

**The Virtues of Ethics Bowl: Do Pre-College Philosophy Programs Prepare
Students for Democratic Citizenship?**

Michael Vazquez¹ and Michael Prinzing^{1,2}

¹ The Parr Center for Ethics at University of North Carolina at Chapel Hill

² Yale University

[Accepted for publication in *The Journal of Philosophy in Schools.*]

Abstract

This paper discusses the rationale for, and efforts to quantify the success of, philosophy outreach efforts at the University of North Carolina at Chapel Hill. We focus on the National High School Ethics Bowl (NHSEB) since the uniformity and structure of this program supports rigorous assessment. We begin by articulating the democratic foundations of Ethics Bowl and reflect on the civic and intellectual virtues that this activity might promote. We then describe our efforts to empirically assess the impact of NHSEB on students' intellectual and civic virtues—including a pilot study conducted in April 2022, which laid the foundation for an ongoing, quasi-experimental study on NHSEB programs across the United States. We then offer recommendations for research design to empower publicly and empirically engaged philosophers to conduct further studies, including alternative methods, outcomes, and study designs. In articulating the possibilities for further research, we discern a related but distinct rationale for philosophy outreach programs that we call the “moral rationale.” We briefly articulate this moral rationale and empirical outcomes possibly associated with it.

The Virtues of Ethics Bowl:

Do Pre-College Philosophy Programs Prepare Students for Democratic Citizenship?

In the last two decades, university-based outreach programs have become an increasingly common fixture among philosophy departments. Many of these programs aim to promote the public character of higher education, or the status of the university as a public good. According to this rationale, philosophy outreach is a form of public service through which the university can fulfill its obligation to promote social and democratic goods, such as civic virtue and intergenerational bonds (see, e.g., Jared 2021). This is a laudable aim, but to what extent do these programs achieve it? In this paper we contribute to answering this question by focusing on the impact of the National High School Ethics Bowl (NHSEB), a large-footprint outreach program based out of the University of North Carolina at Chapel Hill whose uniformity and structure facilitates empirical assessment.¹

In what follows, we articulate the democratic foundations of Ethics Bowl and reflect on the civic and intellectual virtues that this activity might promote. We then describe our efforts to empirically assess the impact of NHSEB on students' intellectual and civic virtues—including a pilot study conducted in April 2022, which laid the foundation for an ongoing, quasi-experimental study on NHSEB programs across the United States. We then offer recommendations for research design to empower publicly- and empirically-engaged philosophers to conduct further studies, including alternative methods, outcomes, and study designs. In articulating the possibilities for further research, we discern a related but distinct rationale for philosophy outreach

¹ For one model of Ethics Bowl outreach that uses an undergraduate service-learning class, including a sample syllabus, see Vazquez (2022). In this paper we alternate between 'Ethics Bowl' as a general form of an activity that is utilized across the lifespan and the 'National High School Ethics Bowl', the program headquartered in and governed by the Parr Center for Ethics at the University of North Carolina at Chapel Hill. NHSEB was founded in 2012 and is modeled on the Intercollegiate Ethics Bowl, pioneered by Dr. Robert Landeson in 1993.

programs that we call the “moral rationale.” We briefly articulate this moral rationale and empirical outcomes possibly associated with it.

The Democratic Foundations of Ethics Bowl

One of the undisputed aims of education is to prepare students to participate fully in democratic life (Allen 2016; Gutmann & Ben-Porath 2014). After all, many theorists consider democracy to be more than a set of procedures and constitutional mechanisms. It is also an ideal of cooperative and egalitarian relations and decision-making at all levels of social life. What does this mean for the kinds of traits and dispositions civic education should foster? Naturally, there is extensive disagreement about the details. Yet the Ethics Bowl relies on a fairly ecumenical vision of democratic citizenship, the core features of which are contained in Gutmann and Thompson’s classic statement of the deliberative ideal of democracy:

...we can define deliberative democracy as a form of government in which free and equal citizens (and their representatives), justify decisions in a process in which they give one another reasons that are mutually acceptable and generally accessible, with the aim of reaching conclusions that are binding in the present on all citizens but open to challenge in the future. (Gutmann & Thompson 2004, p. 7)

In a deliberative democracy, citizens should not simply vote, but also engage in extensive reflection and civic dialogue. They justify their positions to one another with the aim of finding common ground. Moreover, they do so in the face of the seemingly intractable disagreements about morality and politics that are characteristic of pluralistic and multicultural democracies.

Deliberative reciprocity, openness to being moved by another’s claims, and the commitment to minimizing disagreement where possible promote the core value of mutual respect, partly by manifesting “relational” or “democratic” equality (Anderson 1999). Citizens who are seriously committed to the political project of collective decision making will offer reasons that are accessible and intelligible to those who do not share the same overarching moral, religious, and philosophical commitments—“public reasons” as Rawls (1997) called them. Such citizens will

engage in this process of mutual justification with the recognition that their own judgments and the results of collective deliberation are often mistaken and in need of revision (Gutmann & Thompson 2004, p. 6). In other words, *intellectual humility* is a crucial civic virtue.

Intellectual humility has become a topic of active discussion across academic disciplines and there remain disagreements about what exactly it looks like and what forms it takes (Davis et al. 2016; Leary et al. 2017; Tanesini 2018; Whitcomb et al. 2017). Even so, many people are coming to see intellectual humility as important for democratic citizenship. As Jeffrey Rosen, President of the National Constitution Center, argued:

We all have a duty to cultivate our faculties of reason, as Jefferson said, so that we can achieve our fullest potential as human beings and as citizens. *The only way that we can cultivate our faculties of reason is to have the humility to recognize that our first, impulsive, passionate opinions may be wrong.* We have to listen respectfully to arguments on the other side. We have to respect our fellow citizens enough to understand that their opinions too deserve weight...*Intellectual humility can be taught by being modeled.* Students of all ages, citizens of all ages, have to experience the process of a respectful debate and of listening respectfully to arguments on all sides in order to achieve the habits of intellectual humility. *To engage in the civic duty of intellectual humility, we need time, and we also need a space where all sides are respectfully represented* (transcription of National Constitution Center, 2019; emphasis ours).

Rosen seems to conceive of intellectual humility as being part of a network of virtuous habits, including respectfulness, reflectiveness, and open-mindedness. But he also suggests the striking idea that cultivating intellectual virtue is a “civic duty.” This implies that the ideal democratic citizen has certain cognitive traits—having cultivated good habits of mind and of inquiry—but also affective traits. They are not aggressive or hostile towards people with whom they disagree, and they show empathy and compassion to their fellows. These are all traits that a program like the Ethics Bowl might cultivate.

Educating for Citizenship with Ethics Bowl

These virtues must be modeled for and inculcated in students (Gutmann & Ben-Porath 2014). Ethics Bowl is designed to do just that (Ladenson 2001; Richardson 2022). What began as an intramural college competition has become an educational tool that impacts folks across the lifespan (Lee 2021 provides a sampling of Ethics Bowl programs; Israeloff et al. 2022 includes chapters about efforts to use Ethics Bowl for populations beyond schooling years). Our focus is on the National High School Ethics Bowl and the extent to which it effectively promotes the intellectual and civic virtues outlined above.

The basic aims of NHSEB are encapsulated in the program's mission statement:

The National High School Ethics Bowl promotes respectful, supportive, and in-depth discussion of ethics among high school students nationwide. By engaging high school students in intensive ethical inquiry, the NHSEB fosters constructive dialogue and furthers the next generation's ability to make sound ethical decisions. Our collaborative model rewards students for the depth of their thought, their ability to think carefully and analytically about complex issues, and the respect they show to the diverse perspectives of their peers. As a result, it enables students to practice and build the virtues central to democratic citizenship, thus preparing them to navigate challenging moral issues in a rigorous, systematic, and open-minded way. (nhseb.unc.edu/mission-history)

We can think of an NHSEB match as a microcosm of deliberative democracy.² Students deliberate together, in teams, on ethical and political questions pertinent to life in contemporary democracies. The team-based format is premised on the idea that, as individuals, we are highly fallible. Hence, reasonable disagreement about moral and political matters persists even after people have earnestly deliberated on their own. Indeed, there is ample empirical evidence that

² Notably, NHSEB is similar to formats used in recent studies that have investigated how deliberative norms influence the quality and character of political discourse (Fishkin 2018; for studies involving younger people, see McAvoy & McAvoy 2021 and Hess & McAvoy 2015).

people reason better in groups (Dutilh Novaes 2020). Consider, as a simple but striking example, the “Wason Selection Task.” This is a logic puzzle in which respondents are shown four cards and given a (material conditional) proposition about what appears on each side of the cards. They are then asked which card(s) they would need to flip over to determine whether the proposition is true. Only about 10-20% of individuals solve it correctly. By contrast, groups of five or six solve the puzzle correctly about 70-80% of the time (Moshman & Geil 1998). Other empirical work has found that groups reason best when governed by collaborative, egalitarian norms, and when the members have strong socioemotional skills (Trouche et al. 2014; Woolley et al. 2010).

Yet, even groups make mistakes. Hence, Ethics Bowl matches bring teams together for further deliberation and dialogue. The scoring criteria for the matches reward collaboration and penalize adversarial deliberative norms. They are designed to incentivize intellectual virtues like humility, charity, and open-mindedness. For instance, scoring criteria include: “Did the team’s presentation indicate both awareness and thoughtful consideration of different viewpoints, including especially those that would loom large in the reasoning of individuals who disagree with the team’s position?”; “To what extent has the team effectively and directly responded to and engaged the presenting team’s argument?”; and “Did the team demonstrate their awareness that an ethics bowl is about participating in a collaborative discussion aimed at earnestly thinking through difficult ethical issues?”

A typical NHSEB match is broken up into two rounds. Each round begins with the announcement of a case from the set teams prepared in advance and an accompanying “moderator question,” which is the specific question both teams are tasked with attempting to answer over the course of the round. The conversation begins with an initial presentation from Team A. Team B then comments on that presentation. Team A responds to Team B’s commentary. The round then concludes with a real-time Q&A session with a panel of judges (typically three). The second half of the match repeats the process, but with a new case and

moderator question, and with Team B serving as the initial presenting team and Team A serving as the commentating team. The first half of the match has the following structure:

Moderator's Period: The first case and its corresponding moderator question are announced (customarily, neither the case nor the moderator question are known in advance to judges and participants).

Presentation Period: Team A has time to confer before offering a *presentation* in answer to the moderator's question.

Commentary Period: Team B has time to confer before *offering a commentary* on Team A's presentation.

Response Period: Team A has time to confer before *responding* to Team B's commentary.

Judges' Period: The judges lead a real-time *question and answer session* with Team A (i.e., the initial presenting team for that round).

An NHSEB match is modeled on an idealized and structured *conversation* about practical problems. As a conversational form of deliberation, its governing norms are different from those of debate (Laden 2012; McAvoy & Lowery 2022). Teams are not assigned positions, nor are they expected to disagree with one another. Issues are not presented as binary choices, but as complex and multifaceted tradeoffs and dilemmas with several layers of normative ambiguity (the moderator questions selected for the matches are often written with this explicit purpose in mind). The "Commentary Period" and "Response Period" are designed to facilitate a genuine exchange of ideas and to encourage susceptibility to the reasoning of others, rather than digging in one's heels at all costs. In order to 'win' a team must demonstrate that they are *responsive* to the other team and to a panel of judges. There is ample room for vigorous disagreement within this format, but such disagreement must be guided by good will towards others and tempered by humility. Comments and criticism must deepen and strengthen the analysis the other team has offered, and ultimately move the group closer to a satisfactory resolution of the ethical dilemma at hand,

or at least a shared understanding of the reasonable positions on offer. In short, Ethics Bowl creates opportunities for learning how to be democratic equals by modeling it (Laden 2013).

Moving Past Anecdota

Ethics Bowl is self-consciously styled as a moral and civic education program that serves as an antidote to social ills that threaten democracy. As we've discussed, it purports to do so by providing opportunities to cultivate democratic competencies and dispositions that in turn will ensure the stability and resilience of the civic sphere. But how successful is it?

In conversations with students, as well as educators and volunteers, we've heard many stories about the transformative impact of NHSEB on individuals and communities. Yet, despite Ethics Bowl's widespread success at the high school and collegiate levels (and now increasingly at the middle school level), there is little empirical data to shore up these anecdotes.³ Accordingly, the Parr Center for Ethics at UNC-CH has begun a multi-year effort to advance our understanding of the impact of NHSEB, going beyond "anecdota" by gathering more rigorous empirical evidence.

National Competition Pilot Survey

In April 2022, we ran a pilot test with students at the NHSEB Championship. We invited students who attended this event to complete a survey that asked several questions about their experience in NHSEB. These questions gave us some insight into what the students themselves think about the effects of NHSEB. Additionally, in preparation for a larger study, we included measures of some of the intellectual and interpersonal traits discussed in the previous section: reflectiveness, intellectual humility, recognition of one's epistemic limitations, and attitudes towards people who disagree about ethical and political matters (or "affective polarization"). The

³ To the best of our knowledge, there has been no systematic study of the impact of Ethics Bowl on students. There is, in contrast, a substantial body of empirical research on the Philosophy for Children (P4C) dialogic pedagogy. Precedent for assessing the impact of pre-college philosophy goes back at least to Lipman (1973). Relatively recent reviews of research can be found in Millett & Tapper (2012) and Topping & Trickey (2015).

measures we included have been validated in previous empirical research. Yet, because these instruments were designed for adults, we hoped to verify that they would work well with high school students.

All of the 62 students who completed this survey said that their critical thinking skills had improved as a result of participating in NHSEB. Indeed, a majority (73%) said that their critical thinking skills had improved “a lot.” Similarly, 90% of these students said that their ethical and political beliefs had changed as a result of participating in NHSEB. A majority said that their ethical and political beliefs had changed either “a little” (47%) or “moderately” (42%). Nearly all (97%) also said that they had found a community and sense of belonging in the NHSEB program, with most indicating either “moderately” (36%) or “a lot” (47%). Naturally, these retrospective assessments cannot be taken as gospel. But they certainly highlight the promise of NHSEB and invite further empirical investigation into how this program affects students’ intellectual and civic virtues.

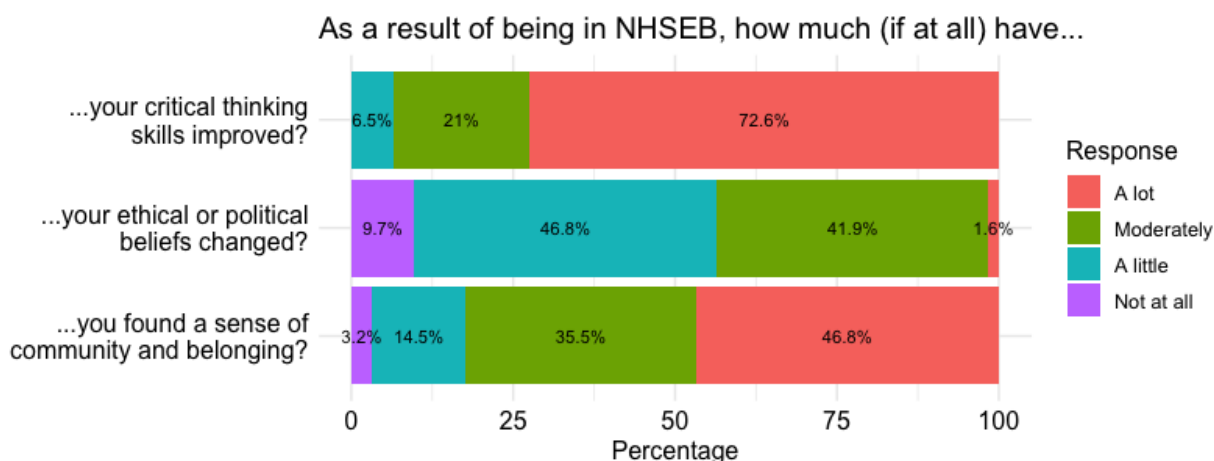


Figure 1. Self-Reported Outcomes of NHSEB Participation from 2022 National Competition Pilot Survey

Of the measures of intellectual traits, two were self-reports. First, the Intellectual Humility Scale (Leary et al. 2017) asks respondents to indicate the degree to which five statements sound like them—for example, “I recognize the value in opinions that are different from my own.” Second, we included an affective polarization “thermometer” measure, which asks about respondents’

attitudes towards people who disagree about ethical and political issues (Iyengar & Westwood 2015). The response scale is a slider, ranging from “Cold / Unfavorable” to “Warm / Favorable.”

The affective polarization measure performed excellently. Responses spanned the full range of the response scale, with an approximately normal distribution. Such a large amount of variance and “bell curve” distribution are ideal for statistical analyses. However, scores on the Intellectual Humility Scale were exceptionally high. In fact, none were below the midpoint on the scale. This may be due to the fact that intellectual humility is frequently and explicitly praised during the National Championships. We should see greater variability and a lower average score in other contexts. Nevertheless, this result is a reminder about the notorious difficulty of measuring traits like humility through self-report. Hence, we decided that in future work we would include measures of “social desirability bias”—i.e., the tendency for a person to answer questions dishonestly in order to look good. By measuring this bias, we will be able to control for it in statistical analyses.

The other two measures in the pilot survey were tests, rather than self-reports. One was a variant on the Cognitive Reflection Test (Frederick 2005), which includes a series of questions that have initially intuitive but incorrect answers. One well-known example is, “A bat and ball cost \$1.10. The bat costs one dollar more than the ball. How much does the ball cost?” Many people initially think \$0.10. However, the correct answer is actually \$0.05. The second test measured the tendency to “over-claim” knowledge (Paulhus et al. 2003). Respondents are presented with a list of people, places, and ideas (e.g., Marie Curie, Kyoto, and gerrymandering) and asked whether they are familiar with each. Some items on the list do not exist. An “over-claiming” score can then be computed from the number of non-existent things that a respondent claims to be familiar with.

The students performed quite well on the reflectiveness test. Most answered at least 3 out of 4 questions correctly. In light of this, we decided that, in future studies, we would extend the length of the test thereby increasing the range of possible scores. We decided to use both the original Cognitive Reflection Test (Frederick, 2005) and a more recently developed alternate

version (Thomson & Oppenheimer 2016). On the over-claiming test, about a third of students received a score of 0 (i.e., didn't over-claim at all). However, most students claimed familiarity with at least one non-existent thing, suggesting that this is a viable, non-self-report measure of students' awareness of their epistemic limitations.

The average scores on these measures were fairly high. We speculate that this may be a result of the fact that these students were finalists at the NHSEB National Championship—i.e., top performers in a program that we hypothesize will cultivate these virtues. Hence, it is perhaps unsurprising that they scored so well on these measures.

2022-2023 Impact Study

In light of the promising results from the pilot survey, during the 2022-2023 NHSEB Season we launched a nationwide, longitudinal, quasi-experimental (i.e., controlled but not randomized) study. The aim is to quantify the impact of the program on the development of students' intellectual and civic virtues by monitoring these traits in students who participate in NHSEB as well as students from the same schools who do not participate. Do students who join NHSEB differ from their peers at the start of their participation? Do they show growth in traits like intellectual humility and reflectiveness? If so, how does the rate of growth for NHSEB students compare with that of non-NHSEB students?

Over the summer of 2022, we began coordinating with teachers, coaches, and organizers of NHSEB teams across the United States. We invited students to complete a baseline survey at the start of the NHSEB season (mid-September 2022) and a follow up survey at the end of the season (mid-February 2023). Based on the results of the pilot test, the survey included: the Intellectual Humility Scale; two versions of the Cognitive Reflection Test; an Affective Polarization “thermometer”; and an Over-Claiming Questionnaire. Additionally, we decided to add two new measures. First, as discussed above, we included a brief version of the Social Desirability Bias Scale (Fischer & Fick, 1993), which assesses the degree to which respondents are inclined to

answer questions dishonestly in order to create a favorable impression. Second, we added the Prosocial Behavioral Intentions Scale (Baumsteiger & Siegel, 2019), which asks respondents how likely they would be to help others in need (e.g., “Help a stranger find something they lost, like their key or a pet”).

Since this is not a *randomized*, controlled trial, it cannot rule out a “self selection” effect, whereby students who join NHSEB are (for other, unknown reasons) more inclined to develop intellectual and civic virtue than students who don’t join NHSEB. However, this study will enable us to test whether students who join NHSEB differ, at the time of their joining, from their non-NHSEB peers. It will also enable us to test whether, over time, NHSEB students cultivate these virtues to a greater degree than their non-NHSEB peers.

Prospects for Future Research

In this section we propose several possible avenues for future research on the impact of Ethics Bowl. We begin by suggesting ways to build on the current efforts by using alternative outcomes, by employing non-survey based methods, and by taking new approaches to the study design. Our reflections on prospects for future research bring us to a second kind of rationale for outreach programs such as NHSEB. This *moral* rationale overlaps with but is distinct from the civic rationale we have so far focused on. We then offer suggestions for how one might assess the program’s ability to promote these moral ends.

Alternative Outcome Measures

I. Specific Intellectual Humility

Empirical research (Hoyle et al. 2016) has found that people can be intellectually humble in one domain (e.g., religion) or with respect to certain questions (e.g., “Does God exist?”) but intellectually arrogant in other domains (e.g., science) or with respect to other questions. Hence, in future studies researchers could assess different forms of intellectual humility. For Ethics Bowl students, in addition to intellectual humility as a general trait, researchers could assess humility

about the kinds of issues raised in the case studies used in the competitions. This would enable, for instance, an examination of the way in which students' overall or general level of intellectual humility relates to their willingness to revise their beliefs about specific moral and political issues such as police reform or lying.

II. Civic Engagement

Ethics Bowl might positively impact civic engagement, within school communities or beyond. This could be measured by looking at patterns and levels of volunteering or involvement in certain school activities. One could also look at the negative side and compare levels of disciplinary referrals, comparing students who participate in Ethics Bowl with those who don't, or comparing disciplinary referrals before and after students join. Under this heading one might also assess civic indicators, such as political self-efficacy, trust in politicians and political institutions, and interest in politics.

III. Behavioral Measures

Another way of building on our existing approach would be to capture students' behavior at organized events, such as Regional Competitions or National Championship. For instance, researchers might record audio or video of the matches and then have trained observers score the teams for behavioral markers of intellectual humility or attitudes towards the other team (see Topping and Trickey 2007c for an example of a P4C study using scored videos). Alternatively, automated natural language processing techniques could be used with the transcripts from matches, or even with students' written assignments from their courses in school.

Alternative Study Design

I. Systematic Comparison of Extracurricular Programs

Ethics Bowl's focus on collaborative deliberation about ethical dilemmas is unique among issue-related extracurricular activities, such as speech and debate. Programs such as Model UN, Academic Decathlon, or the National Speech and Debate Tournament likely promote some of the

same traits as Ethics Bowl and some distinct traits. What, then, is Ethics Bowl's characteristic or unique impact? Some have answered this in a more negative way, suggesting that adversarial debate programs have a vitiating effect on civil discourse (Ellis 2019). One can speculate about the traits each program promotes based on their norms and incentive structure, but empirical data is wanting. Future research could make systematic comparisons between the intellectual and civic traits developed by students who participate in NHSEB as opposed to cognate programs. This would also help to eliminate a potential confound of the design—namely that some students who participate in Ethics Bowl are also involved with these other programs.

II. Developmental Trajectory of Intellectual Virtue

There is a need for a more developmentally sensitive account of intellectual virtues such as humility, and the means to effectively promote them. Ethics Bowl is an exercise in reason giving between equal parties, and as such it presupposes certain developmental capacities (e.g. ability to articulate oneself or to engage with conceptually sophisticated cases) on the part of its participants. Further empirical work on Ethics Bowl, including with the burgeoning Middle School Ethics Bowl program, can help to specify the developmental threshold at which discussion-based activities of this kind are appropriate methods for promoting intellectual virtues like those we have set out to study.

III. Impact on Volunteers

It is important to remember that the impact of NHSEB extends beyond the students who participate. For example, NHSEB competitions across the country rely on volunteer judges from all walks of life. Judges plausibly enjoy some benefits from the experience, including from the training they undergo before matches and also the opportunity to engage in critical ethical reflection alongside the competing teams. We might expect judges' views on particular issues to evolve in light of the discussions they witness. We might also expect their assessment of the agency and competence of younger students to be more positive, or for their own levels of affective polarization to decrease. In many cases, NHSEB coaches are educators (teachers, staff,

counselors, and so on). One could investigate how participation in the activity in that capacity has enhanced their practice and enabled them to think about the normative dimensions of education more clearly.

IV. Impact on Teacher Perceptions of Students

Ethics Bowl as an activity often permeates other aspects of the school environment. Where there are NHSEB teams, for example, Ethics Bowl is often used as a pedagogical tool in the classroom, or even as a school-wide activity to build community and facilitate conversation on difficult subjects. If we focus exclusively on Ethics Bowl as an extracurricular activity, one could investigate the impact of this activity on how teacher perceptions of students who participate change over time, for example by having teachers keep a log of observed behavior, degree and quality of class discussion participation, or intellectual capacity.

A Moral Rationale for Outreach

So far we have focused on NHSEB's promotion of social and democratic aims. In addition to this broadly *civic* rationale, many outreach programs in philosophy also share the *moral* aim of creating spaces and opportunities for philosophical and ethical reflection. According to this second kind of rationale, critical reflection about ourselves and the world around us is an essential ingredient in a good and autonomous human life, and universities have a moral obligation to create spaces where that is possible for folks across the lifespan. The moral rationale includes instrumentally valuable ends such as academic achievement and intrinsically valuable ends such as a sense of meaning and purpose in life. Naturally, these two rationales are closely related. For instance, many of the civic and intellectual virtues discussed above are personally beneficial for those who possess them. Below we describe two possible outcomes of Ethics Bowl participation worth investigating that fall under this heading.

I. Academic Achievement

If Ethics Bowl promotes regimented thinking and clarity of expression, one might expect students to enjoy gains in academic achievement by participating. Empirical work on the Philosophy for Children (P4C) program has produced mixed evidence, with some studies finding that P4C programs improved children's cognitive abilities (Topping & Trickey 2007a; 2007b) as well as reading and math achievement (Gorard et al. 2015), and others finding no such effect (Lord et al. 2021). However, the impact of Ethics Bowl programs on students' academic achievement has not yet been studied, and is therefore one avenue for further investigation into the impact of philosophy programs on academic achievement. One promising method would be to examine Student Growth Percentiles (SGPs), which measure student growth in test scores relative to students who scored similarly on prior tests or "academic peers" (for a technical introduction to this statistical method, see Betebenner 2019). SGPs would indicate relative gains students make when Ethics Bowl is part of their schooling regimen. This form of impact assessment would not require students or coaches to make any effort beyond the ordinary course of annual standardized testing.

II. Personal Well-Being

Participating in Ethics Bowl might be conducive to students' well-being in a variety of ways. First, the kinds of activities that constitute participating in Ethics Bowl are themselves plausible components of a good human life. In Ethics Bowl, students reflect deeply and engage with questions of value and how we ought to live. It's a chance for students to clarify and develop their values, and to refine their conceptions of the good life. More concretely, many of the intellectual and civic virtues that we described above are themselves conducive to health and happiness. For instance, joining a collaborative group with a shared purpose is likely to promote a sense of community and belonging. The results of our pilot survey (see above) certainly support this speculation. Unsurprisingly, there is an enormous body of empirical research, going back decades, on the importance of such social integration for mental health (Baumeister & Leary, 1995). More surprisingly, perhaps, there is also a large body of evidence that it is a powerful

predictor of *physical* health as well (Holt-Lunstad et al. 2015). As another example, if participating in Ethics Bowl makes students more prosocial (as we are currently investigating), it should also foster happiness. Empirical research has found that one of the most effective ways of promoting one's own happiness is to focus on doing good for others (Aknin et al., 2013; Hui et al., 2020; Kushlev et al. 2021). Hence, in future studies, researchers might include measures of mental health and well-being—e.g., a sense of belonging, purpose, or meaning in life.

Conclusion

Social scientific studies have shown us time and again that we are liable to overestimation, closed mindedness, and egoistic reasoning. Preliminary data suggests that Ethics Bowl is an effective and developmentally appropriate intervention for cultivating positive intellectual and civic traits among students. A rigorous understanding of the extent of that impact, and ways pedagogical strategies and extracurricular activities can augment it, would be an invaluable contribution to civic education. NHSEB is a common starting point for university-based outreach programs because of the program's large footprint and supportive structure. It is also a program whose uniformity lends itself to empirical assessment. We hope philosophers, psychologists, and other scholars interested in intellectual and civic virtue will continue to collaborate in ways that enhance our understanding of the impact of Ethics Bowl. Anecdotal evidence and countless testimonials suggest the impact is large and beneficial, but turning those hunches into concrete data is the only way to communicate this message clearly and effectively to administrators and policymakers. In our view, such efforts can shed light on the impact of philosophy-based outreach programs more generally, especially those that utilize discussion practices animated by democratic norms.

Acknowledgments: The research activities described in this paper are made possible by the support of the Parr Center for Ethics and the National High Schools Ethics Bowl at UNC-Chapel Hill. We owe a special debt of gratitude to NHSEB Director Alex Richardson for his input, involvement, and willingness to take the unprecedented step of launching an impact study. We are grateful to Erik Kenyon for his suggestions about academic achievement, to Walter Sinnott-Armstrong for his insight and expertise on measuring intellectual virtues, to Sarah Stroud for her unstinting support, and to Dustin Webster for sharing his research on Philosophy for Children. Finally, we are grateful for the advice about study design offered by Paula McAvoy and the members of the Philosophy Learning and Teaching Organization Research Committee.

References

- Aknin, L. B., Barrington-Leigh, C. P., Dunn, E. W., Helliwell, J. F., Burns, J., Biswas-Diener, R., Kemeza, I., Nyende, P., Ashton-James, C. E., & Norton, M. I. (2013). Prosocial Spending and Well-Being: Cross-Cultural Evidence for a Psychological Universal. *Journal of Personality and Social Psychology*, 104(4), 635–652.
- Allen, D. (2016) What is Education For? (Boston Review). Available from: <https://bostonreview.net/forum/danielle-allen-what-is-education-for>
- Anderson, E.S. (1999) What Is the Point of Equality?. *Ethics*, 109(2), pp. 287–337.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529.
- Baumsteiger, R., & Siegel, J. T. (2019). Measuring Prosociality: The Development of a Prosocial Behavioral Intentions Scale. *Journal of Personality Assessment*, 101(3), 305–314.
- Betebenner, Damian W. (2011). A Technical Overview of the Student Growth Percentile Methodology: Student Growth Percentiles and Percentile Growth Projections/Trajectories. Dover, N.H.: National Center for the Improvement of Educational Assessment. Available from: https://ksde.org/Portals/0/Research%20and%20Evaluation/SGP_Technical_Overview.pdf
- Dutilh Novaes, C. (2020). *The Dialogical Roots of Deduction: Historical, Cognitive, and Philosophical Perspectives on Reasoning*. Cambridge University Press.
- Ellis, J. (2019) Are School Debate Competitions Bad for Our Political Discourse? *The New York Times*. Available from: <https://www.nytimes.com/2019/10/12/opinion/sunday/school-debate-politics.html?smid=url-share>
- Fischer, D. G., & Fick, C. (1993). Measuring Social Desirability: Short Forms of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement*, 53(2), 417–424.
- Fishkin, J.S. (2018) *Democracy when the people are thinking: Revitalizing our politics through public deliberation*. Oxford University Press.
- Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), 25–42.
- Gorard, S., Siddiqui, N., & Huat, B. (July 2015) *Philosophy for Children: Evaluation Report and Executive Summary*. Education Endowment Foundation. Available from: https://d2tic4wvo1iusb.cloudfront.net/documents/projects/EEF_Project_Report_PhilosophyForChildren.pdf?v=1630926028
- Gutmann, A. & Ben-Porath, S. (2014) *Democratic Education*. In *The Encyclopedia of Political Thought*, M.T. Gibbons (ed). Wiley Online Library.
- Gutmann, A. & Thompson, D. (2004) *Why Deliberative Democracy?*. Princeton University Press.
- Hess, D. & McAvoy, P. (2015) *The Political Classroom: Evidence and Ethics in Democratic Education*. New York: Routledge Press.

- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review. *Perspectives on Psychological Science*, 10(2), 227–237.
- Hoyle, R. H., Davisson, E. K., Diebels, K. J., & Leary, M. R. (2016). Holding specific views with humility: Conceptualization and measurement of specific intellectual humility. *Personality and Individual Differences*, 97(July), 165–172.
- Hui, B. P. H., Ng, J. C. K., Berzaghi, E., Cunningham-Amos, L. A., & Kogan, A. (2020). Rewards of kindness? A meta-analysis of the link between prosociality and well-being. *Psychological Bulletin*, 146(12), 1084.
- Israeloff, R., Mizell, K. and Wartenberg, T. eds. (2022). *The Ethics Bowl Way: Answering Questions, Questioning Answers, and Creating Ethical Communities*. Rowman & Littlefield.
- Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3), 690–707.
- Jared, S. (2021) *Philosophy as Public Service: Innovative Carolina programs are bringing philosophy for everyday use to prisons, retirement homes and high schools.* The Well, January 12. Available from: <https://thewell.unc.edu/2021/01/12/philosophy-as-public-service/>
- Kushlev, K., Radosic, N., & Diener, E. (2021). Subjective Well-Being and Prosociality Around the Globe: Happy People Give More of Their Time and Money to Others. *Social Psychological and Personality Science*, 19485506211043380.
- Laden, A.S. (2013) *Learning to be equal: Just schools as schools of justice*. In *Education, justice, and democracy* (pp. 62-79). University of Chicago Press.
- Ladenson, R.F. (2001) The educational significance of the ethics bowl. *Teaching Ethics*, 1(1), pp.63-78.
- Leary, M. R., Diebels, K. J., Davisson, E. K., Jongman-Sereno, K. P., Isherwood, J. C., Raimi, K. T., Deffler, S. A., & Hoyle, R. H. (2017) Cognitive and Interpersonal Features of Intellectual Humility. *Personality and Social Psychology Bulletin* 43(6), 793–813.
- Lee, L.M. (2021). The growth of Ethics Bowls: a pedagogical tool to develop moral reasoning in a complex world. *International Journal of Ethics Education*, 6(1), pp.141-148.
- Lipman, M. (1973). "Philosophy for Children," Report of 1970-71 Pilot Project, submitted to U.S. Department of Health, Education and Welfare, Institute of Education. Available from: <https://files.eric.ed.gov/fulltext/ED103296.pdf>
- Lord, P., Dirie, A., Kettlewell, K., and Styles, B. (2021) *Evaluation of philosophy for children: an effectiveness trial* Education Endowment Foundation. Available from: <https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/philosophy-for-children-effectiveness-trial>
- McAvoy, P. & Lowery, A. (2022) *Structure Matters: Comparing Deliberation and Debate*. In *Making Classroom Discussions Work*. (J. Lo, editor). Teachers College Press.

- McAvoy, P. & McAvoy, G.E. (2021) Can debate and deliberation reduce partisan divisions? Evidence from a study of high school students. *Peabody Journal of Education*.
- McAvoy, P. and Hess, D. (2013) Classroom Deliberation in an Era of Political Polarization. *Curriculum Inquiry*, 43, pp. 14-47.
- Millett, S., & Tapper, A. (2012). Benefits of collaborative philosophical inquiry in schools. *Educational philosophy and theory*, 44(5), 546-567.
- Moshman, D., & Geil, M. (1998). Collaborative Reasoning: Evidence for Collective Rationality. *Thinking & Reasoning*, 4(3), 231–248.
- National Constitution Center (2019) Conversations with Templeton Foundation Grantees: Jeffrey Rosen. Jul 17. Available from: <https://www.templeton.org/news/conversations-with-grantees-jeffrey-rosen>
- Neufeld, B. (2022) *Public Reason and Political Autonomy: Realizing the Ideal of a Civic People*. Routledge.
- Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, 84(4), 890–904.
- Porter, T., Baldwin, C.R., Warren, M.T., Murray, E.D., Cotton Bronk, K., Forgeard, M.J., Snow, N.E. and Jayawickreme, E. (2022) Clarifying the content of intellectual humility: A systematic review and integrative framework. *Journal of personality assessment*, 104(5), pp. 573-585.
- Rawls, J. (1997) The idea of public reason revisited. *The University of Chicago Law Review*, 64(3), 765-807.
- Reznitskaya, A. (2005) Empirical research in philosophy for children: Limitations and new directions. *Thinking: The Journal of Philosophy for Children*, 17(4), pp.4-13.
- Richardson, A.M. (2022) Meeting the Challenge: The Future of Ethics Bowl. In Israeloff, R., Mizell, K. and Wartenberg, T. eds. *The Ethics Bowl Way: Answering Questions, Questioning Answers, and Creating Ethical Communities*. Rowman & Littlefield.
- Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1), 15.
- Topping, K. J., & Trickey, S. (2007b). Collaborative philosophical inquiry for school children: cognitive gains at 2-year follow-up. *The British Journal of Educational Psychology*, 77(Pt 4), 787–796.
- Topping, K. J., & Trickey, S. (2007c). Impact of philosophical enquiry on school students' interactive behaviour. *Thinking Skills and Creativity*, 2(2), 73-84.
- Topping, K.J. & Trickey, S. (2015). Philosophy for children: Short and long term effects. In *The Routledge International Handbook of Research on Teaching Thinking*. Routledge, pp. 103-112.

- Topping, K.J. and Trickey, S. (2007a) Collaborative philosophical enquiry for school children: Cognitive effects at 10–12 years. *British Journal of Educational Psychology*, 77, pp. 271-288.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, 143(5), 1958–1971.
- Vazquez, M. (2022) Ethics Bowl and Democratic Deliberation. Blog of the American Philosophical Association. March 23. Available from: <https://blog.apaonline.org/2022/03/23/syllabus-showcase-ethics-bowl-and-democratic-deliberation-michael-vazquez>
- Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, 330(6004), 686–688.