The relationship between philosophy and evidence in education

Gina Schouten
Illinois State University, USA

Harry Brighouse
University of Wisconsin–Madison, USA

Abstract
This article explores the ways that philosophy and evidence interact in the exploration of normative questions in philosophy of education. First, the authors provide a description of reflective equilibrium, a central method in normative philosophizing. They proceed to describe three tasks of normative philosophy, each of which requires engagement with empirical evidence, albeit in different ways: identifying values, assessing institutional arrangements, and guiding policy action. They recommend increased attention from social scientists to the potential contribution of philosophy in setting the agenda for empirical research, and they recommend increased use by philosophers of real-life case studies, which enable philosophers to better understand decision-makers' feasibility constraints. These recommendations, they argue, can enable both social scientists and philosophers to better equip decision-makers to enact the best policy from among the options available to them.

Keywords
Charter schools, evidence, justice, philosophy, policy, reflective equilibrium

Philosophy of education is shifting, subtly, toward more direct engagement with the details of educational policy. We welcome this shift. Philosophy should be an under-laborer to the social sciences and practical decision-makers. Also, the field of education

Corresponding author:
Gina Schouten, Department of Philosophy, Illinois State University, Stevenson Hall 412, Campus Box 4540
Normal, Il 61790-4540, USA.
Email: rlschou@ilstu.edu

is replete with normative issues, the full import of which often is, we think, underappreciated by both groups.

This article is not driven by a single overarching argument. Instead, we lay out the different ways in which philosophy can and does engage with evidence and policy. Our aim is to clarify the relationships involved and, perhaps, make it easier for philosophers and social scientists to engage with one another.

Normative philosophy, which includes those parts of philosophy of education that will be our focus here, has three central purposes. The first is simply descriptive: Normative philosophy aims to discern, with as much precision as possible, what values actually matter, and how much they matter relative to one another. The second is evaluative: Normative philosophy aims to assist us in assessing the ways in which, and the extent to which, some situation realizes (and the extent to which it falls short of) the ideal. The final task is practical: An adequate normative theory should provide guidance for an agent considering what actions to take.

The article proceeds as follows. First, we discuss the main method of moral philosophy, reflective equilibrium, with the aim of highlighting both its promise and limitations. Next, we address the descriptive function of normative theorizing, and look at the kinds of empirical assumptions that philosophical argument concerning fundamental values relies on. Philosophers disagree about the extent to which this descriptive function requires engagement with evidence.[1] We shall suggest that the task does require *some* engagement with evidence at the level of abstraction at which philosophers of education typically work. We then briefly discuss the assessment function, looking at the way that philosophy and social science need to interact in order for us to evaluate how far and in what ways a given situation departs from the ideal. We turn, finally, to the way that philosophy and social science must interact in order to guide decision-making.[2]

Reflective equilibrium

The central philosophical method used by normative philosophers is elaborated by John Rawls; we follow him in naming it 'reflective equilibrium'. Put simply, reflective equilibrium is a process by which a philosopher sets his or her various judgments about values against each other, searches for inconsistencies among them – because he or she knows that every judgment about value is fallible – and then seeks to get closer to the truth by rejecting, revising, or reinterpreting his or her normative beliefs in light of one another. Specifically, philosophers start with two kinds of normative judgment: judgments about *particular cases* – such as 'It was wrong of Lee Harvey Oswald to assassinate JFK' – and judgments of *principle* – such as 'Political assassinations are always wrong'. What we will normally find, when we list a sufficient number of each kind of judgment, is that we do not have complete consistency: we have judgments that contradict each other. We might, for example, discover that we assent to the claim: 'Claus Von Stauffenberg's attempts to assassinate Adolf Hitler were justified'. But inconsistencies are not always so easy to find, and one central task of moral philosophy is simply uncovering them.

Of course, if two claims are inconsistent, we *could* just choose randomly and abandon one of them. But reflective equilibrium is not simply an exercise in logic; *getting closer*

to the truth is our main aim. So, having identified inconsistencies, we try to discern which of our beliefs is the least likely to be reliable. This we do by considering our relationship to the beliefs, the circumstances under which they were formed, and the perspectives of others. Our judgments are less likely to be reliable, for example, insofar as they align with (rather than contradicting) our self-interest or the interest of those whom we care about, insofar as they result from us having accepted received opinion; insofar as they are disputed by reasonable and morally decent people, and insofar as they are unstable. None of these conditions is sufficient reason to reject a judgment out of hand. Indeed, many judgments which meet these conditions are true. But when we find a contradiction, we have reasons to prefer judgments that meet none of these conditions over those that meet any of them (Daniels, 1979, 1996; Rawls, 1999).

Because none of the conditions *guarantee* that the judgment is false, further consideration can convince us to salvage it. Consider the case of received opinion. Many received opinions are, of course, true: Most of us have not thought much about whether murder or rape is wrong, or exactly what the reasons are that make them wrong – it is obvious that they are wrong and that is normally good enough. Many received opinions are less clear-cut. If we notice that we have no actual reason for holding them, we suspend our belief in them temporarily, knowing that the fact that we have no reasons for them does not mean there *are* no reasons.

Sometimes what appears to be an inconsistency is not, but the *apparent* inconsistency takes some work to dispel. Adam Swift's (2003) book, *How Not to Be a Hypocrite: School Choice for the Morally Perplexed*, has at its core the apparent inconsistency of believing *both* that private schools should be prohibited *and* that one is justified in sending one's child to a private school. In fact, as he shows, there are numerous ways of rendering those two judgments consistent. One of the tasks of reflective equilibrium is finding those ways; another is finding out whether both the two claims are, in fact, true.

The thought experiment is a central tool for identifying inconsistencies. In normative philosophy, the thought experiment consists of the description of an artificial, highly stylized, and usually unrealistic choice situation designed to elicit intuitive judgments relevant to the content of, or truth of, some principles. There are few elaborate thought experiments in philosophy of education, so we shall illustrate with the famous people seed case from Judith Thomson's 'A Defense of Abortion': Thomson asks us to consider a world in which new people come into existence in a peculiar way:

> People seeds drift about in the air like pollen, and if you open your windows, one might drift in and take root in your carpets or upholstery. You don't want children, so you fix up your windows with fine mesh screens, the very best you can buy. As can happen, however, and on very, very rare occasions does happen, one of the screens is defective; and a seed drifts in and takes root. Does the person-plant who now develops have a right to the use of your house? (Thomson, 1971)

Thomson assumes her readers will agree that it does not. But if the person-plant has a right to life but no right to the use of the house, and if the owner of the house can innocently uproot it, thereby causing its death, then its having a right to life does not suffice to show that it cannot permissibly be killed. Moreover, a reader who accepts the stipulation

that the person-plant has a right to life, but nonetheless thinks it is legitimate to uproot the person-plant must, to be consistent, abandon the belief that having a right to life is sufficient to make *aborting a fetus* impermissible.

Thomson is helping us identify a contradiction in our beliefs: We cannot both believe that it is wrong to abort a fetus on the grounds that doing so violates its right to life *and* believe that it is acceptable to destroy the person-plant. Thomson clearly thinks that the most plausible way of resolving the contradiction is by giving up the assumption that having a right to life secures one's right not to be killed in all circumstances whatsoever. But the reader might, instead, resolve the contradiction by rejecting the claim that it is permissible to uproot the person-plant. The thought experiment exposes inconsistencies, but determining how to resolve them requires further reflection.

Reflective equilibrium is not solipsistic. Our knowledge of moral matters is incomplete, and our own perspective is partial. Moreover, some of our judgments will turn out to be distorted by our social position and self-interest. We rely on one another to correct for these deficiencies by engaging with others in reflective equilibrium: reasoned engagement about moral matters with people likely to have access to moral reasons that our situation masks from us. Because those who are most different from us are likely to have greatest access to the considerations our circumstances prevent us from seeing, our ideal interlocutors are those with whom we share as little as possible in common, consistent with our actually having a reasoned dialogue.

Reflective equilibrium is nothing new; Rawls simply explicated a philosophical methodology already in practice. In fact, we think it is a methodology used by ordinary, morally thoughtful, *non*-philosophers when thinking about values. While, as specialists, philosophers may be able to offer useful resources to others, they have no special authority regarding the results of reflective equilibrium. When engaging with social scientists, decision-makers, or the general public, their job is not to present authoritative judgments, but to offer resources for their interlocutors to use in their own deliberations. In the next three sections, we discuss the three roles of normative philosophy, drawing attention to the kinds of resources philosophical theorizing can provide for decision-makers, and the ways in which those resources both depend upon and guide the production of empirical evidence.

Identifying values

G.A. Cohen says that philosophers'

> disposition to notice things in ordinary experience that other people miss means that they can . . . make a contribution to an immediately practical question. They can contribute by identifying a value that bears on choice and is being neglected. Consider an analogy. A bunch of us are trying to decide which restaurant to choose. Suppose everybody talks a lot about how good the food is in various restaurants, how much it costs, and how long it takes to get there. Someone, hitherto silent, is uneasy. She feels we have been leaving something out of account. Then she realizes what it is. 'Like, nobody', she says, 'is considering the décor'! (Cohen, 2013: 145–146)

In fact, philosophers typically see their job as being more demanding than merely *noticing* what is being left out. They *investigate and make precise* what is being left out. Because we rarely know exactly how precise we might sometimes need to be about our

values for practical purposes, we aim for more precision than will be necessary in order to achieve at least as much precision as *is* necessary. For example, we want to know what makes an education valuable for the educated person *as precisely as possible*, because we do not yet know what educational interventions might yield some, but not all, of the possibly valuable outcomes of education; we do not yet know what possibly valuable outcomes might be realizable only partially; and we do not yet know what possible valuable outcomes we might need to sacrifice for the sake of others. By clarifying the values at stake, philosophical theorizing can inform decision-making about what we ought to aim for. By clarifying those values *as precisely as possible* and by clarifying the *relative importance* of those values, philosophical theorizing can inform judgments when options are constrained and trade-offs are required.

How much engagement with evidence does this part of normative theorizing require, and with what kind of evidence? Even at the abstract level of identifying values, philosophers of education do typically need some engagement with evidence in establishing what values matter. Think again about the purposes of education. Suppose someone has the view that education should aim at promoting wellbeing. One can imagine a theory of wellbeing that is sufficiently abstract not to make much reference to, or many assumptions about, human nature. But as soon as we apply that theory to the purposes of education, we will want to know which knowledge, skills, dispositions, and attitudes should inform the design of educational opportunities in order for education to promote wellbeing. Should schooling promote personal autonomy, for example? Any answer to that, in this context, must make empirical assumptions about what human beings are like and whether autonomy does, generally, enhance wellbeing.[3]

Or consider debates about the just distribution of educational opportunities. Even if the value of equality can be established without reference to empirical facts, its relevance for the just distribution of *educational opportunities* depends on facts about the way the world works. In an environment like ours, in which education plays a large role in determining the allocation of goods across the entire life course, the distribution of education is more morally important than it would be in an environment where education is less consequential. So, in an environment in which education plays such a consequential role, the educational implications of a *general* egalitarian principle of distributive justice might be different than they would under other circumstances.

Claims about what rights children have in the realm of education also depend on empirical claims about their capacities. While it is widely thought that racially segregated schooling violates a right to an equal or an adequate education, there is less agreement on the question of what a right to an education, or to an equal or an adequate education, consists in. Establishing *that* at any level of specificity depends on empirical assumptions about children's capacities, and about the ways in which children learn, in addition to higher level normative facts regarding what human beings owe one another in general.

Merely identifying values may not require any empirical information at all. Thomson was able to clarify considerably the nature of a right to life without empirical data, and we can make some progress here by careful reflection alone. But even at the most abstract level of value-identification, getting to a level of specificity sufficient to be action-guiding requires some knowledge of empirical facts.

Assessment

A second role of normative philosophy of education is assessing or evaluating particular states of the world. Here too, philosophy of education relies on evidence. If we want to know in what ways and how far the world falls short of realizing the ideal, then we need to know not only what the ideal is, but also a good deal about the world. And precisely *what* we will need to know about the world requires careful consideration.

For example, our education system falls far short of the widely discussed meritocratic ideal which holds that students' educational success should not be influenced by their race or social class background. But in what ways does it fall short? There are considerable inequalities in average educational success between students of different races and social classes in the United States, as measured by conventional indicators like test scores and graduation rates. But it is difficult to know with any precision how much of those differences to attribute to the influence of race and how much to attribute to social class. On average, 17-year-old Black Americans perform at about the same level as 13-year-old White Americans, on standardized reading and math tests (Magnuson and Waldfogel, 2011). But this gap is not due entirely to the influence of race. Black children are, on average, much more disadvantaged socially than White children, and when we compare Black and White children from families with the same income level, the gap is much reduced. Even some of the residual gap may be due to social class differences: Black families have considerably less wealth than, and live in more disadvantaged neighborhoods than, White families with the same income levels; and even with the same income and wealth levels, Black families are more likely to be spending money on – or to face the prospect of spending money on – elderly parents (because their parents are, on average, less wealthy). The meritocratic ideal condemns inequalities regardless of whether they are due to race or to class. But for the purposes of understanding the world, and for judging which political strategies and institutional reforms will most effectively address the inequalities, we need to know the precise ways in which the world falls short of the meritocratic ideal.

In some cases, it is difficult even to assess *how far* the world falls short of the ideal. Performance on standardized math and reading tests is an imperfect proxy for educational success properly conceived. Among disadvantaged children, test performance is an imperfect predictor of the longer term life-course outcomes that education aims at, including, but not limited to, graduation from high school.[4] Additionally, the accuracy of tests in measuring educational performance depends on the regime in which the tests are introduced. Test questions are expensive to develop and must be re-used to ensure reliability, and to some extent questions are predictable, so it is always possible to teach to a test.[5] If the accountability regime in which the test is embedded has high stakes for schools and teachers, the incentive to teach to it will be hard to resist. Because test scores are imperfectly correlated with educational outcomes of value, we know that the differences in test score performance do not measure differences in success *with precision*. And because inequalities in test scores are the primary indicator of unequal average achievement between races and social classes, their unreliability undermines our capacity to discern how far the world falls short of the meritocratic ideal.

One lesson from these observations is that philosophers who want to do more than identify ideals need to be able to consume social science and need to be able to do so

critically. Another lesson, though, is that philosophical work on what is valuable in education – that is, philosophical work of the first kind that we have described – should inform the way that social scientists develop their instruments and what they choose to study, insofar as *they* want to contribute to an understanding of the ways in which the world falls short of the ideal.

Guiding decision-making

We turn now to the action-guiding function of philosophy. Insofar as we care about crafting good policy, detailed evidence matters. As Nancy Cartwright and Jeremy Hardie observe, the issue for a decision-maker is not 'What effects does Policy X, which I am considering implementing here, generally have, averaging over large numbers?' but 'What effects will Policy X have *here and now* if I adopt it and it is implemented as well as I have reason to expect it will be implemented?' (see Cartwright and Hardie, 2012). This requires a very careful reading of evidence and an understanding of the context one inhabits: in particular, judgments should be informed by an understanding of the relevant ways in which a context resembles and departs from those about which we have evidence.

Consider, for example, how children are allocated to classrooms. Suppose you are a principal and that the research literature shows that, in schools with the same student demographics as yours, on average, children with lower math skills benefit from being in classrooms with diverse skill-levels, with teachers differentiating among them. Suppose, though, that you know that *your* school has teachers who are unskilled at differentiating their instruction, and lack professional development resources to improve their capacity. While one group of teachers is highly skilled both at instruction and classroom management for remedial mathematics, another group lacks those skills but excels at teaching the Advanced Placement (AP) mathematics curriculum. Under these circumstances, de-tracked classes would probably result in substantially worse math outcomes for all students, especially for the lowest achievers. Should you de-track mathematics classrooms anyway? Unless you think that children have a right to a de-tracked classroom regardless of the consequences, you will probably think not, because you understand both the research and *your* context.

In the above illustration, conveniently, de-tracking harms many students academically, and benefits none academically. So there is no need for the decision-maker to think hard about precisely what distributive principle he or she should be guided by. But in more realistic scenarios, de-tracking may benefit some children and harm others, and the decision-maker may have reasonably good evidence about which kinds of students fall into which category. In that case, the decision-maker would need to have a more fine-grained distributive principle. Suppose, for example, that she has one group of teachers who are highly skilled at teaching both remedial and AP Math, but only to homogeneous groups, and another group who are much less skilled at teaching both remedial and AP Math. Who should get the highly skilled teachers – the remedial or the AP students – and who should be subjected to the less skilled teachers? Should the principal, in this case, impose losses to the higher achievers for the sake of gains to the lower achievers?

Philosophers cannot address every empirical situation, but they can contribute to decision-making by addressing these kinds of *conditional* philosophical questions regarding the range of *likely* effects of policies. In some cases, we can discern from the armchair whether the conditions obtain. Even when we cannot, reflection on the conditional statements enables us to identify certain considerations as relevant to policy, and identify others as *ir*relevant. And we can clarify the moral salience of various possible consequences of policies. When philosophers make judgments about social policies that are conditional on the consequences of those policies, philosophy's contribution to decisionmaking is only partial. But in *harness with* empirical data, philosophy plays an integral and consequential role: We can *use* the available evidence to make judgments about what conditions are likely to obtain, and we can *guide* the production of additional data that are important to assessing policy options. In this way, philosophical reflection can helpfully guide decision-making by drawing out implications of conditions that the available data suggest are likely actually to obtain, and by drawing out implications of conditions for which the available data are lacking, thereby illuminating underappreciated but relevant considerations and helping to set the agenda for future empirical research.

To make the best decision under the circumstances she confronts, the decision-maker needs to know *both* what values matter, to what degree, and in what relation to one another, *and* what the evidence tells her about the likely effects of whatever decision she makes on the relative realization of those values. In the remainder of this section, we use three cases to explore the relationship of philosophy and evidence in guiding decisionmaking. We show that philosophers often need to know a good deal about the details of the empirical debates in order to identify what philosophical questions need to be addressed, and in order to address them fruitfully. We then clarify philosophy's distinctive contribution to decision-making. Our three cases are varied: a debate about whether higher education should be subsidized by the state; consideration of the normative issues arising from the debates about charter schools; and Meira Levinson's discussion of student assignment within Boston Public Schools (BPS) (see Levinson, 2015).

We start with a debate among egalitarians over higher education subsidies. In most wealthy democracies, the state uses tax revenues to subsidize higher education. But those societies also contain inequalities that are (according to the participants in this debate) unjust. Furthermore, the recipients of the subsidies are, on average, more likely to be among the unjustly advantaged than non-recipients; a disproportionate number of middle-class and wealthy students, rather than poor students, attend state-subsidized colleges and universities. So subsidies appear, effectively, to deepen already unjust inequalities.

Stuart White argues that, rather than subsidizing higher education for those with the resources and capabilities to take it up, governments should give every young adult a sum of capital to be spent in whatever way the recipients choose, to 'assist their first steps into adult life' (39). White assumes that the justification of funding higher education is to meet young adults' interest in having the necessary time and resources to reflect well on the question of what they want to do with their lives. The basic capital proposal is a fairer, more egalitarian, alternative to higher education subsidies, he argues, because it enables *all* young adults to realize this interest in ambition formation, rather than favoring only those who are willing and able to undertake higher education.

Richard Arneson agrees that policies should be fair to people with diverse preferences for higher education consumption, including those who are unable to attend or uninterested in attending college. He argues, though, that we should think of fairness across a broader time frame: from infancy to adulthood. Seen from this broader perspective, a just comprehensive education policy might involve channeling extra educational resources to those less academically inclined during their *compulsory* schooling, and then – as under the status quo – favoring the *more* academically inclined through higher education subsidies. The additional educational support early on can enable those who are less academically inclined – who, in modern industrialized and post-industrial economies, are likely to be less well off all things considered – to hone the skills necessary for job market success and for converting resources to genuine wellbeing throughout the life course. Meanwhile, supporting the *more* academically inclined additionally through higher education subsidies will preserve both the personal and social benefits of higher education that, according to Arneson, a basic capital policy risks forfeiting.

White and Arneson both start with a set of empirical observations concerning who takes up, and gets the most private benefit from, higher education subsidies. Part of their disagreement concerns values: White assumes that personal autonomy facilitation is the fundamental value of higher education, while Arneson gives more weight to the wider societal benefits. But their disagreement is due, too, to their making different assumptions, based on the available evidence, about how the world actually is. For example, Arneson is more optimistic than White about the extent to which higher education serves valuable personal and social ends, rather than being merely an elaborate and socially wasteful credentialing mechanism. Because nothing very similar to White's basic capital proposal has ever been enacted, empirical conjecture is unavoidably part of their debate. Both White and Arneson are explicit about what assumptions are being made, and are careful to show how those assumptions are supported – even if not *fully* supported – by whatever relevant data *are* available.

Now turn to the debate over charter schools in the United States. Charter schools were introduced in the early 1990s as alternatives to traditional public schools; they are authorized and funded by the government, but typically have much more autonomy than traditional public schools in terms of hiring practices, curriculum and instruction, and school admissions. Currently, most states have chartering legislation, and as of 2012, 2.1 million students, or 4.2% of the public school population, attended charter schools.[6] The most interesting debates about charter schools focus on urban areas in which many children are ill-served by the traditional public schools. A small but significant set of organizations have sought to provide academically rigorous charter schools that target high-need populations in these urban districts.

Consider a specific decision-point in such a district: the decision whether to authorize a charter school affiliated with a Charter Management Organization called the Knowledge Is Power Program (KIPP).[7] High-quality (lottery-based) studies show that KIPP schools, on average, yield considerable academic gains for children who attend them. But there is considerable indirect evidence suggesting that the introduction of such schools may academically *harm* students in the community who are left behind in traditional public schools. In a separate article about this case (Brighouse and Schouten, 2014), we use these data to tease out the various considerations that bear upon the moral case for and

against charter schools. In particular, we point out that in non-ideal circumstances, many of those who are badly off are *unfairly* badly off, and so are owed some benefit as a matter of justice. Policy-makers, however, will often have to make trade-offs among those to whom recompense is owed, as circumstances often make it impossible to ease the disadvantage of some without imposing a cost on others. Society cannot make good on all obligations of justice at once, and making good on some will invariably worsen the plight of others to whom we are similarly obligated. How should we balance benefits and costs of policy options when both costs and benefits befall those to whom benefit is owed? In this case, how should we weigh the likely academic benefits experienced by high-need students who would get to attend a new KIPP school against the likely academic losses experienced by high-need students who will be left behind in the regular district schools? The correct weighting of obligations might be discernible without reference to a particular, urgent case like this, and without reference to the evidence pertaining to that case. But we suggest (and hope we demonstrate in Brighouse and Schouten, 2014) that the questions can be made sharper – and the answers clearer – by consulting and organizing that evidence in light of a specific and even high-stakes challenge like charter school authorization.

Meira Levinson's contribution to this volume concerns children's assignment to elementary schools in Boston. Levinson describes the new Home-Based Plan regulating the district-wide school choice program. All families with children entering kindergarten are assigned a 'basket' of 10–18 schools to apply to in a BPS wide lottery. The baskets must contain schools meeting various criteria of measured quality, but the system is designed so that families in more affluent neighborhoods within the city have more high-quality options available to them than families in less affluent neighborhoods. In this way, equal opportunity is sacrificed, and already-advantaged families are privileged even further. The district panders to middle-class families in response to the data that more affluent families move out of the district when their children are young because they fear their children will be assigned a bad school. Pandering is intended to staunch this flight. Pandering is, in fact, a common practice within urban districts, which often allow decisions about Talented and Gifted Programs and about drawing catchment area boundaries to be influenced by concern not to alienate more affluent families. Levinson's paper essentially conducts a moral audit of the BPS instantiation of the practice, showing what you would need to know in order to make a reasonable judgment about whether to continue with, adjust, or drop the plan. If middle-class parents can be enticed to stay, and *if their continued presence enhances the educational experiences of other students*, that provides a reason to pander to those parents. That reason must be set against the cost to equal opportunity within the district (Levinson, 2015).

All three cases involve philosophers using evidence to inform assessment of which philosophical questions need asking and to inform the answers they defend. In turn, each of the three cases is intended to guide practical decision-making. This they do, first, by clarifying what sorts of policies the existing evidence favors; second, by identifying bits of crucial evidence that are not yet available; and third, by helping to delineate the set of possibly feasible policy options.

The first and second of these tasks – determining the policies evidence supports and identifying missing evidence – are themselves difficult to disentangle. Should the three

cases be regarded as attempts to clarify what policy the available evidence supports, or as requests that additional evidence be produced? Because we are always working under conditions of *some* uncertainty that even the most rigorous and comprehensive empirical research cannot dispel, we must make probabilistic assumptions. Indeed, one role of philosophy is simply to make those assumptions explicit. While both White and Arneson take themselves to be arguing in support of a particular public policy, neither claims to have established their empirical assumptions definitively. How plausible the assumptions are – and how costly it would be to generate more definitive data – determines where we should go from here. Suppose the assumptions are questionable. If the missing empirical data can be produced, and if it can be produced at a relatively low cost, we might take the import of White's and Arneson's projects to be a call for the production of further relevant data. If, on the other hand, the empirical assumptions are highly plausible, and if the cost of producing additional data is high relative to the cost of implementing the prescribed policies absent that data, the arguments might more fruitfully be taken to support policy on their own merits. Enacting policy and attending to the consequences afterward may sometimes be less costly than producing data that would show – in conjunction with normative theorizing – that the policy is all-things-considered justified.

In our work on charter schools, we undertake to help identify the kinds of decisions regarding charter schools that the available evidence favors. We argue that proponents of charter schools like KIPP are wrong to think that documented academic benefits are sufficient to vindicate the charter school agenda, and opponents are wrong to think that possible harms to children who do *not* attend charters are sufficient to vindicate their opposition.

But we also explicitly seek to guide future empirical research. Debates over charters have focused on the academic performance of children who attend them relative to similar students who do not attend them. That is also what most of the empirical social science has focused on: the best quality studies are lottery based, and these studies, which compare students who attend the schools with other students who applied for the schools but were not admitted, tell us *only* about the performance of the students who attend the school (because the non-admitted students are the control group; there is no control group against which to compare *their* performance). So this is what decision-makers have the most evidence about. But if improving the education of disadvantaged students is indeed morally urgent, then we should also seek evidence about the academic performance of students who remain in the schools from which the charter school students are drawn. Moreover, which measurements we should rely on to assess educational outcomes depends on just *why* improving the education of disadvantaged students is so urgent. If the answer is that improving the quality of the lives *generally* of disadvantaged people is morally urgent, then we should seek evidence concerning the effects of improving education for that group of students on their *overall* life prospects. This would be especially important, and a more complex task, if we discovered that academic benefits to charter school attendees do, as we conjecture, come at the cost of academic losses to the students who continue to attend regular schools.

Additionally, in our article, we call for the production of evidence concerning the daily-lived experience of children in the different kinds of schools. Because the specifically *academic* outcomes of disadvantaged students are unlikely to be the only

outcomes of moral importance, data concerning the quality of the daily-lived experience of children in school (and the effects that experience has on their experiences outside school) must be weighed against evidence about academic gains and losses. Philosophical reflection is required in order to know how to seek that evidence: The researcher will need to know what it is about daily experiences in childhood that is really valuable before he or she can operationalize it. Finally, we call for the production of additional evidence concerning the effects of charters on students' *citizenship* education. We do not educate children simply for their own sakes: we educate them in order that they will become good citizens, able and inclined to contribute to the good of other people. That outcome is not well captured by the standard measures of academic success (typically, test scores, years of attainment, and high school graduation) that social scientists use. In all these ways, our project helps orient empirical researchers who seek to supply decision-makers with the evidence they need to inform their judgments.

Now consider the BPS case. Evidence is highly incomplete about two matters crucial to the justification of pandering (if, as we believe, the permissibility of pandering turns on its effects on the performance of lower income children). First, it is not known how many middle-class families will remain in the district thanks to the decision to pander. We are not experts, but we suspect this evidence will be hard to gather even after the fact. Second, it is not known whether the continued presence of those middle-class families will benefit lower income children who would remain in the district anyway. There are several mechanisms by way of which their ongoing presence might improve the education of lower income children: peer effects in schools, the maintenance of property values (and thus district revenues), and the political capital the district has thanks to middle-class parents lobbying to increase resources and improve educational provision. It is not given, however, that these effects will be significant, or even positive. The effects on property values may be small, and it is possible that middle-class parents will use their lobbying power in ways that simply transfer existing resources to their children (e.g. by lobbying to have their own children put in the classes of the better teachers, at the expense of other children). Again, understanding the normative terrain helps researchers orient themselves to the missing evidence, and helps decision-makers understand what they need to make judgments about in the absence of better evidence.

The third way that philosophy contributes to decision-making is by helping to clarify the set of feasible policy options we must choose among. Contemporary philosophy of education – and educational theory more generally – contains a good deal of work that is critical of contemporary developments in education policy. School vouchers, charter schools, No Child Left Behind (NCLB), the Common Core State Standards, and high-stakes tests are all frequent targets of criticism.[8] But criticism can be ambiguous between *assessment* and an *attempt to guide action*.

Suppose it is true that, because charter schools are in place, the world falls further short of the ideal than a world in which charter schools were not in place. It follows neither that we should now eliminate charter schools, nor that someone should vote against any particular charter school. We have to ask what the available alternatives are. Rarely do our feasible alternatives include any that will fully realize the ideal. Consider the BPS case. Pandering is, of course, morally problematic, and it would be better if the district

could benefit lower income children without pandering to more affluent families. Such aims could conceivably be achieved by state-level action that reduces the relative advantage of suburban districts near Boston, or that consolidates Boston with surrounding districts. It is also within the power of the State to create incentives to change patterns of residential segregation. But the school district itself cannot engineer these changes; its options are highly constrained. Given these conditions, the school district may continue to need to pander in order to benefit lower income children.

Attention to the decision-maker's feasibility set gives *focus* to critique. Even assuming that one has exactly the right values, and has made exactly the right assessment of the extent to which the situation realizes those values, the moral authority of one's criticism of that situation is, other things equal, directly proportional to the clarity and feasibility of one's alternative proposal. Focusing on a case helps us to see how assessment and action-guidingness are quite different functions; success at the former does not predict success at the latter task.

Note that philosophical reflection can sometimes *help make* an option feasible by alerting us to unconsidered possibilities. With no political coalitions advocating it, Stuart White's basic capital proposal is not, now, feasible. But working out the details and the normative rationale, and bringing the proposal into policy debate may help to make it feasible, or trigger other proposals to mitigate the unfairness built into higher education subsidies.

Our work on charters, in contrast, offers reasons for *skepticism* that a certain policy option is really feasible. In response to worries that charters, on average, harm the students left behind in traditional public schools, charter enthusiasts might propose taking charters to scale so as to enable *all* schools to adopt the educational innovations that drive the success of charters like KIPP. In fact, several factors make it unlikely that the benefits of charter schools could feasibly be provided at scale: First, successful charter schools appear to rely on a scarce supply of very talented and energetic teachers and administrators. Second, many of these schools appear to rely on a scarce supply of private and philanthropic funding. Finally, we conjecture that some of the signature features of a charter education to which their academic successes are most plausibly attributed serve subtly to exclude the very most disadvantaged students. These features include admission by an opt-in lottery enrollment; a rigorous academic and disciplinary regime that is maintained by demands on parents and students that the most disadvantaged may have difficulty meeting; and expulsion as a consequence for misbehavior – a consequence that, in all likelihood, effectively excludes some of the most difficult-to-educate students. Because of the importance of peer effects, it is plausible to think that the success of charters is due, in part, to their exclusion of some extraordinarily disadvantaged students.

By cataloging reasons to be skeptical that charters can be brought to scale without forfeiting the academic gains that make them so appealing in the first place, we hoped to bring additional focus and clarity to the discussion over charter schools. The policy option that we argue is infeasible is one that may initially have appealed to many theorists, researchers, and decision-makers committed to making the provision of education more just. We hoped to redirect attention to policy options more likely to be viable in our social and political context.

In each case considered, the philosopher takes a very specific decision-point, and asks what values are pertinent, and how much weight those values should have. By adopting that point of view, and appreciating the details of the case, philosophers can inform decisions by aligning normative investigation with the practical needs of the decision-maker, thus enabling the decision-maker to assess the different weights of the values at stake in the specific circumstances. Unless pandering is always and absolutely impermissible, or always and absolutely morally acceptable, we need to weigh the bad that it both represents and does, against the good that it does. This requires us to reflect on the salient values, and scrutinize (and, perhaps, call for more) evidence. Familiarity with the empirical context shows us which normative questions need answering.

The illustrations also reveal something important about the difference between philosophical reflection informed by case studies, on the one hand, and philosophical reflection informed by thought experiments on the other. Thought experiments are *entirely stylized*. They are typically aimed at prompting the kind of reflective equilibrium that helps us to illuminate the exact content of some principle or value. Case studies, in contrast, elucidate what is at stake for an actual agent. They focus our attention on a specific and textured context, in which competing values merit consideration and must be weighed against one another. Insofar as philosophers aspire to inform decision-making, actual case studies informed by empirical data can help by alerting us to the most pressing questions that require attention, and by alerting us to salient features of the circumstances we confront. When we presented a different case about discipline to an elementary education major who has some background in philosophy, this was her illuminating if depressing response:

> I wish they would give us more readings like those in the school of ed, they are much more realistic while most of the readings we do are more idealistic . . . Actually I think that tension is something I struggled with a lot throughout the program-but didn't fully understand why it was so frustrating to me. In my practicum I would see my teachers facing problems like this one-and the other behavior case every day-multiple times a day. Then in our content classes these very real problems were almost watered down, and approached in terms of ideal theory. We talked about the benefits of all-inclusive classes, being preventative, and reflecting in action. But we never really had conversations about how this looks in imperfect practice.[9]

Conclusion

This discussion reveals certain insights about the relationship between philosophy and empirical evidence, both of which, we have suggested, are necessary to inform morally responsible decision-making. To guide decision-making efficiently and well, normative theorists and empirical scholars must work 'on both ends'. Philosophy clarifies fundamental values, identifies relevant bits of missing data, and illuminates feasible policy options likely to be overlooked. In this way, philosophy *informs and guides* the empirical research agenda. But even in so doing, it *uses* empirical data. As we have seen, even the task of elucidating and clarifying the relevant values requires familiarity with some empirical evidence. But as they move from the task of identifying values, to assessing existing arrangements, to guiding policy-making, philosophers should be attuned to

more, and increasingly sophisticated, empirical evidence. In most cases, to produce work that is of use to decision-makers, philosophers need to know a good deal about the details of the empirical landscape.

Informed by an understanding of the relevant empirical data, philosophy has much to offer policy-making. We conclude with a brief survey of some educational policy issues, the discussion of which we think would be enhanced by the engagement of philosophers. This is not meant as an exhaustive list; in fact, we think that philosophical reflection could probably contribute productively to most live debates pertaining to educational policy. But mentioning a few in particular will be illustrative.

First, what contribution can philosophy make at the level of *identifying values*? We think that there is important philosophical work to be done in clarifying precisely what makes an education valuable for the student herself or himself. The academic outcomes of education that are typically used to measure the quality of an education are only imperfectly correlated with the quality-of-life outcomes that are part of what justifies large-scale public expenditures on education – especially *compulsory* education – in the first place. Plausibly, some educational interventions might impact on the quality-of-life outcomes without affecting immediate academic outcomes, and some may impact on the immediate academic outcomes without changing quality-of-life outcomes. Philosophical reflection that aims at uncovering and clarifying values can help us discern the outcomes that are ultimately important in our circumstances, and thereby help us discern the most desirable interventions – or, at least, it can help us discern the criteria by which we should go about assessing interventions. If, as we suspect is the case, quality-of-life outcomes are ultimately what matters morally speaking, then philosophical reflection on wellbeing becomes crucially implicated in morally responsible decision-making.

Second, how can philosophy help us *assess existing institutions* to determine how far and in what ways they fall short of the ideal? Philosophy can fruitfully be brought to bear on discussions of achievement gaps. There are inequalities between groups of students along the dimensions of social class, race, gender, and academic aptitude, and philosophers can help in discerning which inequalities constitute injustices, and in assessing the relative urgency of those injustices. For example, should scarce educational resources be expended to enhance the achievement of naturally less academically inclined students, to reduce the achievement gaps between students of different academic aptitudes? In what ways is the answer to that question dependent on more general features of the social environment, including the extent to which *other* social goods are unequally distributed, and the extent to which their distribution is mediated by unequal academic success? What is the relative urgency of academic benefits owed to students disadvantaged along various dimensions? How urgent are inequalities in academic engagement that are created when children with high aptitudes get bored in classes with poor differentiation? What do we owe children who have the capacity to make great achievements, but need special circumstances in which to realize these capacities? How should we balance the individual rights of children against the achievement of societal goals? Answering these philosophical questions and others is crucial if we are accurately to assess how far short our institutions fall of achieving educational justice, and in what precise ways they fall short. And that assessment is, in turn, crucial if decision-makers are to prioritize projects and allocate scarce resources in morally responsible ways.

Finally, how else might philosophy contribute at the level of *guiding decision-making*? Consider affirmative admission practices as just one example. We might worry that the primary beneficiaries of affirmative admission policies are often not those whom the policies are intended to help,[10] or that such policies will result in unintended and harmful consequences like tokenism or hostility toward the intended beneficiaries. What importance would any of these possible consequences have in assessing how, if at all, educational institutions should take demographic group memberships into account in allocating scarce admissions spots? One important consideration, surely, is what the feasible alternatives are; as we have seen, philosophy can play a crucial role in either narrowing or expanding the range of possibilities under consideration. Are there ways to attain the goods of integration while avoiding some of the ills of typical affirmative policies? Philosophy can also help by clarifying more precisely what the goods of integration *are* in the first place. Some social goods plausibly attainable by affirmative admission policies might be worth pursuing even at the costs conjectured above. Does society more broadly reap benefits insofar as its educational institutions are integrated across race and class?[11] Answering these questions, and determining the relative importance of their answers, is an exercise in normative theorizing, and decision-makers will be better able to make morally responsible choices insofar as they have available to them fuller, more consistent, and more accurate answers to these philosophical questions (see Blum, 2015).

We think that philosophers of education have much to contribute in terms of guiding research and decision-making. But philosophers must attend to facts about the world in order to perform this task well. They must pick representative cases; learn the evidential and institutional terrain of those cases; take the perspective of the decision-maker; and discipline any critique with appreciation of what is, really, within the decision-maker's control.

Acknowledgements

Funding

Notes

1. G.A. Cohen (2003) argues that, strictly speaking, describing values at the highest level of abstraction involves no engagement with facts at all, for example, whereas John Rawls thinks that it must be informed by the basic facts about human nature.
2. Note that we think our approach is very close to Swift (2008).
3. For elaboration on this point, see Brighouse et al. (in preparation).
4. We know this from various interventions, like Perry pre-school and the Abecedarian Project, in which test scores improved in the short term, but faded out quickly, while other benefits, not captured in tests scores (e.g. lack of involvement in the criminal justice system, early and middle adulthood employment and earnings, levels of physical and mental health) persisted. See, for example, Heckman et al. (2010).

5. See, for example, Jennings (2012). See also Rothstein et al. (2008), chapter 4, for an elegant discussion of a general problem for accountability regimes: that whatever measures are chosen as proxies for performance cease to be such good proxies once organizations are held accountable for them.

6. US Department of Education Institute of Education Sciences, National Center for Education Statistics. Accessed on 07/24/14 at http://nces.ed.gov/fastfacts/display.asp?id=30.

7. For a book-length account of Knowledge Is Power Program's (KIPP) development, see Mathews (2009).

8. For example, Howe (2008), Siegel (2004), White (2013), Apple (2006).

9. From personal (Brighouse) email correspondence with Katherine Nahn, dated 21 May 2014 and 10 June 2014; used with permission.

10. For example, if the primary beneficiaries of affirmative policies are relatively advantaged members of the targeted group, but the aim of those policies is to improve inclusion of the *disadvantaged* members.

11. See Anderson (2007).

References

Anderson E (2007) Fair opportunity in education: A democratic equality perspective. *Ethics* 117(4): 595–622.

Apple MW (2006) *Educating the 'Right' Way: Markets, Standards, God, and Inequality*. New York: Taylor & Francis.

Arneson R. *Justice in access to higher education* (Unpublished manuscript).

Blum L (2015) Race and class categories and subcategories in educational thought and research. *Theory and Research in Education* 13(1): 87–104.

Brighouse H and Schouten G (2014) To charter or not to charter: What questions should we ask, and what will the answers tell us? *Harvard Educational Review* 84(3): 341–364.

Brighouse H, Ladd H, Loeb S, et al. (in preparation) *Educational Goods*.

Cartwright N and Hardie J (2012) *Evidence-Based Policy: A Practical Guide to Doing it Better*. Oxford: Oxford University Press.

Cohen GA (2003) Facts and principles. *Philosophy and Public Affairs* 31(3): 211–245.

Cohen GA (2013) *Finding Oneself in the Other*. Princeton, NJ: Princeton University Press.

Daniels N (1979) Wide reflective equilibrium and theory acceptance in ethics. *The Journal of Philosophy* 76(5): 256–282.

Daniels N (1996) *Justice and Justification: Reflective Equilibrium in Theory and Practice*. Cambridge: Cambridge University Press.

Feinberg and Lubienski C (eds) *School Choice Policies and Outcomes: Philosophical and Empirical Perspectives on Limits to Choice in Liberal Democracies*, pp. 61–78. Albany, NY: State University of New York Press.

Heckman J, Moon S, Pinto R, et al. (2010) The rate of return to the High/Scope Perry Preschool Program. *Journal of Public Economics* 94(1–2): 114–128.

Howe K (2008) Evidence, the conservative paradigm, and school choice. In: W Jennings JL (2012) The effects of accountability system design on teachers' use of test score data. *Teachers College Record* 114(11): 1–23.

Levinson M (2015) The ethics of pandering in Boston Public Schools' school assignment plan. *Theory and Research in Education* 13(1): 38–55.

Magnuson K and Waldfogel J (2011) *Steady Gains and Stalled Progress*. New York: Russell Sage Foundation.

Mathews J (2009) *Work Hard. Be Nice: How Two Inspired Teachers Created the Most Promising Schools in America*. Chapel Hill, NC: Algonquin.

Rawls J (1999) *A Theory of Justice*, rev. edn. Cambridge, MA: Harvard University Press.

Rothstein R, Jacobsen R and Wilder T (2008) *Grading Education: Getting Accountability Right*. New York: Teachers College Press.

Swift A (2003) *How Not to be a Hypocrite: School Choice for the Morally Perplexed Parent*. New York: Routledge.

Swift A (2008) The value of philosophy in non-ideal circumstances. *Social Theory and Justice* 34(3): 363–387.

Thomson JJ (1971) A defense of abortion. *Philosophy and Public Affairs* 1(1): 47–66.

White J (2013) Philosophy, philosophy of education, and economic realities. *Theory and Research in Education* 11(3): 294–303.

White S (2010) A modest proposal? Basic capital vs. higher education subsidies. *The British Journal of Political and International Relations* 12: 37–55.

Author biographies

Gina Schouten is Assistant Professor of philosophy at Illinois State University. Her interests include educational justice, political legitimacy, and gender in the family. Her most recent research has concerned the moral foundations of objections to achievement gaps, what principles should be included in a theory of educational justice, and how best to conceive of the basic structure of society for the purposes of theorizing about justice. She is currently working on projects concerning the legitimacy of political interventions to alter the gendered division of labor.

Harry Brighouse is Professor of philosophy at the University of Wisconsin–Madison. His recent research has focused on the ethics of interactions between parents and children, the moral questions concerning school choice, and justice in the distribution of educational opportunities. His books include *School Choice and Social Justice* (Oxford, 2000), *On Education* (Routledge, 2006), and, with Adam Swift, *Family Values: The Ethics of Parent-Child Relationships* (Princeton, 2014).