VIRTUE IS NOT EASY

Professor Robert E. McGrath

Virtue is Not Easy
Robert E. McGrath

I attended my first Jubilee meeting here at Oxford nine years ago. At that point, I had been working with the VIA Institute on Character for about a year. I was still pretty new to this idea of studying character and virtue empirically. When I checked in for the conference, I was given a key that let me into both my dorm building and my room. When I got to the building, I found there was already a key in the lock. Clearly someone had arrived before me who left the key behind after they opened the front door. As I entered, I realized the doors to the rooms had been left open, so I was pretty sure this person didn't know their key was missing yet. I went to track down its owner.

I was there early enough that it turned out only one other person was in the building, so I brought him his key. He thanked me, and then said, "now that's an example of virtue in practice." I happened to attend his session the next day, which he began by telling this story and repeating what he had said to me. I know he was trying to express his gratitude, but I would have preferred a simple thank you. Coming so early in my thinking about what virtue is, his words have bothered me ever since. To be honest, I was just being civil. It was something that cost me almost nothing to do. He made virtue sound like something that was pretty feeble. If virtue is that easy, I really had to wonder why we were spending so much time talking about it, and why we expend so much effort and so much money trying to teach kids to be more virtuous.

From there I began on a long journey of trying to think about virtue and how virtue has evolved over history. Work by Steven Pinker, Norbert Elias, and others convinced me that virtue had grown over time, but how and why? Answering that question has led me to look at a variety of literatures not generally associated with character and virtue before. In this presentation I want to share some of that journey with you.

I want to start by making explicit some of the assumptions underlying the ideas I'll be presenting to you, though I suspect none of the assumptions I'm about to enumerate are likely to be particularly controversial among the people in this room. First, I assume it is reasonable and useful to use the concept of virtue in connection with thinking about the well-lived life, the life that optimizes self in the context of a matrix of communal obligations and expectations. I also assume it's reasonable and useful to conceptualize virtue as a composite of more discrete elements referred to as the virtues. As many of you know, my own work has led me to believe that virtue can be largely decomposed into three key virtue domains or virtue sets that I often refer to as the moral, self-control, and intellectual. Since I will return to these as a basis for my discussion of virtues later, I'll spell out here why I have settled on these three.

First, my personal interest in these three as key domains of the virtues emerged out of a number of latent structural analyses of 24 character strengths from the VIA system I've conducted. You may all be familiar with the VIA character strengths, but these were an attempt by Chris Peterson and Martin Seligman to generate a comprehensive set of the elements of positive personality. Regardless of the measure used of the VIA strengths, the population, or the analytic method used, these three consistently appear in the solution. No other latent factors emerge with such regularity.

Second, here is a listing of discussions of virtue in largely philosophical works that have proposed similar classifications for the virtues. Third, if it is an inherently meaningful taxonomy of the life well-lived, it should emerge in discussions of character or virtue education or of discussions about teaching so-called tranversal skills, skills that are globally useful across life situations and challenges. This slide includes several examples of that. Fourth, there is a body of literature that suggests they are key elements of how we judge people when we first meet them, indicating these features are important to perceptions of worth. Finally, I think it's meaningful that echoes of these three appear even in children's literature. That said, what follows doesn't depend on this taxonomy, it's just helpful to have a taxonomy I can use to make what I'm going to say from here on more concrete.

But back to my assumptions. My final assumption is that it is reasonable and useful to think about this concept of virtue as having a dimensional character. This third assumption in particular seems to me to be the starting point for all efforts to expand virtuous behavior. After all, if there is no dimension of virtuousness, how can people be expected to achieve greater virtuousness than they or their predecessors have achieved in the past?

But this is where things get more complicated. What does it mean to say that virtue is structurally dimensional? As Nancy Snow has tutored me, Aristotle was quite clear that more behavior consistent with a virtue is not necessarily more virtuous, and can even become invirtuous. A classic example here is the distinction between bravery and recklessness, or between curiosity and intrusiveness. In terms of the three virtues that interest me, there is the difference between the deeply moral and the Christian moral ideal, the person who gives up everything to follow Jesus. Some of you may be familiar with the story of George R. Price, for example, the brilliant chemist and theoretical evolutionist who converted to Christianity in 1970 and decided to pursue a truly Christian life. He essentially gave up everything to take care of the homeless, to the point that he himself was left homeless and alone, and ultimately committed suicide. Price's story always puts me in mind of Susan Wolf's statement about moral saints: "I don't know whether there are any moral saints. But if there are, I am glad that neither I nor those about whom I care most are among them." Looking beyond morality, it's also possible to be productively self-controlled to the point of burnout, or intellectually focused to the point of not being able to function effectively or connect to other people in a way that they can benefit from your thinking. I've known a few statistics teachers that fall into this last category.

As someone who has devoted a great deal of time to developing measures of character and virtue, I want to be clear that this description of invirtuous extremes is not intended to imply that scores on self-report scales of virtues or character beyond a certain score indicate invirtuous behavior. This is because people evaluate themselves within the range that is socially normative. That is, when asked how moral they are, most people assume the highest score is meant to indicate the maximum of moral commitment as normally practiced and socially accepted, which is well below the level at which that commitment lapses over into self-destructive selflessness. It is likely that if moral saints exist, they probably would give themselves the highest score on a self-report scale of moral identity. But so would the far larger population of individuals who are simply very, very moral but still live a conventional life. I mention this here both to clarify this issue, and because social norms will be an important part of what follows.

Getting back to the issue of what is the key underlying dimension, this discussion suggests it is not best understood in terms of frequency of behavior consistent with the virtue, as I suspect is commonly assumed. This begs the question, what should we be encouraging more of in our efforts to encourage virtue?

I think Vincent Ng and Louis Tay in 2020 provided a starting point for an answer when they suggested virtue can be understood in terms of situational optimization, the degree to which we think and feel in ways that motivate us to want to maximize the outcomes of our actions in terms of the good of the self-community matrix, and the degree to which the behaviors we choose reflect those thoughts and feelings. As my previous point about the virtues should make clear, I don't mean by this to limit the discussion to moral virtues. And even within the moral context, this is not intended to suggest a purely consequentialist framework. I think doing things because they are the right thing to do, or because of emotional sentiments such as compassion, can be equally valid considerations when thinking about what is best for myself and for those around me.

This leads me to propose that what we should be doing in our efforts to encourage virtue is to get people to think more deeply in general, and in particular about the implications of their actions for others as well as for themselves, and to choose a virtuous option when it is available. I like to think of this as *steadfast* virtue, or what Alasdair MacIntyre described as constancy. When that kindly individual told me my returning his key was virtue in practice, without first finding out whether my favorite hobby is in fact axe murdering, I'm afraid he was rushing to judgment. What really matters is whether I have an on-going commitment to being caring and fair, to being self-controlled with others and in my pursuit of my goals, and to obtaining and effectively processing information.

And here is where virtue, or the effort to encourage virtue, becomes hard. It involves cross-situational consistency. How do we help people become more steadfast in their virtue? There is another assumption that emerges here. School-based programs are based on the assumption that the solution and responsibility lies in the individual, and in particular, making that person more aware of virtue. I think these efforts are worthwhile, I have been involved in them myself, and I think they are part of the answer. But I'm afraid this assumption overestimates the contribution of personal effort to increasing virtuousness.

For example, many of you may be aware of the work by Schwitzgebel and Rust on the ethical behavior of ethicists. Among other topics, they studied differences in behaviors that could be considered moral at American Philosophical Association meetings during presentations on ethics versus other philosophical topics. They found no difference in terms of frequency of talking in the audience, letting doors slam, or leaving trash behind at the end of a session. Those who attended ethics sessions were no more likely to pay the registration fee for an APA meeting when they could get away with not paying it. Outside that context, ethics professors were no more likely to respond to undergraduates' emails, be an organ donor, donate blood, or donate more than 3% of salary to charity than other professors.

In case you think I'm being too hard on ethical philosophers, I'm spelling their findings out because it's a line of research I have available to me. I fully suspect the same level of compromise characterizes psychologists interested in wisdom and virtue and character educators. Are the people in this room, including myself, particularly more virtuous and less susceptible to temptation than the typical person in the street? I'm not convinced. How many of us here today have skimmed on our taxes in one way or another, or engaged in an emotional or perhaps even physical affair? Don't worry, I'm not asking for a show of hands. But I have to wonder if the rate is any different than the 3.4% of the general American population who admit to cheating on their taxes or the 13.4% who admit cheating on their spouse.

On a more serious note, American slavery co-existed for hundreds of years with a deep commitment to a Christian life. In fact, the Bible was used as one of the tortured justifications for slavery. There is a story in the Bible that Noah's son Ham caught Noah drunk and naked. In revenge, Noah cursed Ham's descendants with servitude. What matters here is that Ham is traditionally thought to have been darker skinned. This flimsy foundation was used extensively by good Christians in the South to justify enslaving millions of Africans. Religions have as often been the products of the morality of their time as its shaper.

And yet there is overwhelming evidence that the commitment to behave virtuously has grown over time. We no longer accept slavery, animal abuse, child abuse, or domestic violence as givens. Though wars have become horribly destructive, there are far fewer of them. Fewer people die of violence that at any time in recorded history. We are far better at pursuing complex, collaborative projects of massive scale such as the Webb telescope, the mapping of the human genome, or the construction of supercolliders. Despite major bumps in the road we are a more thoughtful and self-reflective people than our predecessors.

So, if personal awareness and personal intervention is not the primary vehicle by which virtue steadfastness has grown, what is? I've come to believe it is what the psychologist B. F. Skinner referred to as contingencies of reinforcement in the social environment, and economists talk about in terms of incentives. Specifically, I would like to make the case that, over time, shifts in environmental incentives encourage people to start to think and behave in ways we consider to be more virtuous direction. These changes are ultimately incorporated into social norms, and then they become the status quo. We bound ourselves by those norms.

Imagine an individual has been exposed to the best character and virtue education possible. They are taught the value of moral thinking, self-control, and thinking deeply. But then they graduate into a world in which gangs run the streets, or assertive action is treated as a better sign of a true leader than thoughtfulness, or only those born with wealth are likely to succeed. How useful is all that education in these social contexts? Is it reasonable or optimal to behave in deeply virtuous ways if social rewards argue against it. How long would a commitment to virtue survive as a social norm?

To move this analysis to the next level, I want to offer a mathematical approach to how changes in environmental circumstances can lead to changes in acting in ways we might call virtuous. Let's start with morality. Of the three virtue domains I consider key, it has been studied by far the most. What follows is inspired by work from the evolutionary psychologist Gilbert Roberts.

Some of you may be aware of W. D. Hamilton's very influential work attempting to understand the evolution of self-sacrifice in social systems, particularly among Hymenoptera species such as ants and bees. Using Hamilton's work as a starting point, Roberts developed a formula intended to describe circumstances under which genes that support self-sacrifice for unrelated conspecifics, other members of your species, could survive in a gene pool. I think his formula can be expanded and used just as well to describe circumstances under which we as individuals are likely to decide to act morally towards others, that is, what is rational in terms of the survival of genes can also be rational for the survival of a behavior in a social context. Here is a formula I think can be used for this purpose, again based on one offered by Roberts:

$$sb^o + b > c.$$

The variable c represents my estimate of the cost to me of acting morally. The variable b is my estimate of any benefit to me for acting morally. This can include many components: maybe the act is likely to improve my reputation or self-image as a nice guy, maybe it will reduce my distress over the suffering of another person, maybe I have faith that it will lead to some reciprocal kindness in the future. b^o is my estimate of how much the act will benefit others, the targets of my moral act. Finally, s is what Roberts referred to as stake, my personal investment in those others.

Putting these pieces together, the left side combines the personal benefit of the act to me with the benefit of the act to others, with this second factor adjusted for how much I care about those others. The right side is the cost. The formula suggests that the probability of engaging in a moral action increases as the left side exceeds the right side, the cost to me.

I can also introduce a complementary rule applying to circumstances of harm to others:

$$b > c + sc^o.$$

That is, the probability of immoral decisions increases as the benefit to me exceeds the anticipated cost to me, plus the cost to others adjusted for my stake in those others. For example, deliberations about whether to engage in infidelity are often based on its assumed high value, and low potential cost under the assumption that the infidelity won't be discovered and will stay, as it were, under the covers.

I also think the decision to pursue long-term goals, which is a key aspect of self-control virtues, can be rationally formalized by

$$b - d > c$$

In this case, rational decisions of whether or not to pursue a long-term goal involve evaluating the goal's benefit to the actor versus its cost, after discounting that benefit based on several factors. Specifically, the discount is a function of at least two factors:

$$d = f(t, u)$$
$$u = f(o, i)$$

That is, the discount increases as the length of time involved in the activity increases, and as the uncertainty of ultimate success increases. This uncertainty is in turn a function of obstacles to success and/or social instability. Unfortunately, I don't think the decision to use clear reasoning is as easily expressed as a formula, since it is a more internal activity.

I should make a couple of things clear here. I'm not arguing for a classical economic view that assumes individual human decision-making is exclusively or even largely based on rational deliberation. The evidence is far too strong that we fall prey to cognitive shortcuts that allow for suboptimal decision-making. The economist Gary Becker is credited with popularizing the idea of trying to understand the personal decision to engage in crime in terms of a rational analysis of costs and benefits. His work led to a substantial line of research on the decision-making in criminals. The conclusion of that research is that criminals are not particularly driven by rational model consideration. So I don't intend these formulas as models of individual decision-making.

In fact, even if we wanted to be purely rational about good and evil, we're simply incapable of estimating the values in these formulas accurately or doing this kind of math in our heads. So, given that most people hate formulas, why am I bothering you with them? It's because I believe they have some validity as a representation of what rational decision-making about whether to make the or a moral choice would entail. Though we're not perfect in making these decisions, I do think that communities are based around social norms, and those social norms are significantly influenced by what is rational given the current environment, though they are influenced by other factors as well such as symbols of status and common desires. Finally, I'll also highlight these formulas indicate that those rational considerations can and very often do include emotional factors such as caring for others.

From here on I'm going to focus mainly on the rational analysis of whether to act more morally. If social norms surrounding morality are in part shaped by what is considered reasonable, how would moral virtue become more steadfast? What moves communities to expect a higher level of commitment to moral action? What the formula suggests I would argue is that social expectations for moral virtue would tend to increase if:

(1) The cost of moral action declines. This can happen when cultures become more affluent, for example.
(2) Our moral stake in others increases. This can happen with increased information about others, increased interaction with others, or increased dependency on others.
(3) The potential benefit of our moral acts to others increases. This can happen when resources exist for people to use our kind acts more effectively.
(4) The potential benefit of moral actions to ourselves increases. This can happen when interdependencies increase in a society.

To demonstrate how this can happen, I'll return to the historical example of slavery. The modern slave trade in Europe began in the 15th century, starting in Portugal and later spreading to include Spain, France, the Netherlands, Denmark, Sweden, and Britain. There were enemies of the slave trade almost from its inception, but for most of 300 years it was socially tolerated, and even considered sound business practice. John Locke, the father of liberalism, was one of many investors in the slave trade. But then in a period of less than 100 years it went from good

business to anathema. Today, we couldn't imagine any society actively advocating for slavery. What changed?

The first modern novels appeared in the 17th century, and the historian Lynn Hunt believes they set off a dramatic cognitive and emotional shift in the lives of literature Europeans. She points to many signs that emerged shortly thereafter of a society that was more self-reflective, thoughtful, and attuned to the suffering of others. That is, the general population's stake in previously marginalized populations started to grow. The starting point for the popular rejection of slavery in both the United Kingdom and the United States seems to have been the publication of hugely successful narratives about the evils of slavery. In the U.K. the most important contributor was the 1798 publication of the autobiography of former slave Olaudah Equiano. In the U.S. it was Harriet Beecher Stowe's *Uncle Tom's Cabin* in 1852 that set things going. These books created a greater sense of emotional connection to the plight of slaves; a sense of emotional stake became widespread.

The second factor was the cost of ending slavery. For the British, slaves were used primarily for the Caribbean sugar industry. By the time Britain ended slavery, in 1833, competition had driven that industry into steep decline, so shutting down slavery could occur without substantial cost. Similarly, in the United States the seat of abolitionism was the industrializing North, where slaves had always had little economic impact. The fact that the economy of the South relied heavily on slavery to fill agricultural labor needs meant that armed conflict became necessary to resolve the matter.

I find the case of Britain's Lord Palmerston in the early 19th century particularly interesting for what it says about how environmental circumstances can selectively shape ethics. His letters and actions suggest he was a zealous opponent of slavery, and an important contributor to the effort to abolish slavery in the United Kingdom. Yet barely 6 years later, he was the driving force behind the first Opium War between Britain and China. When Britain was suffering a trade deficit with China based largely on the love of Chinese tea, the Chinese were forced to open their borders to opium dealers from British India. The result was that something like 3% of the entire Chinese population became addicted to opium. These were two ethical sides of the same man. I have to wonder how things might have been different had the British public widely read an autobiography of a Chinese opium addict, or if the Chinese had not been so successful at feeding the British desire for tea.

I will stop with slavery, but I have many similar vignettes having to do with the decline of violence as a social option, the expansion of protections to women and animals, increasing objectivity in social judgments, the growth of long-term goal pursuit, and other activities that we would associate with the growth of virtue. In each case, changes in social structures ultimately changed social norms.

The final point I want to make about this process of changing environmental circumstances changing social norms is that those social norms in turn reshape personal cognitive schemas for what are acceptable behavioral options and beliefs. Returning to the example of self-rating my level of morality, I choose the highest option because I don't even consider the moral saint to be

a consideration in my self-evaluation. They fall so far outside the norm that I treat them as irrelevant to my judgment of scale.

Here's one of my favorite examples of this reshaping of thinking by social factors. Based on a hypothesis first raised by the economist Ester Boserup, Alberto Alesina and his colleagues compared cultures that traditionally used plows to those that relied on alternatives such as digging sticks. Plows require a great deal of strength, so males would have come to dominate traditional agricultural production, while digging sticks can be wielded effectively by either gender. They found that plow cultures still have fewer women participating in government and business in the 21ˢᵗ century, long after plows have been replaced by pens and personal computers. An even more interesting finding was that traditional sexist attitudes were stronger in children of parents who emigrated to Europe or the United States from a plow culture, suggesting an effect on beliefs and values even after removal from that environment.

The shaping of cognitive schemata by environmental incentives can be for the positive as well as the negative, and very often has been. Here we see another part of the process by which virtues can become more steadfast. Slavery is not just a prohibited practice, it has become a reviled practice. We are less prone to act violently than at any known time in human history, and for many of us we no longer even think of it as a viable option except in cases of immediate self-defense. We make plans for family security in our retirement 40 years hence with no second thought about it. We assume that people should be judged and treated fairly. Greater virtue becomes part of the skin in which we live. To use an Aristotelian concept, we move from being continent, considering invirtuous options but ultimately deciding to act for the good, to being truly virtuous, where we no longer even consider the invirtuous option possible. I'm pleased to say this influence of norms on the range of options was recently supported in a study by Kalkstein et al. in press in the *Journal of Personality and Social Psychology*. They found that socially counter-normative behaviors can fall out of consideration as options.

So far everything I've said is an attempt to provide a descriptive model for the growth of virtuous tendencies in the past. But we are all interested in how we can get people to be more virtuous, and things always get messier when you try to look to the future. I suspect everyone here agrees with me that social stability, progress, and accord is enhanced by a more virtuous orientation to life's challenges. Is it possible for us to outline a prescriptive approach to making that happen, an intentional agenda for social revision that would produce the kinds of changes in normative expectations associated with greater virtue?

Here are some social changes I think could contribute to advancing such an agenda. The first is the development of a complete curriculum spanning the length of the required educational system that focuses on issues of virtue. Such a curriculum would:
    (1) increase awareness of key moral, self-control, and intellectual skills.
    (2) be designed to be developmentally appropriate for different ages
    (3) ideally include both dedicated sessions and use of virtuous skills in the general curriculum.
I know pieces of such a curriculum already exist. I also know this can seem like I'm contradicting myself, since earlier I suggested that awareness is generally not enough to ensure steadfast virtue generalizes beyond instruction. I do think it's not sufficient, but it would at least

prepare people for a world that elevates its expectations surrounding virtue. And who knows, if it were a universal curriculum, perhaps that could become part of shifting social expectations for adults. That said, I have to admit I suspect this kind of training tends to have its greatest impact on those students who were already amenable to a more steadfast approach to virtue in their lives.

But I think the more important issue is whether the social environment will continue to change in ways that encourage a more virtuous stance, and here I find myself getting particularly vague. I do think there are some general trends in how we operate that could encourage a more moral, self-controlled, and intellectually sophisticated approach to life. For example:
  (1) As transparency in any settings increases, there is greater social pressure to act immorally. This is particularly true in the work can both preclude opportunities for misbehavior, and on the positive side increase the potential to establish a positive reputation.
  (2) Innovations that enhance social welfare, and commitment to social welfare, over time have the potential to reduce certain forms of crime.
  (3) Involving a broad spectrum of stakeholders in major decision-making can improve intellectual virtues. Personally, I'm particularly drawn to the potential for movement from stockholder-centric of stakeholder-centric capitalism, where there is an expectation that all major constituencies influenced by a corporation's actions should be involved in governance.
  (4) Enhancements to basic safety and security contribute to both non-violence and goal pursuit.

Some of you may recognize that I'm coming awfully close to the idea of nudges, though I suspect things like a comprehensive and universal virtue curriculum or increasing transparency in business transactions are closer to a shove over a cliff than a nudge. I accept the similarity. Where I see a difference in emphasis is that discussions of nudges tend to focus almost exclusively on behavior change. As I've indicated, I am just as interested in the normative and cognitive changes that accompany substantial social innovations and concomitant behavioral changes, because it is these norm changes that are the foundation for a greater personal commitment to choices we see as virtuous.

I realize there will be resistance to all of these possibilities. Since this is a conference about virtue in the professions, I'll focus on the work setting. Unfortunately, I consider that setting involves some substantial, inherent obstacles for the steadfast application of virtue. The profit motive exceeds the desire to do good. Businesses adopt strong hierarchical structures that can undermine deep thinking and goal pursuit. They also work to encourage loyalty to the organization and to colleagues rather than the larger community. Similarly, any efforts to increase transparency tend to brush up against concerns about privacy and proprietary information.

That said, I'm going to go out on a limb here and suggest that many of the things that will contribute to social virtuousness are going to occur gradually, developing in parallel without necessarily being connected with each other as part of a coherent effort to advance virtue. The truth is that from traffic cameras, to accounting disclosures, to requiring a nurse in the room

during a physical exam of a woman by a male physician, to online revelations of wrongdoing, we are already gradually moving towards a more transparent society. There are people far more influential than I am already pushing for a shift to stakeholder capitalism. How far this process will go I can't say, though I do think modern technologies, particularly the Internet, will play a powerful role in the long run pushing us in virtuous directions, even if in the short term that seems not to be the case. It is already making it easier for us to pursue goals, to consider populations of people and animals we have never considered before in our moral speculations, and to expose ourselves to alternate viewpoints. In fact, I believe much of the political and social conflict we have experienced in recent years is a reaction to how quickly ways of thinking are changing.

I'll finish with a challenge. If my analysis here makes sense, is it possible to move this idea from prediction about the future to prescription. That is, would it be possible for the community of people, like those in this room, who see the value of social change in ways that will contribute to a more virtuous, stable, and flourishing future to identify mechanisms by which those changes could be pursued more quickly and effectively? It's a question we have barely addressed at this point, and I really don't know the answer. But I think it's an interesting question to discuss.

To summarize, achieving virtue is not easy. It requires a steadfast commitment to approach problems from the perspective of what people in this room would call practical wisdom, in terms of what is the best for self and for community in each unique situation. Encouraging more steadfast virtue is also not easy. Efforts to boost virtue through intentional, time-limited educational programs are helpful, but I doubt they are effective over a lifetime in the face of social situations that reward invirtuous behaviors. I think the real solution lies in social innovations making more virtuous choices easier and more desirable to pursue. These choices are then incorporated into more virtuous normative expectations, which in turn shape the set of behavioral options we consider acceptable and worthy. Looking forward, it is unclear to me whether these concepts can be the basis for a social agenda intended to enhance morality, self-control, and intellectual proficiency, or whether that change is likely to occur organically. But I think it is the question we need to think about if we are truly committed to the hard work of encouraging a more steadfast commitment to virtue.