

Compromise despite conviction: Curbing integrity's moral dangers

Hugh Breakey¹

Integrity looks dangerous. Passionate willpower, focused devotion and driving self-belief nestle all-too-closely—both conceptually and psychologically—to extremism, narcissism and intolerant hubris. How can integrity skirt such perilous terrain? This question's significance extends beyond mere definitional matters. It speaks to the perennial ethical question of whether devout, driven devotees can guard themselves from antisocial extremes.

Existing proposals to inoculate integrity from moral danger hone in on integrity's reflective side. I argue that this 'epistemic approach' disarms integrity's dangers only by stripping the trait of everything that made it worthwhile in the first place. Instead, I propose a direct solution: integrity contains substantive moral principles that serve to surgically target the dangers integrity would otherwise pose. The person of integrity avoids extremism not by questioning whether her values are right, but by recognizing that in a social world whether her values are right is not the only factor that bears on how she should act. Normatively, the proposed account allows integrity to retain its intuitive allures of willpower, direction and unified character. Descriptively, it explains the surprising capacity for principled compromise displayed by canonical figures of integrity, as exemplified through acts of conscientious objection and civil disobedience.

In short, integrity empowers us to be fit for society, even as we are true to ourselves.

My argument proceeds as follows: When theorists discuss integrity's dangers, they have foremost in mind a character trait I will call 'self-directedness'—roughly, being 'together' or 'driven'. I begin my argument by drawing on the literature to develop a coherent account of self-directedness, and I detail the (by my count) six distinct moral dangers such self-directedness poses. Yet integrity differs from self-directedness in several ways. The next sections describe these differences: showing how integrity possesses elements of sincerity and morality as well as (I will argue) a remarkable capacity for principled compromise.

With this groundwork laid down, I interrogate the epistemic approach to avoiding moral danger, contending that this approach does not answer the core moral problem so much as sidestep it. In its place, I construct an account of integrity that seizes self-directedness' attractions while dodging its dangers, arguing that integrity's ethical elements constitute a targeted response to the perils it would otherwise engender. Principled compromise with others' contrary positions turns out to be not only consistent with integrity, but essential to it. In the final section I sketch four normative arguments for integrity, so understood, showing why societies might construct the virtue and why individuals might embrace it.

¹ Institute for Ethics, Governance & Law, Law Futures Centre, Griffith Law School, Griffith University, Australia. Correspondence to: h.breakey@griffith.edu.au. This is the Accepted Manuscript version of: Hugh Breakey. 'Compromise despite Conviction: Curbing Integrity's Moral Dangers' *Journal of Value Inquiry* 50, no. 3 (2016): 613-29. The final publication is available at Springer via <http://link.springer.com/article/10.1007%2Fs10790-016-9541-1>

1. Self-directedness

We possess no perfect, pre-existing term for the trait of ‘self-directedness’, though many colloquial expressions capture its gist. Self-directedness implies being a ‘together’, ‘focused’, ‘spirited’ or ‘driven’ person—someone who ‘is his own man’, ‘has the courage of her convictions’ or ‘marches to the beat of his own drum’. For expositional purposes, I will stipulate ‘Selena’ as our archetypal self-directed person.

Almost all major philosophical accounts of integrity include an element of self-directedness, so understood, and it is easy to see why. Dictionaries define integrity as being ‘whole’, ‘unified’ and ‘undivided’. The term’s etymology invokes ideas of being ‘integrated’ and cleaving to what is ‘integral’.¹ Accounts of integrity incorporating self-directedness capture many intuitive uses of the term, and rule out many of integrity’s well-known contraries, such as being torn, weak-willed and inconstant.² Clearly, integrity has a lot to do with self-directedness.

How then should we understand self-directedness? I submit that the trait incorporates three distinct qualities of (what I will term) ‘driving-values’, ‘reflective-ownership’ and ‘defining character’. Knitting these three qualities into a single character delivers us a together, driven person—exactly the character feared by those who stress integrity’s dangers. (While I go on to present a fairly specific account of self-directedness, note that my larger thesis does not hinge upon this account’s particulars. The overall argument requires only that self-directedness involves some form of driving values and defining character.)

Driving-values requires that Selena’s principles, projects and desires knit together to form a stable, functional whole, providing her life with direction, meaning and energy. The integrity literature describes two qualities of values—‘integral’ and ‘integrated’—of use to us here.³ *Integral* values center on Selena developing and pursuing those values integral to her identity as a specific person.⁴ Selena must possess a strong idea of who she is, and what she endorses: she understands and cleaves to the pursuits, histories, standards, communities, desires and relationships that define her identity. For *integrated* values, what matters is that all Selena’s values press in the same general direction. Selena’s preferences—and her preferences about what preferences she wants to have—accord with one another.⁵ This integration requires that there are no deep-seated clashes between her values.

In sum, pursuing what is integral makes Selena’s life feel more invested, special and worthwhile. Integrating her preferences with those integral values allows Selena to pursue her values whole-heartedly, with less felt tension, sacrifice and regret. Combining the two qualities gives Selena driving-values—providing her life with direction, energy and meaning.

Self-directedness’ second element, *reflective-ownership*, captures the level of cognitive attention and action required for Selena to be genuinely *self-directed*—as opposed to manipulated or compartmentalized. Reflective-ownership relates to the first element (driving-values) by requiring more than that Selena just-so-happened to possess such values. To the contrary, Selena herself needs to have deliberately worked out what values are integral to her life, and how best to adapt her preferences to the integrated pursuit of those values. Thus, the *outcomes* that constitute Selena’s driving-values (integral and integrated values) become reflective *processes* that she must continuously perform.

Reflective-ownership also includes critical attention, demanding that Selena has thought for herself about her values and her reasons for holding those values. The idea here is

to set down a ‘basic critical rationality’ that guarantees that Selena is indeed self-directed (not manipulated by others, or blindly conformist), stable (not fragile or pliable in the face of criticism) and cognitively integrated (avoiding inadvertent compartmentalization).

To ensure these outcomes, Selena’s basic critical rationality requires that she makes sensible attempts to ensure her empirical beliefs’ accuracy, and that her values cohere with those beliefs where appropriate.⁶ And in the spirit of a Millian respect for her own values, we might think that Selena would inform herself about others’ contrary values and beliefs. ‘He who knows only his own side of the case,’ as Mill observed, ‘knows little of that.’⁷ This basic critical rationality allows for Selena to hold mistaken and unjustified convictions—just not recklessly uninformed ones. Combined with the deliberate construction of her driving-values, reflective-ownership thus ensures Selena avoids being wantonly obsessed, ‘in denial’, self-deceived, conformist or manipulated (all signature failures to take charge of one’s life).

The final component, *defining-character*, gives Selena the strength of mind and courage of action to guarantee that her actions follow from her values. Having decided what she wants to do on the basis of her values, Selena follows through on her decision. Selena’s defining-character differs from her driving-values by referring to her psychological character⁸—specifically, her disposition to feel certain emotions in certain situations, and how she responds when gripped by those emotions. Defining-character thus includes the emotional dispositions of willpower, patience, diligence, courage, grit, assertiveness, dedication and conscientiousness. These qualities immunize Selena against the common human frailties of cowardice, diffidence and akrasia, ensuring she can pursue her task without being tempted or distracted, or too fearful or lazy to follow through with it.

Combined together, Selena’s driving-values, reflective-ownership and defining-character consolidate into her over-arching trait of self-directedness. She is, we might say, a ‘together’ person. (Note the conceptual parallel between ‘togetherness’ and ‘integrity’.) It is not hard to see why Selena might have wanted to be self-directed, and why we might envy her for attaining that status. Selena’s self-directedness empowers her to take her calling seriously. By organizing her other preferences around her integral commitments, by having the tenacity to follow through on them, and by thinking carefully about them, Selena does justice to her values’ importance. Selena’s self-directedness also promises her personal benefits.⁹ Because her projects and values knit together, she is not torn or indecisive—and her successes one day feed into her projects the next. With her strong will and defining-character, Selena acts on her values and goes for what she wants. And because she reflectively endorses her values, Selena identifies wholeheartedly with her successes. (Many of self-directedness’ features dovetail with other attractive psychological qualities, such as emotional intelligence and ego-strength.¹⁰) Selena knows who she is, and she is nobody’s fool.

2. Moral Danger

Selena is dangerous. A veritable Pandora’s Box of threats lurk within her signature trait.¹¹ In fact, we can distill no less than six distinct ways self-directedness poses moral dangers by impelling Selena to commit serious ethical violations. I will term these the dangers of empowerment, egoism, personal calling, extremism, investment and unilateralism.

The danger of *empowerment* stresses that Selena’s self-directedness makes her more powerful. Tenacious, decisive and single-minded, shrugging off threats that would worry lesser

types, and admired for holding the courage of her convictions, Selena makes things happen. ‘I’m a can-do person,’ she enthuses, ‘and I get results.’ This means that if Selena decides to harm others, she might be very successful at it.

The danger of *egoism* arises when Selena’s self-directedness makes her think she is being principled as she pursues her naked self-interest. ‘It’s my life,’ she insists, ‘and I must make the best decision for me.’ Selena might even be blind to her egoism, with cognitive dissonance gulling her into forming values that vindicate her prior selfish actions.

In the danger of *personal calling*, Selena feels she is justified in pursuing the values she has chosen purely because *she* has chosen those values. ‘This is my path,’ Selena proclaims, ‘and right or wrong I will follow it.’ But even if Selena’s values avoid simple egoism, there is no guarantee they will be morally sound, and her quest for self-actualization might overthrow ethical obligations that would otherwise restrain her.

Alternatively, Selena’s self-directedness might steer her in an opposite direction, towards an unyielding faith in the objective truth of her convictions: the danger of *extremism*. True, Selena’s basic critical rationality prevents her dogmatically demanding she cannot possibly be wrong. But this does not dispel her overwhelming conviction she is right. Open-mindedness does not dispel surety—it only requires that Selena remains receptive to revisiting her surety when presented with new evidence. ‘It’s simple,’ explains Selena. ‘There’s my way and there’s the wrong way.’ Again, the danger looms that Selena’s confidence in morally mistaken convictions might see her spurn social restraints.

Even if Selena dodges these dangers, the fact remains that she is emotionally, materially, socially and cognitively committed to her chosen values. Call this the danger of *investment*. ‘This is what I have become,’ recognizes Selena, ‘and I can no sooner surrender my values than stop breathing.’ Through her driving-values, Selena has consolidated her commitments, invested in projects, taken up role-identities, and cultivated personal relationships. Through her defining-character, Selena has shaped her emotions and intuitions; her snap-judgments and gut-feelings chime with her values and support her endeavours. And through her reflective-ownership, Selena has ensured that all these personal investments cohere with her beliefs and rational deliberations. All this means that if Selena cleaves to morally fraught values, she will have little capacity to release herself from them. Selena has boarded a juggernaut—and there is no stopping it now.

All these dangers contribute to Selena seeing her values as supremely authoritative, triggering the final moral danger of *unilateralism*. We live in a social world, and public peace and cooperation depend on respecting rules that others have a hand in creating—even if we ourselves strenuously oppose those rules. But Selena’s self-directedness might make her resistant to compromise, to sharing authority, to being bound by other’s choices, and to acknowledging the legitimacy of social decision-making processes. ‘Why should I cede any authority to others’ opposing values, and their fallible democratic institutions,’ Selena demands, ‘given my strong commitment to my own convictions?’ When society prohibits what Selena wants to do, she does it unilaterally through force or secrecy.

Self-directedness thus stokes a hornet’s nest of moral dangers. I noted earlier how integrity shares many features with self-directedness. Yet we rarely apply the term ‘integrity’ to extremists who display what Greg Scherkoske aptly terms ‘vicious steadfastness’.¹² A puzzle

thus arises (the first of three puzzles I aim to answer): Given that integrity resembles self-directedness, how does integrity avoid self-directedness' moral dangers?

This puzzle extends beyond how we happen to use the word 'integrity'. History shows all-too-clearly that fiercely held convictions can and do erupt into brutal extremism, unyielding unilateralism and self-righteous egoism. In searching for a theory of integrity, even as we search for purpose and direction in our own lives, we want an account of how convictions can give people reasons to live without giving them reasons to kill.

Before continuing, observe one background assumption: these concerns with moral danger presuppose *either* that reigning community mores are somewhat on the right track (so spurning existing social norms invites moral danger) *or* that general peace, security and social cohesion carry some moral weight (so trampling these invites moral danger). To be sure, these are thin assumptions, as almost all societies will fulfil at least one of these minimal conditions. But in extreme cases—such as Nazi Germany—these assumptions may not hold. For brevity, I will limit the following analysis to more ordinary cases where societies are flawed but not irredeemably evil. But if we think that integrity in extreme cases remains possible (consider Oskar Schindler), then we will at some point need to expand our philosophical ambit to address these more brutal contexts.

3. Integrity's Sincerity

If we stipulate *Ingrid* as an archetypal person of integrity, then what other features does Ingrid possess, beyond those held by Selena?

Perhaps most obviously, Ingrid must be *sincere*. Hypocrisy and deception (including their institutional corollaries of corruption and the abuse of power) stand as singular failures of integrity. Ingrid must stand for her values, even at personal cost and in the face of social hostility. People may disagree with Ingrid, but they know what she stands for and they trust her to comply with her commitments.¹³ Integrity here lies in the integration of word and deed, the authentic accord between the public persona and the internal bearing.

4. Integrity's Morality—and its Puzzles

Another difference between Ingrid's integrity and Selena's self-directedness lies in the link between integrity and correct moral action. Without triggering too much controversy we might agree that Ingrid must have a personal moral stance. She possesses moral commitments, reasoning and sensibility, and she holds firm to her principles in the face of challenge. This does not tell us anything about the contents of Ingrid's moral life: only that she *has* (what she pictures as) a moral life.

Beyond this rather bland claim about the relation between integrity and morality, however, it proves difficult to nail down integrity's precise moral requirements. Some commentators define integrity as connoting exemplary performance of objective moral duties.¹⁴ Others resist any collapse of integrity into moral sanctity – even allowing (in principle) Nazis to possess integrity.¹⁵ Complicating affairs is that ordinary usage allows us to apply the term to people we disagree with morally; we can praise another's integrity even as we disavow their principles.¹⁶ Let us call this the *moral disagreement puzzle*; we want an account of integrity that explains how acknowledging Ingrid's integrity counts as high moral praise even though it does not imply she has avoided significant moral error.

5. Integrity and Principled Compromise

One of integrity's less-recognized puzzles is the *exemplary compromise puzzle*. This puzzle, which I will argue ultimately paves the way for a resolution of integrity's core nature, hones in on the relationship between integrity and compromise.

We take as exemplars of integrity those people who display professional integrity or who engage in conscientious objection or civil disobedience. Yet each of these practices involves compromise. Each activity involves a person who cannot agree with a particular practice, but who reconciles herself to the practice or to the society that countenances the practice. Let us survey in turn the elements of compromise resonant in these three practices of professional integrity, conscientious objection and civil disobedience.

The person displaying professional integrity says, 'On the basis of my core values I have reservations, but I will do my duty anyway.' The defense lawyer ably defending a client he believes is guilty, the judge ruling in accord with a law she thinks is unjust, and the public servant setting aside his own cherished investments as he selects public expenditures, all display great professional integrity.

The conscientious objector says, 'On the basis of my core values I disagree, and will not comply with this practice. But I will not myself unilaterally prevent this practice occurring, and I may even facilitate it.' The doctor morally opposed to abortion conscientiously objects to performing the operations herself—but she must not sway her patient to her preferred outcome. Indeed, in alerting the patient to this alternative, and even referring the patient, she plays a causal role in facilitating the abortion itself.¹⁷ Similarly, the conscientiously objecting soldier does not perform treason or sabotage, but surrenders himself to incarceration. Standing resolutely against the war, the soldier only withdraws his own contribution, rather than attacking the war effort itself.

The person undertaking civil disobedience says, 'On the basis of my core values I disagree, and I will prevent or disrupt this practice. But I will not act secretly. I will peacefully cooperate with the lawful authorities, and stand legally accountable for my actions.' Peaceful public protestors who face physical, social and legal risks make a similar claim. Rather than simply flouting a law, or covertly undermining its operation, the protestor and the civil disobedience provocateur raise awareness and accept the costs of doing so.

To be clear: The exemplary compromise puzzle is not: How are these three practices *consistent with* integrity? That puzzle—if it ever really was a puzzle—has been definitively laid to rest.¹⁸ In particular, Martin Benjamin has argued at length that even people with great integrity will often have values that—in some situations—must clash with each other.¹⁹ For instance, while Ingrid may have firm views on a contentious topic, she may also possess strong views on the importance of not securing settlements through violence or intimidation, of achieving settlements that respect the sincerity of others' views, of engaging in constructive and inclusive dialogue, and so on.²⁰ Confronting disagreement on the disputed issue, Ingrid will desire a way forward that respects *all* these values. To be true to herself, Ingrid will need to bring these conflicting values into alignment—to compromise between them. And as Benjamin observes: 'It is one thing simply to compromise (or betray) one's principles; quite another to compromise between them.'²¹

So let us agree that the types of compromises noted above do not necessarily compromise integrity. The vexing question, though, is why we would see these practices as *exemplifying* integrity. Why are our greatest exemplars of integrity those whose practices follow a statement of principled disagreement with a ‘but’? As in, ‘I disagree, *but* I will nevertheless support in various ways the practice or the social system that endorses it.’ For despite these compromises, we unhesitatingly recognize these actors as paradigms of integrity. Popular discourse and theoretical analysis both highlight the links between *professional integrity* and ordinary integrity.²² Etymologically, the ‘professional’ is literally ‘one who professes’—and influential theories of professional ethics place this act of integrity at their very core: the act of professing what one stands for and living by it.²³ *Conscientious objection* remains the fundamental policy arena in which a concern for practitioners’ personal integrity shapes law and policy.²⁴ And *civil disobedience* stands as the defining feature of so many of our greatest exemplars of integrity: think of Emmeline Pankhurst, Gandhi, Rosa Parks, Nelson Mandela and Martin Luther King Jr.

Indeed, we can easily forget that the noble deaths of venerated paradigms of integrity—such as Socrates and Sir Thomas More—were thick with compromise. Socrates chose to drink hemlock rather than flee the Athenian justice that condemned him to death. Yet he submitted to that punishment out of respect for Athens itself—despite his steadfast view that the sentence handed down by the Athenians amounted to a flagrant injustice. As for Sir Thomas More, despite his unswerving religious objections to the union, the Saint attempted compromise after compromise on Henry VIII’s marriage to Anne Boleyn. Thomas tried repeatedly to resign his post, and then held his silence on the matter, knowing that legally his silence had to be understood as consent. It was only when—at the last—Henry demanded the Saint explicitly swear the oath of supremacy of the crown that Thomas was left with nowhere to turn. Both Socrates and Sir Thomas were willing to acquiesce and forgo positive resistance to, or direct breach of, laws and norms with which they profoundly disagreed. Far from uncompromising zealots, each knowingly worked with and within the systems that put them to death.

Here, then, lies the puzzle: Why are those who act with the deepest professional integrity, those who conscientiously object and those who engage in civil disobedience upheld as paragons of integrity? Why do we not instead (or at least similarly) valorize the person of integrity as the one who follows through unblushingly on their chosen purposes? Why—to speak concretely—do we laud Emmeline Pankhurst’s integrity when she says she would die for women’s freedom, but would not kill for it?

6. Putting the Pieces Together

Let us take stock. We began by considering the quality of self-directedness, and its moral dangers. We noted that, despite its parallels with self-directedness, integrity somehow avoided its dangers. We observed too that integrity contains elements of sincerity, morality and principled compromise.

The attentive reader might already glean how these puzzle-pieces might snap together. Integrity’s curiously-shaped elements of sincerity, morality and compromise *turn out to be the very qualities that specifically target the dangers of self-directedness*. Integrity emerges not as ‘ethical’ in a general way, but in a quite particular way, surgically targeted at the dangers that would otherwise render it morally perilous.

Before we explore this moral approach to responding to the dangers, it is worth examining an alternative response, currently prevailing in the literature.

7. Responding to Danger: The Epistemic Approach

One way of responding to the challenge of integrity's moral dangers inflates the role of rational reflection in integrity's workings. We saw earlier (in Section 2) that self-directedness includes basic critical rationality, and that this included some recognizably epistemic virtues like being open-minded, informed and logically consistent. We noted, however, that such minimal reflectiveness failed to quash the potential dangers.

But what if we expanded the ambit of these epistemic virtues? A tempting solution to the problem of moral danger aims to deflect integrity's threats by moving epistemic concerns to the concept's core. The work of Chesire Calhoun, Greg Scherkoske, Andrew Edgar and Stephen Pattison all adopt this 'epistemic approach'. Though these accounts differ in some details, they agree in proposing three key qualities of integrity that work to dissolve its moral dangers.²⁵ In explaining these, I invoke *Eddie*, who we will stipulate possesses the key qualities of this epistemologically-branded integrity.

The first quality is *rational justification for convictions*: On the epistemic approach, integrity requires having rational reasons for one's views, and holding confidence in these views only to the extent warranted by those reasons. Rather than Selena's basic critical rationality (which merely ensures she is not operating in denial of obvious realities), Eddie must positively possess rational, informed reasons for all his convictions. This quality injects fallibilism into many of the beliefs that Selena could take as articles of faith.

The second quality extends the first, requiring of Eddie rational justification for his *values*. The epistemic approach demands that Eddie's values and personal commitments, as much as his factual beliefs, must be rationally justified through his ongoing and impartial interrogation. As well as extending Eddie's fallibilism to his values, this rationalist demand aims to oust any of Eddie's more egoistic or parochial commitments.²⁶

The third quality requires a commitment to a collective, intersubjective reflective process. Eddie's epistemic virtues do more than ensure that he, *Eddie*, possesses justified values, but also see Eddie playing a conscious part in a larger social-epistemic project of developing more justifiable values. By publicly standing for his values, and explaining the inter-subjectively defensible reasons he holds them, Eddie aims to help his community develop improved convictions. This feature—redolent of a progressive Millian social epistemology—explains why the community has a stake in Eddie's integrity, and thus why integrity presents as a 'social virtue'.²⁷

The epistemic approach injects these three qualities into the heart of its account of integrity, ruthlessly purging any role for the other elements of self-directedness (such as possessing willpower, a defined character and driving-values). As such, we can see how the epistemic approach quashes integrity's moral dangers. The first quality (demanding rational justifications) hammers home to Eddie the singularity of his own views. Eddie acknowledges his prized convictions are just one position among many, and their resulting fallibility must be branded into his mind. The second quality's expansion of the domain of rational justification to personal value judgments increases Eddie's epistemic hedging, ensuring he presupposes that even his 'most deeply held personal values may be flawed and imperfect'.²⁸ The requirement

for rationally defensible values also rules out self-serving commitments that cannot be inter-subjectively defended, expunging loyalty to one's self or one's projects.²⁹ Finally, the third quality (of commitment to a social epistemic process) forces Eddie to be sincere about where he stands. Eddie's values are thus avowedly fallible, rational, public and inter-subjectively justified. The danger, surely, has passed.

Despite this apparently happy result, the epistemic approach to integrity suffers from fatal flaws. In fact, some general philosophical concerns can be raised about several of the approach's fundamental assumptions. The second quality's extension of epistemic demands from beliefs to values looks particularly fraught. In order to prevent the moral dangers, these rationalist demands must apply not only to impartial moral principles, but also to what appear as highly personal value choices about life-projects, callings, enterprises, relationships, private desires and aesthetic pursuits. So when the epistemic approach demands a rational grounding for values, it is not merely requiring Eddie provide rational, public reasons for his allegiance to the impartial moral principles of human rights (say), but also that he rationally ground all his other deeply personal value choices about relationships and private desires. As Robert Noggle argues in a related context, the parallels between value theory and epistemology fail to vindicate this sort of demand.³⁰ Plausibly, if a scientific fact is true, then it is true for you as much as for me, and it makes sense to apply universal, rational and public norms about reasons and evidence to both our beliefs about that fact. But there is no analogous notion of 'truth' applicable to personal values—which is precisely why we tend to think of such values *as* personal, as indexed to the person who holds them. 'It makes me feel good and reflects who I am as a person,' may be a lousy rational defense of a belief about (say) the existence of global warming—but it is a perfectly sensible explanation of why a person devoted his life to poetry rather than boxing.

One might also harbor misgivings about the epistemic approach's sanguine presumption that shifting one's convictions in a rationally-justified direction will serve to eliminate selfish and loyalty-preserving values.³¹ The epistemic approach seems to equate being rationally justifiable with being morally justifiable, supposing that giving sensible, inter-subjectively understandable reasons amounts to providing a moral apology. But these two processes differ, and someone can sensibly argue for the rationality of self-centered behavior and parochial priorities. Indeed, many of the arguments endorsing 'loyalty-to-self' in the integrity literature (such as Williams' integrity-based attacks on utilitarianism) amount to exactly this: sincere, public, rational defenses of prioritizing one's own values.³²

Even if we set aside all such general philosophical qualms, the epistemic approach still possesses a fundamental flaw: the quality it advances as constitutive of integrity bears almost no resemblance to the term's plain meaning. Setting aside philosophical theorizing for a moment, the Oxford Dictionary defines integrity as:

1. The quality of being honest and having strong moral principles;
2. The state of being whole and undivided;
3. The condition of being unified or sound in construction.

Apart from the reference to honesty, an impartial observer would struggle to see these three elements as constitutively equivalent to 'believing only in values that one can rationally and publicly justify, as part of a collective epistemic pursuit'. The epistemic approach's core qualities actually *oppose* many of the qualities invoked by the dictionary definition. Prima

facie, vigorous fallibilism clashes with strong conviction; undertaking collective co-deliberative projects opposes beating one's own path to personal unification; and relentless open-mindedness stymies the consolidation of one's personal commitments.

The problem extends beyond the fact that the epistemic approach to integrity has changed the subject. The epistemic approach's deepest flaw is that it fails to acknowledge that we face a real ethical conundrum. We want to know, as a question of practical reason and personal ethics, if we can harness the obvious attractions of self-directedness without risking moral corruption. In proposing a response to moral danger that does not reform self-directedness, but jettisons it, the epistemic approach pays no heed to its advantages and so fails to engage with the ethical dilemma at hand.

In sum, the epistemic approach cannot see the forest for the trees. In honing in on just one facet of integrity (asking 'how does integrity avoid moral danger?'), the epistemic approach somehow forgot that integrity requires other ingredients – including integration, willpower, character, steadfastness, and personal commitment.

8. Responding to Danger: The Moral Approach

Taking a different tack, let us imagine a new figure, *Michael*. Let us stipulate that Michael shares Selena's qualities of driving-values, reflective-ownership and defining-character. Unlike Selena, however, Michael comes to realize the moral danger of his self-directedness, and he wants to inoculate himself against it. (We will consider in the next section some reasons why he might want to do this.) Not wanting to extirpate the trait entirely, he rejects the epistemic approach. Instead, Michael turns to a moral solution for what is, after all, a moral problem. He commits to the type of principles that guarantee his self-directedness no longer threatens those around him. What might these principles look like? I submit that Michael's task requires principles of trustworthiness, non-coercion and sociability.

The principle of *trustworthiness*: If Michael is fully public about what he stands for, what he does and what he intends to do, then this allows other individuals—and the society itself—to respond to his stance in a pre-meditated way. If Michael's direction becomes dangerous, they will have the opportunity to limit the harms he might cause. While others may disagree with him, they can trust him implicitly. When Michael agrees to a way forward, others know it not a mere *modus vivendi*, to be overthrown once his power grows. This principle requires more than Michael adhering to his agreements' letter and spirit; it mandates that he proactively ensures that others are not mistaken about his values or intentions.

The principle of *non-coercion*: Michael commits to respecting other people's pursuit of their own integrity, which amounts to not coercing or threatening them (he respects their 'liberty rights'). Michael will not allow his parochial commitments to determine what others do with their own lives. If Michael wants to change others' behavior, his primary recourse will be by persuasion, and by exemplifying and sacrificing for the values he holds dear.

The principle of *sociability*: Michael acknowledges the legitimacy of social decision-making processes and multilateral actions in which he plays only an equal part to others. By this commitment, Michael ensures that the privileging he gives to his own values does not metastasize into a conviction that his values should dominate others' lives as well. When it comes to dealing with others, Michael's respect for social processes forbid unilateralism. If he

must, upon pain of conscience, depart from social norms and political decisions, he does so in a measured and public way, transparent and accountable to his society.

If Michael welds these three interwoven principles into his integral values, then he disarms the six dangers of his self-directedness. His *empowerment* no longer presents a concern, as his respect for rights and social decision-making prevent him using his powers in coercive ways, and his trustworthiness ensures that society's strength can restrain him if necessary. His *egoism* and *personal calling* are now directed towards governing his own life. Even combined with his *investment*, these qualities can no longer cause him to ride roughshod over others' rights and social processes. With the principles of non-coercion and sociability forbidding his *unilateralism*, Michael's *extremism* must now act through persuasion and exemplification if it is to win adherents.

The result? In disarming the moral dangers of self-directedness, Michael has added the very traits that defined Ingrid's integrity. This result explains the first puzzle noted above (How does integrity avoid moral danger?). Through the principle of trustworthiness, Michael has added sincerity to his existing attributes of defining-character, driving-values and reflective-ownership. Michael's three-pronged morality also resolves the remaining puzzles about integrity's morality and its capacity for principled compromise.

Beginning with the *moral disagreement puzzle*, Michael's three new principles ensure he cannot be evil, nor a violent extremist or vicious ideologue. For as well as being committed to the principles of trustworthiness, non-coercion and sociability, Michael's self-directedness (in particular, the integrated nature of his driving-values) requires that he coheres these three principles with his other values and goals. This places substantive restraints on the content of those other values; the three principles function as an immovable moral anchor around which his other values must orbit.

Yet Michael is no saint: decent people may still reject Michael's values and lifestyle, even as they acknowledge his integrity. Michael might be located anywhere along the political spectrum, except perhaps the very extremes. Eschewing coercion or insincerity as he acts, and respecting the rule of law, Michael might endorse conservatism or progressivism, enlightened egoism or profound altruism, devout faith or resolute atheism, scientific rigor or romantic aestheticism, public activism or private dignity. With his commitment to the three principles, as well as to sincere pursuit of his other integrated values, we can see how Michael could achieve a reputation for admirable moral rectitude even from those who disagree with his convictions. Like Socrates, Thomas Moore, Emmeline Pankhurst and Martin Luther King Jr., Michael works with and within his social system.

Turning to the *exemplary compromise puzzle*, when a conflict arises between Michael's values and (what he recognizes through the principle of sociability as) legitimate democratic outcomes, Michael works out how he can remain steadfast in his values without disrespecting that social decision-making process. If the tension is small, Michael might go along with the policy in ways consistent with his professional integrity, even as he lobbies publicly for its repeal. If the conflict is greater, he might conscientiously object, distancing himself from the policy, but without preventing others from implementing it. When the conflict becomes unbearable, Michael might finally decide he must break the rule. But even here, he respects the process itself and the legal system that created it, by being public, non-violent and accountable

in his civil disobedience. No longer is Michael a treasonous fanatic. Instead, his behavior parallels the archetypal figures of integrity we canvassed earlier.

Summing up, we saw reasons why a person might be attracted to the personal quality of self-directedness, yet we noted the trait's serious moral dangers. When we developed a set of principles targeted to deflect these specific dangers, we saw that their addition suffices to transform self-directedness into integrity, and solve the puzzles described earlier. Ultimately, integrity contains conceptual unity because its morality shapes around moral concerns internal to the risks it would otherwise pose. Integrity's morality responds to its own dangers.

9. Normative Arguments for Integrity, So Understood

I have argued that integrity comprises the driving-values, reflective-ownership and defining-character of self-directedness, combined with ethical principles that specifically target the moral dangers that self-directedness would otherwise engender. In this final section, I show that integrity, so understood, enjoys normative appeal. I do not presume to demonstrate that all people should universally desire integrity, but instead sketch four accessible reasons why people might pursue the quality, and why societies might laud them for doing so. The reasons are: prior morality; social construction; respect for self-directedness; and human completion.

Prior morality is philosophically unexciting, but important nonetheless. One reason people pursue integrity rather than mere self-directedness is because they are already moral by the time they are in a position to begin the process. Integrity is an adult notion, a capstone of simpler childhood virtues. For this reason, most people deciding what they stand for already possess well-entrenched moral commitments and well-developed moral sensitivities.³³ When they start to reflect on what things they want to commit to, and build their lives around, they do not look forward to becoming a brutal extremist or vicious ideologue. They want to be personally driven and moral too—not least because their moral attachments would prevent them being truly integrated if they tried to repress those attachments. Adopting the three principles allows such people to ensure their newfound self-directedness will not corrupt their existing moral compass.

The *social construction* grounding holds that virtues can be formed from the outside in. In David Hume's moral theory, many virtues are formed this way.³⁴ Third-parties observe and appraise others' characters, especially in situations where they have no personal stake. Through the operation of emotions like sympathy and empathy, observers applaud those who improve the lives of their neighbors, and recoil from those who foment discord. Indeed, even a wholly self-interested agent has reason to prefer peaceable, accommodating neighbors to rapacious ideologues. As the wider society constructs its normative concepts, these countless decentralized appraisals filter its general perspective on self-directedness. True, the society can benefit from self-directed people: they are energetic, enthusiastic characters who get things done. As Mill argued, even if such determined iconoclasts swim against the tide, society still gains from them.³⁵ But society has no reason to desire self-righteous egoists who flout basic moral rules, or extremists who attack their communities. Such people present a clear threat to everyone outside their narrow coterie. This explains why societies cultivate qualities like integrity rather than self-directedness. Integrity empowers personal benefits while limiting social dangers. From this 'outside-in' basis, the concept of integrity can become normative from the individual's perspective. On Hume's telling, this occurs because of a rational-

emotional inconsistency in despising in others a trait one displays oneself. If we abhor dogmatic extremism in others, then we will not want to find it in ourselves.

Respect for integrity occurs when Michael comes to value his pursuit of self-directedness for its own sake. In this process, self-directedness becomes part of what Michael sees as the good life, valuable not just in his own life, but intrinsically. Michael admires any person's quest to unearth their own task, and respects their attempts to fashion their own rule of life. After all, if valuing self-directedness is one value we can expect to find in a person like Michael who personally pursues self-directedness, then there is a natural progression to him respecting that value when he finds it in others' lives. But as Michael moves from cherishing his own self-directedness to respecting others' free development of their personalities, he is driven to respect social peace and others' liberty rights, replacing his naked pursuit of self-directedness with the enriched goal of aspiring to integrity.

Finally, on the *human completion* grounding, we strive for integrity because we pursue the excellence specific to the type of creatures we are. As Aristotle argued, this requires becoming a complete human.³⁶ Becoming complete requires honing all the signature human faculties, rather than culling any of them. To ignore our moral dimensions and social relationships when we integrate ourselves means rejecting a natural part of ourselves, rather than cultivating the whole. At best, it is to remain incomplete. At worst, it is to leave a fissure that may later expose a hidden vulnerability. While being self-directed cultivates our personal values, character and reflective faculties, the danger looms that it can distort our public persona and wreak havoc with our moral compass. Integrity, however, allows us to unify and perfect all our human attributes: our values, our emotions and decision-making faculties, our reflective capacities, our moral sensibilities and reasoning, and our public persona, making us complete axiologically, psychologically, epistemologically, socially and morally. To forgo any of these elements to pursue the supremacy of the others would be to fail in the excellence specific to the human being. Integrity completes us in a way that self-directedness cannot.

10. Conclusion

All of us have reason to desire self-directedness in ourselves, and to envy it in others. But we have reason to be wary of it as well. Yet I have argued integrity is not similarly dangerous. Integrity presents the answer to the moral dangers that self-directedness poses.

In the final analysis, the great virtue of integrity is that it allows us to give a rule to ourselves in a world that cannot function if everybody lives only by their own rules. It empowers us to be fit for society, even as we are true to ourselves.

Endnotes

1. David Luban, "Integrity: Its Causes and Cures," *Fordham Law Review* 72, no. 2 (2003): 298.
2. Martin Benjamin, *Splitting the Difference: Compromise and Integrity in Ethics and Politics* (Kansas: University of Kansas, 1990).
3. Drawing on major strands of the integrity literature, Dudzinski includes these two features of identity formation (integral values) and principled coherence (integrated values) in her

-
- description of formalist integrity. D. M. Dudzinski, "Integrity: Principled Coherence, Virtue, or Both?" *The Journal of Value Inquiry* 38 (2004): 301-5.
4. See Bernard Williams, "Utilitarianism and Integrity," in *Utilitarianism: For and Against*, ed. Bernard Williams and J. J. Smart (London: Cambridge University Press, 1973).
 5. See Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* LXVIII (1971).
 6. See Sharon Dolovich, "Ethical Lawyering and the Possibility of Integrity," in *Professional Ethics and Personal Integrity*, ed. Tim Dare and W. Bradley Wendel (Newcastle upon Tyne: Cambridge Scholars Publishing, 2010), 146-7; Jody L. Graham, "Does Integrity Require Moral Goodness?" *Ratio* XIV (2001): 242-44.
 7. John Stuart Mill, *On Liberty*, ed. Elizabeth Rapaport (Cambridge: Hackett, 2003), 35.
 8. Benjamin, *Splitting the Difference*, 48-51.
 9. See Damian Cox, Marguerite La Caze, and Michael Levine, "Should We Strive for Integrity?" *The Journal of Value Inquiry* 33 (1999).
 10. See Colin James, "Seeing Things as We Are: Emotional Intelligence and Clinical Legal Education," *International Journal Clinical Legal Education* 8 (2005): 133; James R. Rest, "A Psychologist Looks at the Teaching of Ethics," *The Hastings Center Report* 12, no. 1 (1982): 34.
 11. See Cheshire Calhoun, "Standing for Something," *Philosophy* 92, no. 5 (1995); Luban, "Causes and Cures.,"; Greg Scherkoske, "Integrity and Moral Danger," *Canadian Journal of Philosophy* 40, no. 3 (2010); Andrew Edgar and Stephen Pattison, "Integrity and the Moral Complexity of Professional Practice," *Nursing Philosophy* 12 (2011).
 12. Greg Scherkoske, "Could Integrity Be an Epistemic Virtue?" *International Journal of Philosophical Studies* 20, no. 2 (2012): 185.
 13. Graham, "Does Integrity Require Moral Goodness?" 246.
 14. *Ibid.*
 15. See Tim Dare, "Distance, Detachment, and Integrity," in *Professional Ethics and Personal Integrity*, ed. Tim Dare and W. Bradley Wendel (Newcastle upon Tyne: Cambridge Scholars Publishing, 2010), 117-18.
 16. Cox, La Caze, and Levine, "Should We Strive for Integrity?" 520.
 17. Dudzinski, "Principled Coherence," 309; Morten Magelssen, "When Should Conscientious Objection Be Accepted?" *Journal of Medical Ethics* 38 (2012): 19-20.
 18. See Dudzinski, "Principled Coherence," 305-09.
 19. See Benjamin, *Splitting the Difference*.
 20. *Ibid.*, 36-43.
 21. *Ibid.*, 37.
 22. See Edgar and Pattison, "Moral Complexity," 96; Dolovich, "Ethical Lawyering."
 23. See Michael Davis, "Professionalism Means Putting Your Profession First," *Georgetown Journal of Legal Ethics* 2, no. 34 (1988).
 24. See Magelssen, "Conscientious Objection."
 25. Scherkoske, "Epistemic Virtue?" 196-201; Calhoun, "Standing for Something," 253-60; Edgar and Pattison, "Moral Complexity," 102-3.
 26. Calhoun, "Standing for Something," 258; Scherkoske, "Moral Danger," 354.
 27. Calhoun, "Standing for Something," 253-58.

-
28. Edgar and Pattison, "Moral Complexity," 103.
 29. Scherkoske, "Epistemic Virtue?" 201.
 30. Robert Noggle, "Integrity, the Self, and Desire-Based Accounts of the Good," *Philosophical Studies* 96, no. 3 (1999): 310.
 31. For example, Scherkoske, "Epistemic Virtue?" 201-5.
 32. See Williams, "Utilitarianism and Integrity.,"; Damian Cox, Marguerite La Caze, and Michael Levine, *Integrity and the Fragile Self* (Aldershot: Ashgate, 2009), 73-100.
 33. Dare, "Distance, Detachment, and Integrity," 118.
 34. David Hume, *A Treatise of Human Nature* (Baltimore, Md.: Penguin Books, 1969).
 35. Mill, *On Liberty*.
 36. Aristotle, *Nicomachean Ethics*, trans. Joe Sachs (Newbury, MA: Focus Publishing/R. Pullins, 2002), 1097a-98a; Christopher V. Mirus, "Excellence as Completion in Aristotle's *Physics* and *Metaphysics*," *The Review of Metaphysics* 66, no. 4 (2013).