*'s ranking formula—like, say, supporting local underserved minority communities. But following such a distinctive mission invariably meant dropping spots in the rankings, which promptly resulted in precipitous drops in donations and student interest. Most schools, report Espeland and Sauder, have since abandoned their original missions and reoriented their admissions process and educational methodology toward performance in *USNWR*'s ranking calculations. And what matters the most to that ranking is the grade point average (GPA) and Law School Admission Test score of the incoming class and the employment rate of the outgoing class.[3]

In the case of the law schools themselves, the change in goals could be understood as a case of perverse incentives. Law school administrators were forced to align their efforts with the rankings, even if their own values were unchanged. But with prospective law students, the problem seems to run much deeper. The rankings seem to exert a magnetic pull over students' values. Some students, of course, were merely responding to incentives—since potential employers also care about law schools' rankings. But a majority of students, say Espeland and Sauder, seemed to care directly about those rankings. Instead of exploring their own values and desires for their legal education, they seem to presume that the process of going to law school should be oriented toward getting into the "best" law school, where "best" is determined strictly by the rankings. The existence of that clear, vivid, objective-seeming list offers an easy substitute for the process of personal value deliberation.

The effect on students I take to be a clear example of value capture. The fact that value capture occurs I take to be an empirical matter—and its existence is well-documented.[4] My goal here is to think about the harms of value capture.

3 Espeland and Sauder, *Engines of Anxiety*, 43.

4 Beyond Espeland and Sauder, see Porter, *Trust in Numbers*; Scott, *Seeing Like a State*; and Merry, *Seductions of Quantification*, for good entry points into the literature.

2. VALUE CAPTURE

Value capture happens when:

1. An agent has values that are rich, subtle, or inchoate (or they are in the process of developing such values).
2. That agent is immersed in some larger context (often an institutional context) that presents an explicit expression of some value (which is typically simplified, standardized, and/or quantified).
3. This explicit expression of value, in unmodified form, comes to dominate the entity's practical reasoning and deliberative process in the relevant domain.

If you would like a portable version, try this: *value capture happens when a person or group adopts an externally sourced value as their own, without adapting it to their particular context*.

Let us take a moment to get clearer on what, exactly, counts as value capture. First, notice that value capture includes both voluntary and nonvoluntary adoptions of an external value. It certainly counts as value capture if, say, you were brainwashed, and an external value was somehow injected into you against your will. But it equally counts as value capture if you willingly and voluntarily adopted that external value—perhaps because it is easier or helps you to fit more easily within your profession or because it lets you avoid the painful process of value deliberation. The target of my criticism here is not simply those cases of involuntary value transformation. I am interested in the problem with letting externally sourced values dominate one's practical reasoning—even if that dominance was established knowingly and consensually.

Next, my definition of value capture is narrowly aimed at those cases where the entity uses the external expression of value precisely as given. It is aimed at those cases where we internalize and deploy an external value just as we found it, *without further adjustment*—without further contouring it, interpreting it, or fine-tuning it to ourselves. Value capture does not include cases where you get the seed of your values from the outside and then start fiddling with them. If you get the starting seed for your values from your family, your culture, your religion but then tweak them to fit your personality and place in the world—that is not value capture. Value capture is when an externally sourced value, like a metric, comes to dominate your practical reasoning in its given form—when your goal is simply to get to that higher ranking, those higher citation rates, those more likes.

I have been speaking so far about the value capture of individuals by large-scale institutions. Such examples are vivid and familiar. But they can invite a

simplistic reading of the problem: that real values are somehow original inventions of the individual, and that socially generated values are somehow fake. It is tempting to think that what is going on here is, say, a battle over individual authenticity—some conflict between the solitary free spirit and the forces of social conformity. But the problem is much more complex than that. For one thing, our values are often acquired in their initial seed from social sources—parents, teachers, friends, colleagues. For another, we often develop our values in community with others.

And crucially, value capture is a problem that can afflict groups too. A philosophy department can be captured by the larger university's focus on student evaluation scores. In my own experience, the clarity of an institutional metric can quickly come to dominate the attention of a deliberating group. Even when a group agrees that they care more about some inchoate value—like, say, fostering curiosity—the actual day-to-day decisions end up driven by whatever clear metrics happen to be on hand. Merry, in *The Seductions of Quantification*, offers a good example.[5] The United Nations (UN) publishes the Human Development Index (HDI)—a quantified ranking of all countries in the world in terms of how they supported quality of life. Merry says that the committee behind the HDI published it with a very clear and loud set of qualifications. They published it with a lengthy report on the complexity and multidimensionality of "quality of life" measures and clearly stated that the HDI ranking was simply a gross oversimplification. Unsurprisingly, says Merry, the full report was largely ignored. Once the HDI was published, governments the world over became incredibly invested in advancing their ranking—even though the score was not attached to any concrete real-world incentives or rewards. Here is a case where entire governance cultures have been value captured by an external metric. So, in my account of value capture, I intend "agent" to be in a broad sense, including individual persons and group agents.[6]

Next: condition 3 in my account specifies that the external expression of value "comes to dominate" the entity's practical reasoning and deliberative process in the relevant domain. I mean the notion of "dominate" to be quite substantive here. Value capture occurs when an external value becomes the dominant source of reasons for action in a domain. It is not value capture if I adopt an external value in a controlled manner—as temporary instruments,

5 Merry, *Seductions of Quantification*.

6 I use "entity" rather than "agent" here because, while "agent" includes group agents, I think the category is not large enough. I suspect that some loosely organized communities can qualify as having values but not have sufficient internal cohesiveness to count as a group agent. See discussion of shared values in Hedahl and Huebner, "Sharing Values," and of loose community values in Nguyen and Strohl, "Cultural Appropriation."

accountable to my own richer values. I mean to exclude here from the category of value capture those cases where we use external values as proxies and heuristics under full reflective control—when we select, monitor, and adapt those heuristics in the light of our own richer values.

Suppose that I want to get healthier and more fit. By "healthy and fit," I mean something complex and textured and difficult to express—something about feeling good in my body, being more capable of comfortably executing complex physical tasks, and getting rid of this feeling of awkward clumsy brokenness that too much laptop time has left me with. But such inchoate and airy expressions of value are pretty hard to use in the rush of daily life. Beings like us need heuristics—simple and clear rules of thumb to use in the day to day. And I can pick a heuristic, like increasing my step counts, as a quick-and-easy decision procedure to use in my daily life as a way of pursuing that richer notion of health.

But such heuristics are not usually supposed to supplant our full values entirely. We are supposed to use them with the knowledge that they are mere proxies for our full values. They are supposed to serve our dominant values, which means they should be revisable and discardable under the light of our full values. What I really want is health in this richer sense, but I also know that I need an easier target to aim at on a daily basis in order to get myself motivated. So, I start using a Fitbit and just aim at getting step counts. But after a few months, I step back and reflect on my time with the Fitbit. Has pursuing step counts made me happier? Is my body performing better? Do I feel less broken and awkward? Perhaps the answer is affirmative and I keep going with the Fitbit; perhaps the answer is negative, and I abandon it and try some other proxy goal. Perhaps the answer is a qualified yes, and I modify my approach, adding a few more goals to the mix beyond just maximizing my steps. This *controlled use of a proxy value* is not a case of value capture since the externally sourced value does not dominate my deliberative procedure. The dominant value is not the Fitbit's step measures but something else, and this can be seen by the fact that I do sometimes adopt a reflective stance where I decide whether adopting the Fitbit's goals is serving my real values and decide whether to continue or discontinue my use of that simple proxy in the day to day. The Fitbit is not in charge.

Similarly, it is not value capture when I am merely *taking the metrics into account* in the pursuit of my own rich and textured goal. Suppose I want to be a legal activist working for immigration reform. I know that going to a high-ranked law school will be important for getting the influence I need for this kind of work. In that case, I will pay attention to the law school rankings—but they do not dominate my practical reasoning. I may see the instrumental value to going to a highly ranked law school, but I can also trade off that instrumental

value against other things I value. I will not, for example, just go to the highest-ranked school possible. I will use the rankings for my purposes—trying to find a compromise between a well-ranked school that will get me the power I need and a school that will help me learn to do the activist work I want to do. That kind of instrumental awareness of ranking systems is a long way off from a real case of value capture—where somebody's primary goal was, say, simply to go to the highest-ranked law school.

Next: value capture can happen at different points in an agent's life arc with values. Sometimes, an agent has already established their values, and then they come to be *replaced* by some external metric. Other times, the agent does not yet have their own articulated values; they are in the process of figuring them out. But the existence of a prefabricated value offers them a *shortcut* in the process of value deliberation. They can simply adopt a ready-made value instead of going through the slow and oftentimes painful process of figuring out and adjusting their values to their own personality and circumstances. The definition of value capture is intended to include both replacement and shortcut cases.

Finally: value capture can happen at some different loci. One kind of value capture involves the *wholesale* capture of the entire value—such as when you got into this career for joy but came to care only about the money. In wholesale value capture cases, the agent systematically changes how they think of their values; they come to describe their values differently and report them differently. But just as common as these wholesale cases, I suspect, are cases of what we might call *application capture*. In such cases, an external expression of a value does not replace how the agent conceives of their original value—in how they would think about and report their values in the abstract. But the external value dominates how they act by setting the practical criteria in day-to-day applications of their values in particular decisions and evaluations. Say that I, an academic, care about the pursuit of truth, wisdom, and understanding. Across my career, if asked, I would describe my core values using those same terms. But suppose that in the course of my professionalization, the way I apply those terms changes. Now, whenever I try to evaluate the success of my articles, I turn to certain metrics, like the citation rate or the status of the publication venue on some ranked list. And when I evaluate my overall success as an academic, I turn to metrics like my total citation rate or the status of my institution on some ranked list. In that case, it is those institutional metrics, and not the vaguer values I report upon reflection, that effectively dominate my actual actions and self-evaluation. Here, the metric gains dominance by capturing, not the general terms in which I articulate my values, but the more specific application criteria I use when the values hit the ground. The metric fills out the process by which

I determine whether I have fulfilled my core values. And suppose that I guide my actions based on those evaluations: I start writing papers that are more like the ones that have succeeded, in these terms, and start taking actions that might advance my general success, in these terms. Then those external criteria have come to effectively dominate my practical decisions.

There is a crucial difference between controlled use of a proxy as I described it earlier, and application capture. When we *use* a metric as a mere proxy, our richer values are in charge. We will regularly reflect on the proxy from the perspective of our full values, and modify, discard, or adapt that proxy. In the application capture cases, we let the proxy take charge. It functions as the effective practical translator, connecting our abstract expressions of value into specific cases of evaluation—controlling how we apply our values to the world. (Though, at least from my own observations, many cases of value capture start as innocuous-seeming uses of a proxy. I have heard many people say that they put on a Fitbit in order to pursue some other goal, like health or happiness, but then years later they found that they had forgotten about that larger goal—that doing well in the Fitbit's terms had come to occlude all else.) From here on out, I will speak of "values" being captured, for brevity's sake—but I mean to indicate both wholesale value capture and application value capture.

To sum up: value capture does not include every interaction with rankings or metrics. It does not include the controlled use of proxies and heuristics, nor the informational use of metrics. Value capture occurs when an externally sourced value plays the dominant role in practical reason—when it gets put in charge for some domain. This looks like: people who pursue step counts even when it hurts their knees and exhausts their spirit; academics who pursue publications in the highest-ranked journals even when their work feels boring and meaningless; universities that pursue high rankings in the USNWR over richer understandings of education; newspapers that pursue clicks and pageviews over their own sense of newsworthiness and social importance. And, as I have noted: the empirical work indicates that this sort of robust value capture is actually quite common.

Value capture is distinctive because we do not change or adapt the particular externally sourced specification of a value to our particular context. Compare this with other, more open-ended and dynamic relationships we might have to externally sourced values. We often get our first grip on a new pursuit—and its value—with another's help. A friend shows me the wonders of horse riding, the beauty of jazz, the depth of haiku. They talk about what they find meaningful and rich in the activity; they guide me into the actions and attentions that will help me get onto its distinctive value. As Tal Brewer puts it, the values of activities are often obscure to the outsider or novice; it takes a long process of

immersion in the activity to get onto its true value.[7] And we often need help to find our way in.[8] The friend who taught me to see jazz talked me into it—into the particular thrill of seeing a live improvisation. But I suspect she would have been very disappointed if, ten years down the line, the value I found in jazz still precisely mirrored her own. Like any good art friend, she hoped that I would eventually fly on my own wings and sharpen the details of my love of jazz in my own way.

And I did—I used her guidance to find my way in and then slowly began to develop my own relationship with jazz, finding out what thrilled and moved me in the music. This is not a case of value capture; I have used external guidance to get my first grip on the terrain of value in an activity, but then significantly tailored my sense of value in light of my own experiences. Value capture cases are the cases where I internalize, wholesale, an externally sourced value and permit it to dominate my reasoning in its unchanged form. This is where the "outsourcing" metaphor is particularly useful. The harms of outsourcing do not depend on any involuntariness. I can wholeheartedly consent to outsourcing. The harms come from the particular content and nature of outsourced objects—of their inflexibility and prefabrication. If you want a slogan: our values should be tailored to our particular selves and our particular context—but in value capture, we buy our values off the rack.

3. THE PROBLEM OF VALUE CAPTURE

What, then, is the cost of outsourcing one's values? First, to be clear: I am not trying to argue that value capture is always wrong. Value capture, as with any other form of outsourcing, involves a trade-off between efficiency and fine tuning. I think that we are often clear on the benefits of that trade-off but fail to plainly see the costs. My goal is to articulate more precisely the costs.

But it turns out to be rather hard to articulate the problem with value capture. First, what is wrong with getting our values from external sources? It seems utterly naive to think that our values need to spring fully formed from some magical inner place, wholly devoid of social origin. We are deeply social beings, and we often seem to get our values from our culture, our community, our social context. Second, how could value capture undermine autonomy? Many cases of value capture are entirely voluntary and consensual. We know that Fitbit motivates because it presents information in public and shared terms—though the

7 Brewer, *The Retrieval of Ethics*.

8 See Nguyen, "Trust and Sincerity in Art," for a discussion of how trust in others is often required to provide the motivation for attending to difficult or obscure art forms.

full implications of that publicity may not be entirely obvious. Consensual value capture can seem like an aid to autonomy. People often seek out such gamifications in order to overcome weakness of the will. People buy Fitbits or use Duolingo precisely because the gamified structure—in which they are awarded points and levels for progress—makes them more able to get certain things done, like start exercising or learn a language. They are *hoping* for value capture, and they are choosing the effect with some understanding of the basic mechanism. As Jane McGonigal puts it, gamification is a force for good because it can turn monotonous tasks into fun.[9] If value capture can help us overcome weakness of the will, then it helps increase our autonomy and agency. So, what is the harm?

There are at least three ways to think about the potential harm of value capture. First, it might be that *autonomous* participation in the formulation of our values seems good in and of itself—and not just mere one-off consent to a big package, but a fine-grained and ongoing autonomous control of the details of our values. If that were true, then value capture would undermine our autonomous control over our values.[10] Second, institutional values are subject to demands for *hyperexplicitness*, and hyperexplicit values seem unlikely to adequately capture the full richness and subtlety of human values.[11] Third, the kinds of external values we encounter are typically formulated according to the interests and perspectives of large-scale institutions. They are, we might say, *standardized* values. Such values seem unlikely to fit the varying and peculiar interests and situations of particular people and smaller-scale groups.

9 McGonigal, *Reality Is Broken.*

10 Of the three options outlined here, the autonomy option is the one I am most undecided about. While intuitively appealing, developing such an account depends on walking the tightrope between specifying a substantive condition of autonomy, while keeping a grip on the social sourcing of many of our values, even in the most autonomous cases. An earlier, and much simpler, version of this paper attempted to offer an analysis of the harms of value capture in terms of a violation of autonomy, but that version of the argument could not survive contact with the insights from the literature on relational autonomy, especially from feminist critics of idealized pictures of autonomy. See Buss, "Valuing Autonomy and Respecting Persons"; Christman, "Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves"; Khader, "The Feminist Case against Relational Autonomy"; Superson, "Deformed Desires and Informed Desire Tests" and "Deferential Wife Revisited"; and Westlund, "Selflessness and Responsibility for Self." It remains to be seen whether there is some more refined version of the autonomy worry that can be made in light of this discussion.

11 I am exploring this possibility in other work. See Nguyen, "Value Collapse," for a discussion of the possibility that hyper-explicit values represent a bad epistemic attitude toward the world of value—that they discourage exploration of the space of value by making it easy to dismiss new candidates for value.

A full account of value capture would need to address at least these three approaches, and I hope one day to provide such a full account. Here, I can only take a first step. I will concentrate on the issue of *standardized values*—in part, because I think it highlights the unique problems of value capture.

Here is the worry in a nutshell. Value self-determination is important for all sorts of reasons. Here is one: value self-determination yields values that are finely tuned to our particular context. By substantively participating in the detailed process of formulating our values for ourselves, we can get values that nicely fit our particular circumstances—our individual psychologies and phenomenology, our group culture, our local context. Value capture intrudes on that process of value self-determination, *substituting prefabricated and standardized values for finely tailored ones*. Note that this is not an argument that autonomous value formulation is a good in and of itself, but rather an argument that substantively participating in the process of shaping one's values is instrumentally good in that it yields better, more finely tailored values.

Here is an example from my own life. For the first two decades of life, I avoided most physical activity. I had an incredibly simplistic conception of the value of exercise. I thought that exercise was basically pounding out some miles on a treadmill to burn some calories. Eventually, I came to see the vast and varied joys of athleticism. But in order to get there, I took a long meandering journey through many different sports, each of which paid off in profoundly different ways. Long-distance running turns out to be Zen-like and calming. Trail running requires more attention but offers this thrilling sense of reactive flow to the difficulty of the trail. Deadlifting is brutal and intense, a pure shot of grueling focus. And rock climbing turns out to be a fascinating fusion of bodily aesthetics and puzzle solving, where you solve thorny movement puzzles through elegant motion.[12] And even inside one of these activities, there is not some singular value on offer. Rock climbing can be pursued in radically different ways, each of which offers very different rewards. You can seek out thrills and risk; you can do easy climbs in rapturously beautiful terrain; you can focus on finding climbs with graceful movement; or you can go for gruelingly athletic climbs on a cave roof just eight feet off the ground. Each of these different ways of valuing rock climbing suggests a different way of approaching it, which in turn yields richly different textures of activity.[13] This is a process of *exploration,* where you try things out, figuring out how they fit with you, and

12 For more on the aesthetic qualities of movement in rock climbing and other games, see Nguyen, *Games* and "Arts of Action."

13 This description has been deeply influenced by Tal Brewer's account of how the formulation of the value of an activity and the way we do an activity form a feedback loop as we explore and refine our understanding of the activity (*The Retrieval of Ethics*). Agnes Callard

changing around your approach in response, seeing how it goes in an ongoing loop of feedback and adjustment.

The worry then is that when you are value captured by a Fitbit, you do not go through that process of exploration and fine-tuning. My claim here is not that one puts on a Fitbit and is automatically value captured. One could simply use a Fitbit as a data-gathering system to pursue one's own values. But Fitbit does present its step counts as a *score*. It is a gamified system, which openly employs design features from games.[14] A Fitbit does not force value capture, but it certainly invites it.

We might even call this an extended value system. Some philosophers have been very excited to claim that our minds are extended beyond our bodies—that our minds can include various technologies as parts of their internal functioning.[15] Most of the discussion of extended mind has focused on adopting, as part of our extended mind, various value-neutral cognitive resources—like using a notebook or Google Docs as an extended memory. Some of the discussion has gone so far as to suggest that we can extend our mind to use various technologies as part of our emotion regulation system, such as Joel Krueger's suggestion that we use our portable music devices for mood regulation.[16] My suggestion is one further step: in some cases, a *standard for evaluation* is embedded in a technology, as in Fitbit's step counts or Twitter's likes. When we integrate that technology into our cognition, our extended mind now includes a *value system* which was created externally, and which is sustained through external technologies.

Of course, one might respond, we get our value systems from external sources all the time— from our parents, our community, our culture. But, in the unproblematic cases, we can use external values as a starting seed, which we can adapt and tailor to ourselves. The worry is that we simply plug in these external values and use them as is. In particular, some external values can *resist further tailoring*. This is especially likely when adopting a particular prefabricated value is appealing precisely *for the standardization*. Then we will be quite tempted to leave them as they are, or else lose out on the promised efficiency.

---

has also written on such proleptic ends, though her account adds a requirement that the process is triggered by a desire to become the kind of person who so values (*Aspiration*).

14 See Nguyen, "How Twitter Gamifies Communication," for a discussion of how certain user interfaces can present metrics as scores, and thus as forms of evaluation.

15 Alternately, if we want to avoid the endless tussles about what exactly is the line between mind and not mind, we can use Kim Sterelney's locution: that various technologies are scaffolds for agency ("Minds"). Either locution takes us to the same worry.

16 Krueger, "Music as Affective Scaffolding."

Value standardization is like any other kind of standardization: we gain in efficiency but in exchange for giving up localized tailoring.[17]

To really understand the problem here, we need to have a good grip on why we should want to tailor our values to fit. Elijah Millgram offers a useful account of how we adapt and improve our values in *Practical Induction*.[18] He is not talking about abstract, generic renderings of value, like, say, "happiness" or "flourishing." He is interested in the specific, grounded articulation of our values and goals by which we conduct our day-to-day lives, such as a runner's pursuit of a better marathon time, or a filmmaker's interest in playfully subverting genre conventions, or a philosopher's interest in writing deep, rigorously argued papers. Crucially, says Millgram, we do not derive these specific articulations of value by deriving them from some abstract specification of the good. Rather, we acquire our particular values and goals via a process of practical induction. We *try on* particular goals and values for a while. We might enter a profession and try on the goals associated with that profession: a literary fiction writer might start caring about achieving realism of character and setting; a Montessori teacher about fostering autonomy in very young children. And then the person gets *feedback* from the experience of living life under that particular value system. They might get positive feedback, like feeling engaged, happy, or interested; they find themselves savoring the details of their life. Or they might get negative feedback: they feel bored, listless, disengaged. This is feedback about the *fit* between the values we have adopted and the particular circumstances of our lives: our personality, our culture, our place in society. To flourish, we need to be sensitive to that feedback, and use it in fine-tuning our values to fit.

Millgram's own discussion focuses on large-scale value shifts which accompany things like, say, having a midlife crisis and changing careers. But his argument leaves room for smaller-scale adjustments in the articulations of our values. Say I start rock climbing and take up the most obvious standard of success in that hobby. There is a generally agreed upon difficulty scale for climbs; most new climbers just start by trying to advance up that scale. Some climbers flourish under that goal; others do not. When I focused on advancing on the difficulty scale, I found myself miserable, tormented by my sense of inadequacy and my inability to progress. My climbing days were filled with exhaustion and dread. So, I began to change my sense of my goal in my climbing, started pursuing a slightly more personal vision. I started looking for the most elegant,

17 Bowker and Star, *Sorting Things Out*.

18 Millgram, *Practical Induction*. Millgram has since developed some of the ideas in *Practical Induction* further, most notably in an argument that boredom and disengagement is a signal that one's values and chosen roles are a bad fit ("On Being Bored out of Your Mind"). My treatment here relies on both of his discussions.

interesting climbs, and my goal became to climb them as delicately and with as much control as I could. And under that goal, I flourished: I became more constantly sensitized to the details of my movement and more at peace with simply enjoying a bit of lovely climbing; rock climbing trips now left me feeling restored and happy.

So here is a first pass at the *tailoring argument* for the harm of institutional value capture. If Millgram is right, then we will flourish when we have the capacity to adjust and tailor our values in light of our rich experience of the world living under them. When we tailor our values to ourselves in light of those rich experiences, then our values will be better fit to promote our flourishing, as the very specific people we are, in our very specific circumstances.

Perhaps you do not like the references to some ill-defined sense of "human flourishing" or "well-being." We can put the same thought in less mysterious terms. When we adjust our values in light of our rich emotional experience of the world, then those adjusted values will be better suited to support a more emotionally positive life. If we adjust our values, taking interest and engagement as a positive sign and boredom and ennui as a negative sign, then our values are more likely to give us a rich, interested, and engaged life, rather than a bored and listless one. But in institutional value capture, we do not adjust our values in light of our particular experiences. We take values as provided by some large-scale institution and live under them as given. Those values will have been formulated to take deeply into account various institutional interests: like the ability to be counted in a reliable way across a large institution and the ability to be readily aggregated in an institutional bureaucracy. They will not have been formulated in light of the rich feedback of how our particular lives have gone when we live under these values. In value capture, we adopt values that have been formulated in a way that is *insensitive to* and therefore *less able to support* our rich, subtle, and personal emotional experiences.

The first pass emphasized the problems of value capture for the individual. This reflects Millgram's own version of the argument, which emphasized individual values and individual phenomenology. But we can also easily extend the argument to encompass group value capture. Groups, too, have particular articulations of their values.[19] This can look like, say, the community of analytic philosophers' value in rigor or the creative writing community's value in personal expression.[20] The contemporary community of improv comedians, for example, have come to put a strong value in collaboration via automatic agreement.

19 See Nguyen, "Monuments as Commitments," for a further discussion of group valuing, and how public art may function as an alternative to metrics for expressing group values.

20 For an excellent discussion of the scientific community's interest in enlarging the collective data supply, see Strevens, *Knowledge Machine*.

The core rule is "Yes, and . . .": you always accept other people's suggestions and build on them. This expresses a value, we might say, of *radical acceptance*—of never refusing ideas, and always integrating every proposal and building upon it. This value works extremely well in the context of improv comedy. And the precise articulation of this value has obviously evolved over the years through trial and error in countless acts of improv comedy. But its success is context-dependent. (Imagine trying to center such a value in analytic philosophy. Analytic philosophy, one might think, is a century-long social experiment in the value of harsh criticism and the radical refusal to accept anything.)

So long as there are accessible signs of a groups' flourishing or of a community's well-being, then Millgram's account of practical induction should also apply to the development of group values. Such group-level value tailoring is unlikely to center internal emotional phenomenology as strongly. But groups can tailor their values in response to their particular nature and context. Jane Jacobs offers a particularly vivid example of how we might tailor a specific value to a specific context.[21] Dwellers in dense urban environments, she says, have learned to cherish *privacy* in a way that suburban and rural people do not. So much of one's life is conducted in dense public environments, that city dwellers have developed a profound devotion to maintaining privacy: of not making unnecessary eye contact, of not intruding into nearby conversations. Valuing a certain kind of eager "friendliness"—easy eye contact, being willing to start conversations with anybody at anytime—makes perfect sense in lower population density areas without that constant press of humanity. But in a dense city, without that collective devotion to the practice of privacy, city dwellers would be utterly overwhelmed by constant social interactions and demands.

To sum up: it is good for agents—individual and group—to tailor their values to their particular context. Those values will be better suited to support the well-being and flourishing of individuals, groups, and communities by being adapted in their formulation to the particular nature of the agents and to their particular context. Value capture interferes with that tailoring. It does so even when the value capture is the result of a fully informed and consensual process since the problem lies in the *content* of the values and not in the bare fact of their voluntary adoption.

Here is another way to put it: value capture, even when consensual, involves a low degree of granular control over the details of the contents of one's value. It puts you in the same relationship with your values as you have with, say, your iPhone's end-user license agreement (EULA). When you clicked to sign a EULA, you did, technically, consent, and you are, technically, responsible. But

21 Jacobs, *The Death and Life of Great American Cities*.

you only have one binary choice: accept the whole package or not. When we permit ourselves to be value captured by institutional values, we have the same low granularity of control over our values: we either accept the whole package, or not. You cannot get control over how your Fitbit counts steps or how the edifice of higher education counts citation rates and impact factors.

This low granularity arises directly from the core functioning of large-scale collective values. To better understand why, let us look at the processes that drive the creation of institutional metrics.

4. METRICS AND THE STANDARDIZATION OF VALUE

Let us focus on value capture by institutional metrics. I think this is the starkest case of value capture and a good starting point for thinking about other forms of value capture. My goal in this section is to make clear why institutional metrics resist value tailoring as part of their essential functioning.

One response to the case studies I have offered so far—the *USNWR* case and the UN's human rights metrics case—is that the particular metrics are bad. Perhaps there is nothing wrong with value capture per se; it is just that we need to pick good metrics. But I will suggest that we are unlikely to find any institutional metrics that are good to take on as individual or small-group values. Metrics are formulated to serve certain key institutional interests—to work at large scale—and they need to be relatively inflexible to play their role. Institutions want metrics that are narrowly specified, standardized, and inflexible. Precisely what makes a metric good in the institutional context will make it problematic to internalize as a value for individuals and small-scale communities.

Here, we can turn to a rich and useful empirical literature on the place of quantification and standardization in bureaucracy and political life. Here, we are the beneficiaries of decades of empirical study of quantification culture, performed across a number of fields, including history, sociology, anthropology, and communications.[22] What follows may sound familiar to some philosophical ears; the empiricists I will be discussing are often working in a Foucaultian mode. The field, in particular, has been highly influenced by philosophical figures such as Ian Hacking, Bruno Latour, and Martha Nussbaum.

A foundational work here is Theodore Porter's 1995 history of quantification culture, *Trust in Numbers*.[23] Porter is particularly interested in how quantified forms of justifications, like the cost-benefit analysis, came to dominate politics

22 The study of quantification culture is often associated with the interdisciplinary field called Science and Technology Studies.

23 Though Porter is a historian, he was significantly influenced by Ian Hacking's work in the philosophy of science on the formation of categories and measures.

and management. He is not arguing that quantification is always bad. Rather, his goal is to get clear on the relative advantages and disadvantages of qualitative and quantitative ways of knowing. Porter argues that qualitative ways of knowing are nuanced and context-sensitive. But qualitative information is difficult to manage en masse and difficult to transfer across contexts. Qualitative evaluations usually require significant shared background knowledge to adequately interpret. When we transform information from a qualitative to a quantitative format, we strip off much of the nuance, texture, and context-sensitivity. By doing so, we create a *portable* package of information, which can be easily sent across contexts and understood by people with little shared background.[24] Quantified evaluations can be easily transmitted between people with little shared background, precisely because they have been stripped of context-dependent features. And quantification isolates the more invariant parts of that information so that the results can be readily aggregated. For this reason, quantitative methods are preferred by large-scale institutions, which must pass information across many levels of hierarchy—between distant administrators with low shared context.[25] In other words, quantifications are preferred in large-scale institutions precisely because of their narrowness and their context-invariant stability.

And quantitative evaluations themselves vary according to their nuance and context sensitivity. Once, land in England was measured in *hides*. A hide is the amount of land required to support the average family. The hide is a measure that highlights a highly relevant functional quality. The acre is a measure of land size, rather than land function. Similarly, says Porter, older Polish land measures varied by soil quality, so a given unit of land would approximately represent a similar productive value.[26] When a ruler attempts a fair distribution, the measure they use will determine which quality is evenly distributed—in this case, land size versus land functionality. Think about the difference between, say, a king's giving each of his soldiers a hide of land, versus his giving them each ten acres of land. One might think that the hide is a superior measure of functional worth, and so a vastly preferable measure for providing fair compensation. But hides are highly variable in size, and determining what counts as a hide requires the application of detailed local knowledge. A hide in a fertile river valley is smaller than a hide in a desert. Hides also vary depending on

24 This notion of "portability" as the center of quantified information is alive in more contemporary work in this space. In Sabina Leonelli's crucial work on the philosophy of data, she defines "data" as information that has been prepared to travel to new and unexpected contexts ("What Counts as Scientific Data" and *Data-Centric Biology*).

25 Porter, *Trust in Numbers*, 3–86.

26 Porter, *Trust in Numbers*, 24.

local weather patterns, game animal migration patterns, and more. The hide is a measure that can really only be effectively managed at the administrative periphery—by locals, who know their environment and its inhabitants' typical needs and usage patterns. The hide is impossible to administrate from any sort of distant bureaucracy. So, says Porter, when we shift from small, local, distributed governance to large-scale centralized governance, we inevitably shift from informationally rich—but difficult to manage—measures like the hide to more standardized, but informationally impoverished, measures like the acre.

James Scott calls this the state's view of the world. By "states," Scott means any large-scale institution, including governments, corporations, and the emergent networked institution of globalized capitalism. States, says Scott, can only manage what they can see, and they can only see that information which has been rendered into a form which can be processed bureaucratically—information that has been standardized and quantified. States can only see those parts of the world which have been rendered *legible* to them.[27]

Student grades provide a familiar example. In the modern educational environment, student grades are almost always quantified. But there are other modes of educational assessment. Imagine an educational environment where we only offered qualitative evaluation of their students' work, like written feedback describing its good qualities and its problems. Such evaluations can easily pivot to address different dimensions—like the writing clarity, the originality, the argumentative clarity—without any demand to compress that all down to a single dimension of evaluation. Such evaluations can also be tailored to each student's own particular goals. I might give very different suggestions to a nursing student interested in the practical implications for their work than I would to, say, a future lawyer or future journalist. If our goal is simply to educate the student, we do not necessarily need to provide an overall rating of all our students on some single common scale.

But in our actual world, we must offer a quantified measure of each student's success—a measure which permits us to instantly compare any student with any other: their grade. This quantified ranking of students is extremely useful to administrators. All of a student's efforts in a class can be expressed in a single number. This also enables a further aggregation: all their class grades can be averaged to generate a single number, which represents their entire educational career—a GPA. And the existence of GPAs is enormously useful for the project of administrating a large-scale educational bureaucracy. They enable all kinds of fast, easy, and objective-seeming manipulations. An admissions officer can arrange the data from every single student application into a spreadsheet and

27 Scott, *Seeing Like a State*, 11–83.

quickly sort them by GPA. They can create an automatic cutoff point below which student applications are automatically discarded. Sets of student GPAs can be aggregated in order to yield a single number that can be used as a metric of performance for a particular teacher or a whole school district.

In their study of the history of American grading, Jack Schneider and Ethan Hutt argue that standardized grading schemes were implemented to make grades more legible and usable to administrators and employers.[28] Before grading, there was no communicative "shorthand." Evaluations required intimate communication between teacher and student. Early systems of grading were "low-stakes" affairs; they were set up differently in different schools and built to encourage student learning. But the modern system of grading serves not a pedagogical purpose but an organizational purpose. It enables students to easily transfer between different institutions. Perhaps most importantly, it standardizes a product for future consumption on a market. Standardized grades make possible standardized educational certificates, which are extremely useful for potential employers. It was administrators and employers who "placed a premium on readily interpretable and necessarily abstract grading systems."[29] Qualitative evaluations of student might be nuanced and context-sensitive—but they are illegible to the large-scale administrative institution.

Finally, these various procedures—data collection, transformation into standardized inputs, and aggregation—need to be codified into a set of policies that can reliably executed by very different people. Large-scale institutions need to train up people from different backgrounds to perform the same sorts of tasks. And their performance needs to be assessable and auditable by others—where those auditors also come from different contexts, and their audit procedures themselves are subject to the same demands of explicability and transmissibility.[30] These various procedures need to be *standardized*. That means that the inputs and processing rules of these procedures need to be regulated across many contexts.[31]

We can draw from this mess of observations some underlying themes. Institutions share a basic functional interest, inherent to the functioning of large-scale administrative systems. They need to manage information across a vast domain. This need arises intrinsically from the need for an institution to function as a coherent whole. Notice here that I am not presuming that the institution has some interest in controlling or manipulating individuals. Even

28 Schneider and Hutt, "Making the Grade," 203.

29 Schneider and Hutt, "Making the Grade," 217–18.

30 Du Gay, *In Praise of Bureaucracy*.

31 Bowker and Star, *Sorting Things Out*, 13–16.

the most well-intentioned of organizations—like, say, a charitable nonprofit—has this same functional interest in information management. The interest arises from the basic conditions of coherent group agency, as instantiated in a policy-based, centralized bureaucracy.

This functional interest is served by two standard mechanisms—quantification and standardization. Institutions need to render the world into a format legible to large-scale institutional information processing procedures. So, institutions need information in quantified and standardized format. Because of their institutional function, these mechanisms—quantification and standardization—tend to share some specific features that make them problematic to internalize as personal values. First, quantified metrics are *narrowed* by design. Only certain things count. Institutional measures need to be usable across different contexts. This requires that the measures leave aside highly context-dependent forms of understanding and focus for their inputs on context-invariant qualities. As Scott says, the narrowness of the metric creates a narrowness of institutional vision. Institutions can only see, process, and act on parts of the world that are counted by their metrics. Anything that does not impinge on those metrics is invisible at an institutional level.

In value capture, we internalize those narrowed metrics, thus narrowing our values. And insofar as our values drive our attention, then the value captured will be subject to an analogous effect to Scott's narrowed institutional vision. It is not that we literally do not see things that fall outside our narrowed values, but we will not devote much energy to them or dismiss them as unimportant. Think here of the businessperson who thinks that only money matters and who immediately dismisses from mind any unprofitable ventures—like art or philosophy.[32]

Next, such institutional metrics typically present values in highly explicated, *finished* form. They resist reinterpretation. Preinstitutionalized values are often expressed in an open-ended manner. A concept like "health" or "fitness" or "a good education" admits of different interpretations. Different people may work out their own interpretation of what counts as a good education—and so evaluate their understanding of the term. You want to know more useful things; I want to indulge my sense of curiosity—both are viable understandings of what one might want out of an education. But step counts and law school rankings do not admit of such variability. The method of assessment is rigid. Says Porter, the process of quantification is useful to large-scale institutions, in significant part, precisely because it reformulates information so as to remove the need

32 I further develop this line of thinking in Nguyen, "Value Collapse," which explores the possibility that overly explicit articulations of value can narrow our attention and exploration.

for interpretation.[33] Standardization is required for informational portability—and standardization requires rigidity.

And those off-the-rack values usually come *embedded* in institutional infrastructures, institutional language, and mass technologies so as to resist further tailoring. We do not have the power to fine-tune the innards of such institutional values. We cannot tinker with the way Twitter counts likes, nor adjust with *USNWR*'s ranking algorithm. They are hardwired into external systems. This rigidity and uncustomizability of measures and metrics is no accident. It is essential to their institutional function. Standardization enables easy communication and ready aggregation—but to do so it must resist individual customization.

To summarize: institutional metrics are designed according to *alien interests*.[34] They have, in fact, been fine-tuned and adjusted—but to satisfy interests that are not our own. I am not presupposing here that institutions must have malevolent intent, like an interest in domination, control, or power. To put Scott's discussion into philosophers' terms: all we need to attribute to an institution is a basic interest in *agency at scale*—an interest in gathering information about the world, managing that information, and using it to inform actions. But the scaled-up nature of bureaucratic institutions imposes certain distinctive requirements on that information-gathering process. Institutional metrics are typically formulated to fit the demands of scaled-up informational agency: for easy recording in institutional memory, for transmission across bureaucratic layers, and for manipulability by institutional methods. When we internalize institutional values, we are letting such interests play a powerful role in the formulation of our own values. Value capture gets us to take an institution's eye view on ourselves—to evaluate ourselves and our activities in institutional terms.

We have much to gain by fine-tuning our values, fitting them with our psychology and world. Institutional metrics are tuned, not to an individual's rich and particular experience of the world or a small community's particular context, but to the needs of information processing at a mass scale. In value capture, by institutional metrics, our values become rigidly tied to an external

33 Porter, *Trust in Numbers*, 21–29.

34 Owens and Cribb make a similar point in their analysis of Fitbit technologies ("My Fitbit Thinks I Can Do Better," 32–35). They distinguish, however, between procedural autonomy—which involves internal deliberative processes—and substantive autonomy, which involves one's ability to actually act on and bring to fruition one's decisions. They say that by and large, self-tracking technologies like Fitbit may aid procedural autonomy but cannot aid substantive autonomy since such technologies cannot fix large-scale social inequities. My argument is that such technologies also significantly undermine procedural autonomy.

expression. That rigidity arises, in significant part, from the institution's interest in large-scale informational management. Metrics, by design, resist attempts at digestion and customization by the agent—and they usually come embedded in large-scale institutional infrastructures that make them even more inflexible. Their alienness resists adaptation.[35]

5. The Seductiveness of Metrics

Of course, we often have to use such metrics when we work within, and next to, institutions. But we could use them while also keeping them at emotional arm's length. We could employ them in our reports and our requests for funding, but only as the trade language of bureaucracies. Why might we ever take the further step and internalize them? The answer comes in several stages.

First, quantifications, in and of themselves, are seductive in their clarity and crispness. Many people seem to trust quantified data simply because it is quantified. And we should certainly trust data when it has been generated using reliable methods. However, the mere quantified format itself often seems to generate trust, regardless of quality of the underlying methodology. But obviously, mere presentation in a quantified format does not offer any guarantee of reliability. So, insofar as we trust from the bare fact of quantified presentation, then that trust is unwarranted. And Porter, Merry, and Espeland and Sauder provide evidence aplenty that bare fact of quantification actually does, in fact, generate such unwarranted credibility. To put in the contemporary parlance, the excessive credibility given to quantified data counts as a form of epistemic injustice or epistemic oppression.[36] It harms those who are unwilling or unable to present their information in such quantified form, preventing them from being appropriately recognized as sources of information. And insofar as

35 One might ask what relationship this view has with various forms of alienation critique. Though my analysis here is obviously similar, in spirit, to the general themes of alienation critique, I avoid use of the term "alienation" because my analysis here differs, in key respects and in many details, from traditional alienation critiques. As Rahel Jaeggi says, many forms of alienation critique involve views of the alienated agent as divided against themselves, as unable to identify with their work, as diffident and depressed (*Alienation*). But the value captured agent can be wholehearted (think of the capitalist all-in for money), fully identified with their work, energized, and motivated. They are not divided against themselves; rather, they are simplified, where that simplification has been guided along institutional lines. Notice, furthermore, the difference between my analysis and the traditional Marxist alienation critique. It is possible to be value captured by a fully socialist bureaucracy. Here, I am aligned more with Scott's particular version of neo-Foucaltian critique than with Marx. For Scott, both globalized capitalism and centralized communism share an interest in rendering the world legible into the terms that they can process and act upon.

36 Fricker, *Epistemic Injustice*; Dotson, "Conceptualizing Epistemic Oppression."

quantified data tends to emerge from certain sorts of institutions, then those institutions themselves are the beneficiaries of epistemic injustice.

Why might the mere presentation of information in quantified form invite such excess credibility? One familiar suggestion is that numbers carry with them, through their association with the sciences, an aura of authority. I would like to suggest another mechanism: our use of *cognitive fluency*, a phenomenon well-documented by cognitive psychologists. Cognitive fluency is the "subjective experience of ease or difficulty with which we are able to process information."[37] As it turns out, we often use cognitive fluency as an epistemic heuristic. The easier an idea is for us to comprehend, the more likely we are to accept it as true. This is sometimes a useful shortcut. We are typically better at processing information in domains where we have expertise, so ease of comprehension is somewhat correlated with correctness. But the heuristic is far from perfect, as cognitive psychologists have amply demonstrated. First, we seem more willing to accept an idea simply because it is familiar. Second, we are more likely to accept claims presented in a more legible font. But, obviously, the bare fact of repetition or graphic legibility has no direct bearing on truth. In both cases, using a cognitive fluency heuristic results in a mistaken degree of trust.[38]

We should, then, experience a cognitive fluency effect with anything with which we are familiar. And we are extremely familiar with numbers. They are the universal abstraction. Information presented in quantified form thus wears an extremely familiar face. So, the fluency heuristic can lead us astray with quantification, just as it does with fonts. This offers an explanation for the unwarranted credibility of quantified values. Fluency may bring somebody to accept a quantified evaluation of value over a more inchoate one—like accepting the *USNWR*'s clear presentation of a ranking over one's own internal sense of, say, fit with a law school's culture. And insofar as the quantified presentation is more likely to emerge from external and institutional sources, then the fluency effect gives an unwarranted credibility boost to such sources.

But it is not just that metrics are quantified; it is that they are *standardized*. Once our values are standardized, then we can easily explain our actions and justify our decisions to others. Metrics offer an engineered communicability for values. This engineered communicability grants a further credibility advantage to claims made in the evaluative language of those metrics. After all, our ability to make ourselves understood to others can be a sign that our own

37 Oppenheimer, "The Secret Life of Fluency," 237.

38 Reber and Unkelbach, "The Epistemic Status of Processing Fluency." I offer a sustained discussion of how cognitive fluency plays into our attraction to seductively clear systems in Nguyen, "The Seductions of Clarity."

understanding is good.[39] And metrics are, by their very nature, easier to understand across contexts. But there is a gap between communicability and epistemic worth, and that gap can be exploited. As Porter makes clear, institutional metrics trade away informational nuance, richness, and contextual sensitivity for the sake of easy portability across communicative contexts. Metrics, then, in virtue of their basic institutional function, also function to precisely exploit the gap between communicability and epistemic worth. When we standardize metrics, we engineer in broad-ranging comprehensibility by removing contextual nuance. Because the very act of removing contextual nuance increases communicability and cross-contextual comprehensibility, metrics—by their essential nature—invite excess trust.

Quantifications can also be seductive because they offer us the pleasures of *value clarity*. When we internalize them, our value landscape becomes simpler and easier to navigate. We are tempted to take them on because they offer us hedonistic rewards in exchange for simplifying our values along certain lines. This line of argument draws on my account of the motivational structure of games, which I have developed at length elsewhere.[40] In games, we take on artificially constructed goals. In ordinary life, our goals and values are often complex and subtle. It is often hard to explain our values clearly, hard to adjudicate conflicts between values, and hard to figure out if we have actually achieved what we value. But in games, values are easy. They are clearly articulated, with explicit criteria for application. In games, we know exactly what we should be doing, and exactly how well we have done. Games offer us a momentary refuge from the nauseating complexity of real-world values. They are an existential balm.

This offers us a second mechanism for the seductiveness of quantification. We can gain a hedonic reward for internalizing simplified values. When we come to value a simplified goal in a non-game activity, we bring the pleasures of value clarity into the real world. Our purposes become clearer, our degree of success becomes more obvious, and our achievements become more readily comprehensible—and it becomes easier to compare and rank our respective achievements. But to get those pleasures, we need to simplify the target. And this helps to explain why it can be so tempting to internalize institutional metrics. Metrics are narrowed and finished. When we internalize such clarified metrics, we can eliminate our struggles with the ambiguity and complexity of our values. Metrics may

39 I am relying here on the literature from the philosophy of science's investigation of understanding. According to the standard account, one of the signs of understanding is the ability to communicate that understanding to others (Strevens, "No Understanding without Explanation"). I offer a discussion of how engineered simplicity can hijack our sense of understanding in Nguyen, "The Seductions of Clarity."

40 Nguyen, "Games and the Art of Agency" and *Games*.

not have been explicitly made for gamification, but the institutional pressures on the generation of metrics make them function as pleasingly game-like goals.

And the value clarity effect becomes even more powerful when that clarity is standardized. After all, the existential burden of our complex values is not merely a personal affair; we have to deal with the buzzing tangle of everybody else's values too. Navigating this overwhelming plurality—understanding other people's values and explaining our own—can be grueling. There is, so often, a vast gap between our values. Try explaining to another person your profound love of some weird old comedy, or why a sour cabbage casserole makes you feel so comforted on the bleakest of days. Try explaining why a particularly acid passage of Elizabeth Anscombe's fills you with such glee, or why you never quite got along with running, but rock climbing makes you feel so amazing. Sometimes we can make ourselves understood, but often we cannot. So much of our sense of value arises from our particular experiences—the long life we have led, our twisty paths to self-understanding and world-loving—that explaining the whole mess to others is often beyond our capacities.

But institutionalized values offer us an experience of social value clarity. If an institution offers us a prefabricated metric for some value, and we collectively internalize it, then we will make easy sense to each other. Perhaps that prefabricated metric is citation rates, or Twitter followers, or GPA, or your university's ranking. In any case, once we internalize that value together, much of the existential friction of social life suddenly disappears. Metrics create a common currency for justification. I no longer need to struggle to explain my way of valuing to others, or to understand their way of caring about the world. Justification becomes easy because metrics can function as a preengineered system of aligned value. Metrics offer, not just a personal form of value clarity, but *social value clarity*. When we converge on the same simple, public value system, it becomes so much easier to communicate our values and our justifications. We have gone on the same value standard.

Let us take a step back. Is it some wild accident that institutional metrics turn out to be so seductive? I suspect not, though I can only offer a brief sketch here. Rational agents often need clearly articulated policies to function—including policies about what our goals are and how we are to evaluate our progress toward those goals. Clearly articulated policies ensure reasonably fast decision-making that is consistent over time. As Michael Bratman argues, such policies play an integral part in our being able to maintain coherent agency over time.[41] Policies are desirable for large-scale institutions for similar

41 Bratman, *Intention, Plans, and Practical Reason* and *Shared Agency*; Holton, *Willing, Wanting, Waiting*; Andreou, "Coping with Procrastination."

reasons, since institutions also need to ensure relatively quick and consistent decision-making across a large and scattered structure, in order to enable cohesive collective action.[42]

But the nature of large institutions requires that we heighten the clarity and explicitness of those policies in order for them to function consistently across the whole. The policies I set for myself can hinge on my own peculiar sensitivities and ways of understanding the world. A coherent policy for me is: "Exercise every day until I start to get that pleasant, warm, cheerful feeling." This works for me because I can consistently recognize that pleasant warm feeling. (Another coherent policy for me is, "Cocktails before 6:00 PM only when I really, really, really need it.") But such policies will not work for large-scale institutions because criteria like "a pleasant, warm, cheerful feeling" cannot be written into institutional policy, nor could they be reliably applied by different people across the institution. Institutional policies need to be hyperexplicated so that they may be executed by a wide variety of people hired from a variety of backgrounds. They need to be, to adapt Porter's language, procedures that are *portable* between many contexts. In order to function in institutions, policies need to be easier to apply—and so they can be appealing to internalize. It is very easy to act clearly and consistently when we adopt such hyperexplicit policies. However, in adopting them, we are giving up on the kinds of policies that hinge on sensitivity to subtle internal phenomena.[43]

6. CONTEXT LOSS

Let us step back and summarize the action so far. Values, I have argued, benefit from being tailored to an agent's particular context. In individual cases, that context can involve all the particular details about the person—their personality, their subtle emotional responses to the world. In group cases, that context can involve details about the particular people who make up the group, or the group's ambient culture. It can involve the kinds of subtle considerations that are only adequately comprehensible to a particular community who have gone through a particular set of struggles together.[44] And in all cases, it involves

42 This comment relies on the extensive recent literature on group agency, including List and Pettit, *Group Agency*; Gilbert, *Joint Commitment*; and Rovane, *Bounds of Agency*.

43 For a further discussion of problems with the explicit policies in group agents, see Nguyen, "Games and the Art of Agency."

44 For an excellent recent overview of standpoint epistemology, see Toole, "Demarginalizing Standpoint Epistemology." For a discussion of the tension between standpoint epistemology and the demands of bureaucratic transparency, see Nguyen, "Transparency Is Surveillance."

particular details of the specific context of the agent: the location, the surrounding culture, the environment.

The argument here is an instrumental one, and epistemic in nature. I can easily imagine very different accounts of the harm of value capture. One might wish to argue that autonomous control over one's value was a good in and of itself and that something was intrinsically wrong with ceding that control via value capture. And this argument might also be a good one. But that is not the argument I am making here. Here, I am arguing that fine control over the expression of one's values is instrumentally good. It promotes well-being and flourishing and other such ways of indicating a good life in individuals and communities. It does so because substantively participating in the process of adjusting and fine-tuning one's own values yields more nicely tailored values. Notice this argument works irrespective of whether or not we conceive of the value capture process as consensual or voluntary. It has to do with how much one substantively tailored that value to one's context. One may have voluntarily undergone value capture, but in so doing, one has withdrawn from the process of finely tailoring one's values to fit.

And fine control leads to better well-being and flourishing for epistemic reasons. Individuals and smaller-scale groups have better access to the details of their specific context. There is a useful analogy here to a discussion about the epistemic value of democracy. One reason that democracy is important, one might think, is that self-determination is an intrinsic good. But another reason that democracy is important is that, when appropriately structured, it is the best way to integrate the *epistemic access* of the governed. This is the epistemic defense of democracy. As Helene Landemore puts it, epistemic democracy functions well when it employs a deliberative process that takes into account the specific details known to the relevant communities. Democratic deliberation, done properly, is sensitive to the special understandings of the deliberating citizens. Importantly, Landemore's argument is not that democratic participation is an intrinsic good or constitutive of an authoritative government. Landemore's argument, rather, is that democratic participation is instrumentally good since it yields laws and policies that better fit the circumstances and take better advantage of the various perspectives, expertises, and understandings of the entire citizenry.[45]

45 Interestingly, Landemore does not consider the problems of scale (*Democratic Reason*). This is worth a wholly separate discussion, but I can briefly say: Landemore's argument presumes that the process of democratic deliberation will always preserve the knowledge of the participants and aggregate them. I think there is significant reason to be skeptical about that. I have offered some reasons to be skeptical in my discussion of the epistemic problems with public transparency metrics (Nguyen, "Transparency Is Surveillance"). I think

The tailoring argument works similarly. When an agent tailors its own values, this yields an instrumental good for epistemic reasons. The agent has more access to their context—their psychology, culture, the local details—and so can tailor their way to better-fitting values. Value capture involves adopting values from an external source—typically a massive institution. Such an external source has far less fine-grained access to the local details. The large scale at which such an institution operates imposes a specific demand: that the information they use can be transferred easily across very different contexts. The general insight from the empirical work on bureaucracies and metrics is this: the larger the scale, the less the sensitivity to the details of a particular context. As F. A. Hayek puts it, central decision makers cannot serve each particular person, but only the average person.[46] And, I might add, central decision makers cannot serve local communities but only the average community.

7. VALUE SWAMPING

You might have started to suspect that there are actually two distinct problems running side by side here: a problem involving externality and a problem involving scale. To disentangle them, let us consider a different phenomenon, right next door to value capture, which will isolate the problems of scale. Consider a case where we actively participate in specifying some shared value—but the efforts of coordination at scale color the formulation of that value. Let us call it value swamping.

Value swamping happens when:

1. An agent's values are rich and subtle (or in the process of developing in that direction).
2. The agent participates in a large-scale social process that yields a specification of shared values.
3. Those specifications of shared values come to dominate the agent's practical reasoning (in the relevant domain).

Landemore underestimates the importance of federalism and local governance because of her unwarranted optimism in the possibility for lossless information aggregation at scale. What the empirical work I have discussed on quantification and bureaucracy—especially Scott's discussion—demonstrates is that information aggregation at scale always leads to massive data loss. This, Scott suggests, is a reason to strongly prefer local governance in most situation.

46 Hayek, "The Use of Knowledge in Society." I take Hayek to be a major influence on Scott's analysis. In fact, I take Scott's analysis to be offering a synthesis of Marxist criticisms of capitalism with Hayekian criticisms of central planning.

Value swamping is just like value capture, except for step 2. In value swamping, the agent does not get their values from some wholly external source; instead, they participate in a large-scale social process that yields a specification of a value.

Here is an example, slightly fictionalized from my own life. I took part in a large effort, across the humanities departments of my university, to defend the humanities from constant budgetary incursions from the STEM departments and the business school. We wanted to help the humanities survive and thrive. We ended up in a long discussion about realistic goals, and we decided that we most wanted to push for increasing the number of lines of humanities faculty and increasing faculty diversity. We ended up settling on some targets: we were going to push for a fast increase in the total number of humanities lines by 5 percent and embark on a long-term project to increase the representation of people of color in the faculty by 20 percent.

We needed such clear targets—and such a small number of them—because we needed some specific demands to bring to the upper administration. We also needed highly legible targets—the kind of targets that could be coherently targeted and tracked over the coming years by a revolving set of faculty representatives. Notice what is not on the list, however. I would have loved to push for creative work in new hires and for diversity in intellectual interests—and not just race. But how could we track such fuzzy, inchoate targets over the years? It might be that every humanities faculty cares about "creativity," but since we lacked a readily accessible and scalable measure of creativity, we cannot easily make it a group target.

What happened? Bowker and Star say that any attempt at large-scale collective action creates a demand for cross-contextual informational categories and for data that is readily aggregable.[47] This, in turn, creates a demand for publicly accessible, standardized procedures of measurement, such as metrics. Notice that the pressure for standardization here does not arise from the external sourcing of the metrics but from the demands of large-scale collective action itself. In other words, for some of the pressures, *it does not matter if the metrics are generated by an external source*.

Suppose that our case of value swamping is ideally participatory. Still, as the size of the relevant community scales up, the values that get generated are more and more subject to the demands of cross-contextual communication and consensus. Value swamping admits of more tailoring than value capture. Since a group's values are generated by the group itself, they can still be moderately tailored to the group's experiences. But there are still formidable constraints

47 Bowker and Star, *Sorting Things Out*, 53–161.

on the kinds of values the group can use. The group can only adopt the sorts of values that can be understood in all of the varying contexts across which the group operates. The processes of context stripping and de-nuancing are problems of scale, not of externality. In value swamping, the state-level interests, in information that is aggregable and portable across contexts, are not the alien impositions of an external force. They are necessitated by the process that we signed onto, and for very good reasons. But they do pull us toward less nuance in the specifications of our values as part of the drive for larger-scale cooperative action.

What we have learned is that there are two problems that can lead to badly tailored values. The first problem is that the values are generated by an external source. The second problem is the values are subject to the pressures of scale. In value capture by institutional metrics, we are exposed to both problems: externality and scale. In value swamping, the values may be our own, but they are still subject to the demands of scale.

Something is still lost in value swamping cases. But we cannot quite say that the values are not our own. We consented, we participated, we actively formulated, and we approve of the outcome. What is going on here is not exactly outsourcing. But we are sending our values out for processing at a larger scale and getting them back *filtered*. What comes back is what can survive the large-scale deliberative process intact; the more private, intimate, or small-scale communal reasons get filtered out.

8. The Scale Problem of Values

This suggests a larger picture. What we are starting to expose here is an essential problem with group agency at scale—or at least, a deep tension between smaller-scale agents and the demands of larger-scale agency.

We have lots of reasons to participate in large-scale collective efforts, and some efforts are far more effective when scaled up. Some things are best pursued collectively: reducing carbon emissions, increasing vaccination rates. In many cases, we can pursue those targets most efficiently by agreeing on a precise and shared specification of that target. In those cases, the upsides of having a precise, stable, shared specification of value may outweigh the cost.

But when we scale up our target-setting process, we lose sensitivity, contextual nuance, and granularity. And I take figures like Porter, Scott, and Bowker and Starr to have shown that this is no accident; it is an inevitable cost of scaling up organization for beings like us, using the methodologies we are presently using for informational aggregation. When we need to achieve agreement across a vast scale—across people who do not share all the same context, who

do not share the sensitivities—then what we can agree on will need to be subject to that filter of low-context comprehensibility. And insofar as collective efforts require some kind of shared stability, then these collective values will, of necessity, not admit of tailoring to small-scale or individual contexts.

The tensions between the value swamping and contextual tailoring are not the result of some sloppy process of coordination. They are baked into our very nature as limited beings with varying personalities and contexts who need to coordinate our actions across those contexts. They arise, in particular, from one feature of our finitude: that we each have a special understanding of our own patch and our particular context and a far weaker grasp of distant contexts. So, any attempt to render anything comprehensible across scale involves eliminating those details that require special understanding or contextual sensitivity. The tension between small-scale and large-scale valuing is ineliminable; our lives as both individual and social beings will always involve some kind of tension between our small-scale and large-scale commitments. We are the kind of beings that are perpetually stuck in a painful compromise—between the intimacy of small-scale understanding and the de-contextualized comprehensibility demanded of large-scale shared understanding.

None of this shows that we should not scale up our activities sometimes. Some things are best pursued collectively, as shared projects on the largest scales: reducing carbon emissions, increasing vaccination rates. And the demands of large-scale organizations clearly require clear, legible targets. But there is a cost to scaling up. Some goals—like stopping climate change—are worth the cost. And in other cases, we care more about the goods of local tailoring than the goods of large-scale cooperation. I can see plenty of good that can come from collectively pursuing a clear target that we all understand in the case of climate change or public vaccination. It is much harder to see the goods that come from collectively pursuing the same specification of, say, values in fitness, or musical values, or values in family organization. Some things are best managed at a personal or local scale. There is a trade-off between collective coordination and local specificity—and we may want to make that trade-off quite differently in different domains.

Here is an analogy: in law, we want federalism. That is, we want some of our laws set at the national level, some at the state level, and some at the county or city level. And the explanation is that for some kinds of laws, it is better to coordinate across a vast realm because the goods of standardization and sameness are worth the cost of low local tailoring. And for other kinds of laws, it is best to let them be set at smaller and smaller scales.

What this suggests is that we should want *value federalism*. Some values are perhaps best pursued at the largest-scale level, some at smaller community

levels, and some individually. And the upshot here is not that we should reject all large-scale values. It is that we should maintain a variety of differently-scaled values. There are many cases in which it might be useful to participate in a larger collective effort and so to accept, as part of that collective effort, less finely tailored goals. But, at the same time, we can confine those large-scale, standardized goals to our life inside those collectives and not let them swamp the rest of our values. The problem occurs when we exhibit an excess preference for the largest-scale values and let the largest-scale values swamp too many of our smaller-scale values. The problem comes when we let the demand for large-scale legibility intrude into every aspect of our lives, even the most intimate ones.[48]

University of Utah
c.thi.nguyen@utah.edu

REFERENCES

Andreou, Chrisoula. "Coping with Procrastination." In *The Thief of Time: Philosophical Essays on Procrastination*, edited by Chrisoula Andreou and Mark D. White, 106–15. New York: Oxford University Press, 2010.

Bowker, Geoffrey C., and Susan Leigh Star. *Sorting Things Out: Classification and Its Consequences.* Rev. ed. Cambridge, MA: MIT Press, 2000.

Bratman, Michael E. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.

———. *Shared Agency: A Planning Theory of Acting Together*. New York: Oxford University Press, 2014.

Brewer, Talbot. *The Retrieval of Ethics*. Oxford: Oxford University Press, 2009.

Buss, Sarah. "Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints." *Ethics* 115, no. 2 (January 2005): 195–235.

48 This paper has had a long, tortured, and meandering history. I owe an enormous debt of gratitude to an enormous number of people, and I am sure that more people have been deeply helpful than I have managed to recall and record. But as a start, I would like to thank Natalie Ashton, John Basl, Peggy Battin, Sarah Buss, Elizabeth Camp, Elizabeth Cantamalessa, Mercy Corredor, Steven Diggin, Steve Downes, John Dyck, Melinda Fagan, Will Fleischer, Carolina Flores, Rachel Elizabeth Fraser, August Gorman, Joyce Havstad, Max Hayward, Aaron James, Paul Katsafanas, Tim Kenyon, Eliot Michaelson, Elijah Millgram, Paul De Font Reaulx, Ronnie Sandler, Carlos Santana, Alexander Sarch, Karl Schafer, Jonah Schupbach, Nick Stang, Dustin Stokes, Olufemi Taiwo, Greta Turnbull, Elise Woodard, and the reviewers and editors of *JESP*.

Callard, Agnes. *Aspiration: The Agency of Becoming*. New York: Oxford University Press, 2018.

Christman, John. "Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves." *Philosophical Studies* 117, nos. 1–2 (January 2004): 143–64.

Dotson, Kristie. "Conceptualizing Epistemic Oppression." *Social Epistemology* 28, no. 2 (2014): 115–38.

du Gay, Paul. *In Praise of Bureaucracy: Weber, Organization, Ethics*. London: SAGE, 2000.

Espeland, Wendy Nelson, and Michael Sauder. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability.* New York: Russell Sage Foundation, 2016.

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Clarendon Press, 2007.

Gilbert, Margaret. *Joint Commitment: How We Make the Social World*. New York: Oxford University Press, 2013.

Hayek, F. A. "The Use of Knowledge in Society." *American Economic Review* 35, no. 4 (September 1945): 519–30.

Hedahl, Marcus, and Bryce Huebner. "Sharing Values." *Southern Journal of Philosophy* 56, no. 2 (June 2018): 240–72.

Holton, Richard. *Willing, Wanting, Waiting*. Oxford: Oxford University Press, 2009.

Jacobs, Jane. *The Death and Life of Great American Cities*. New York: Random House, 1961.

Jaeggi, Rahel. *Alienation*. Edited by Frederick Neuhouser. Translated by Frederick Newhouser and Alan Smith. New York: Columbia University Press, 2014.

Khader, Serene J. "The Feminist Case against Relational Autonomy." *Journal of Moral Philosophy* 17, no. 5 (October 2020): 499–526.

Krueger, Joel. "Music as Affective Scaffolding." In *Music and Consciousness 2: Worlds, Practices, Modalities*, edited by Ruth Herbert, David Clarke, and Eric Clarke, 48–63. Oxford: Oxford University Press, 2019.

Landemore, Hélène. *Democratic Reason: Politics, Collective Intelligence, and the Rule of Many*. Princeton: Princeton University Press, 2012.

Leonelli, Sabina. *Data-Centric Biology: A Philosophical Study*. Chicago: University of Chicago Press, 2016.

———. "What Counts as Scientific Data? A Relational Framework." *Philosophy of Science* 82, no. 5 (December 2015): 810–21.

List, Christian, and Philip Pettit. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press, 2011.

McGonigal, Jane. *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. London: Penguin Books, 2011.

Merry, Sally Engle. *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*. Chicago: University of Chicago Press, 2016.

Millgram, Elijah. "On Being Bored out of Your Mind." *Proceedings of the Aristotelian Society* 104, no. 1 (June 2004): 165–86.

———. *Practical Induction*. Cambridge, MA: Harvard University Press, 1997.

Nguyen, C. Thi. "The Arts of Action." *Philosopher's Imprint* 20, no. 14 (May 2020): 1–27.

———. *Games: Agency as Art*. Oxford: Oxford University Press, 2020.

———. "Games and the Art of Agency." *Philosophical Review* 128, no. 4 (October 2019): 423–62.

———. "How Twitter Gamifies Communication." In *Applied Epistemology*, edited by Jennifer Lackey, 410–36. Oxford: Oxford University Press, 2021.

———."Monuments as Commitments: How Art Speaks to Groups and How Groups Think in Art." *Pacific Philosophical Quarterly* 100, no. 4 (December 2019): 971–94.

———. "The Seductions of Clarity." *Royal Institute of Philosophy Supplements* 89 (May 2021): 227–55.

———. "Transparency Is Surveillance." *Philosophy and Phenomenological Research* 105, no. 2 (September 2022): 331–61.

———. "Trust and Sincerity in Art." *Ergo* 8, no. 2 (2021): 21–53.

———. "Value Collapse." *Royal Institute of Philosophy Supplements* (forthcoming).

Nguyen, C. Thi, and Matthew Strohl. "Cultural Appropriation and the Intimacy of Groups." *Philosophical Studies* 176, no. 4 (January 2019): 981–1002.

Oppenheimer, Daniel M. "The Secret Life of Fluency." *Trends in Cognitive Sciences* 12, no. 6 (June 2008), 237–41.

Owens, John, and Alan Cribb. "'My Fitbit Thinks I Can Do Better!' Do Health Promoting Wearable Technologies Support Personal Autonomy?" *Philosophy and Technology* 32, no. 1 (March 2019): 23–38.

Porter, Theodore M. *Trust in Numbers: The Pursuit of Objectivity and Public Life*. Princeton: Princeton University Press, 1995.

Reber, Rolf, and Christian Unkelbach. "The Epistemic Status of Processing Fluency as Source for Judgments of Truth." *Review of Philosophy and Psychology* 1, no. 4 (December 2010): 563–81.

Rovane, Carol. *The Bounds of Agency: An Essay in Revisionary Metaphysics*. Princeton: Princeton University Press, 1997.

Schneider, Jack, and Ethan Hutt. "Making the Grade: A History of the A–F

Marking Scheme." *Journal of Curriculum Studies* 46, no. 2 (2013): 201–24.
Scott, James C. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press, 1998.
Sterelney, Kim. "Minds: Extended or Scaffolded?" *Phenomenology and the Cognitive Sciences* 9, no. 4 (December 2010): 465–81.
Strevens, Michael. *The Knowledge Machine: How Irrationality Created Modern Science*. New York: Liveright, 2020.
———. "No Understanding without Explanation." *Studies in History and Philosophy of Science* 44, no. 3 (September 2013): 510–15.
Superson, Anita. "The Deferential Wife Revisited: Agency and Moral Responsibility." *Hypatia* 25, no. 2 (Spring 2010): 253–75.
———. "Deformed Desires and Informed Desire Tests." *Hypatia* 20, no. 4 (Fall 2005): 109–26.
Toole, Briana. "Demarginalizing Standpoint Epistemology." *Episteme* 19, no. 1 (March 2022): 47–65.
Westlund, Andrea C. "Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?" *Philosophical Review* 112, no. 4 (October 2003): 483–523.