

Beyond the Trifold in Civics Presentations: The Measure of Youth Policy Arguments

Carlos P. Hipolito-Delgado

University of Colorado Denver carlos.hipolito@ucdenver.edu

Dane Stickney

University of Colorado Denver dane.stickney@ucdenver.edu

Ben Kirshner

University of Colorado Boulder ben.kirshner@colorado.edu

Andrew Maul

University of California Santa Barbara amaul@ucsb.edu

Abstract

Youth are increasingly engaging in civic action to address social injustices. Many adult educators are also looking for instructional resources that support youth voice as a way to promote adolescent civic development and community change. Alas, assessment tools to support youth voice and policy argumentation are lacking. Existing tools overemphasize public speaking skills and rely on dated artifacts such as cardboard trifold posters. In this article we introduce the Measure of Youth Policy Arguments (MYPA), a tool designed to aid in the development and assessment of high-quality youth policy presentations. We also describe how to use the MYPA in formative and summative contexts. Additionally, we provide initial evidence for the validity and reliability of the MYPA. Furthermore, we argue that MYPA has applications in preparing youth for policy presentation and in assessing learning outcomes associated with youth voice projects.

Key words: youth policy arguments, student voice, action civics, youth participatory action research

On a mid-May morning, youth from 15 high schools near a major Midwestern U.S. city presented their findings after months of researching problems at their schools. At the culminating event of the yearlong academy aimed at bolstering culturally responsive pedagogy

New articles in this journal are licensed under a Creative Commons Attribution 4.0 License. This journal is published by the University Library System, University of Pittsburgh and is cosponsored by the University of Pittsburgh Press. The Journal of Youth Development is the official peer-reviewed publication of the National Association of Extension 4-H Youth Development Professionals and the National AfterSchool Association.

through youth participatory action research (YPAR), the youth presented by speaking into a microphone and pointing at charts and quotes pasted onto a cardboard trifold presentation poster board. Topics included the effects of substance abuse on youth, schools' inability to address youth's mental health concerns, and low rates of youth motivation in classes. Superintendents, principals, teachers, and university professors applauded the youth for their work. Although the youth did not receive many questions, they were rewarded with handshakes, hugs, and backslaps.

As the event concluded, program organizers asked the youth to share any feedback about the event. After several youth shared appreciations, one of the youth questioned why, in the 21st century, were they asked to present using trifolds? Based in part on this comment, we found ourselves asking how the youth could have been better prepared or provided with critical feedback that would aid them in addressing the complicated topics they explored. Though this May event provided youth a space to share their concerns, we felt that youth required additional support in constructing and delivering policy arguments and an authentic stage where they could exercise their voice.

Though youth civic development efforts are gaining attention, there is a dearth of resources that guide the construction, delivery, and evaluation of policy arguments to real-world audiences (Pope et al., 2011). In many cases, youth invest hours of work throughout an academic year preparing for a final public delivery. Youth deserve real support in crafting policy arguments, including constructive, tangible assessment and feedback after presenting.

To this end, we describe how the Measure of Youth Policy Arguments (MYPA; Kirshner et al., 2020) can be used to support and assess youth in publicly presenting policy proposals. The MYPA assesses basic presentation skills (hooking the audience's attention and clarity of speech) as well as higher order skills, such as identifying the root causes of a focal problem, using appropriate research methods to gather data, and creating a policy solution that addresses said problem. The MYPA avoids the traps of traditional presentation rubrics by focusing on quality of argument, rather than the presentation style, thus providing a form for youth to receive critical and substantive feedback.

Our primary goal in developing the MYPA was for its formative potential, so youth would have a tool to aid them in crafting powerful civic presentations. However, the aim of this article is to provide initial evidence for the MYPA's use as both an authentic formative and summative

assessment tool. In our own experience, the summative potential of MYPA has opened opportunities for partnerships with in-school and out-of-school organizations, ultimately leading to the MYPA being used as both a formative and summative tool in these organizations.

In this article we describe the rise of youth voice and the need for authentic formative and summative civics presentation tools. We then discuss the development, intended uses, and validation of the MYPA. Finally, we share potential applications of the MYPA in preparing youth for policy presentation and in assessing learning outcomes associated with youth voice projects. This article advances the youth development literature base by describing a tool that will aid youth and their adult partners in preparing for and authentically assessing the quality of youth policy presentations.

The Rise of Youth Voice

Internationally and nationally, youth activists have organized large-scale demonstrations, including marches and speeches. Greta Thunberg, perhaps the most visible youth and climate activist, was named Time magazine's youngest ever person of the year in 2019 (Felsenthal, 2019). Thunberg's address to world leaders at a climate action summit in New York in September 2019, was a master class in impassioned speech. Nationally, youth from Parkland, Florida organized the March for Our Lives in response to a shooting at their high school and lax national gun laws (Witt, 2018). The Parkland youth have been innovators: Emma Gonzalez's moment of silence at the March 24, 2018, March for Our Lives is a powerful example of this (NBC News, 2018). More locally, youth from Denver took to social media to critique district leadership and garner support for striking teachers. Though the Denver youth faced backlash on social media platforms and from district personnel, they succeeded in mobilizing their peers and increasing public awareness. None of these youth were confined to the typical tools of civics courses—a trifold, prompt, and 3 minutes to present—and they seized the opportunity. These forms of voice and activism did not emerge from education spaces, but if we embrace the idea that a strong and just democracy requires education, then we want to think about the kinds of curriculum and assessments that support youth's civic development.

With regard to school-based youth agency and activism, the most promising approach is action civics. In action civics, youth "do civics and behave as citizens by engaging in a cycle of research, action, and reflection about problems they care about personally while learning about deeper principles of effective . . . political action" (p. 224, Levinson, 2012). Action civics has

roots in the educational theories of John Dewey but has recently received renewed attention as educators have sought approaches to civic education that bring democratic practice to life through real-world project-based learning, discussion of controversial issues, and sharing of policy ideas with public audiences (Blevins et al., 2016; Gould et al., 2011). Evidencing the impact of action civics curriculum, Andolina and Conklin (2018) found that youth who partook in a weeklong action civics curriculum, where they chose topics relevant to their daily lives and interacted with peers and community members to share their stories, demonstrated higher levels of civic and rhetorical skill. Cabrera et al. (2014), using a logistic regression analysis, found a significant direct relationship between youth participation in an ethnic-studies action civics program and graduation rates and standardized test scores.

The out-of-school time (OST) field, too, has seen a resurgence of programs that facilitate youths' opportunities to critically reflect on their social world and develop projects that make political demands of elected leaders. OST programs use different models, including youth organizing, YPAR, and youth-adult partnerships (Torre & Fine, 2006; Wu et al., 2016). Many OST programs share with action civics an emphasis on opportunities for young people to speak up about issues they care about but vary in the extent to which they prioritize research, critical social analysis, or political action (Akiva & Petrokubi, 2016; Conner & Rosen, 2016). Despite this renewed interest in civics programing, a challenge remains in how to assess youth learning, especially as related to civics policy presentations.

The Challenge of Assessing Youth Policy Presentations

Although located in different institutional contexts, with different pressures related to accountability, schools and OST programs both rely on assessment systems to evaluate youth learning, improve programs, and report to external stakeholders. In this sense, youth voice programs share a common interest in the availability of tools for assessing youths' capacity to construct and deliver high-quality policy arguments. Alas, existing resources are inadequate.

In the school context, assessments of action civics are limited in two ways. The first pertains to a mismatch between typical federal- and state-mandated assessment systems, which are standardized and based on bodies of pre-established knowledge (codified, for example, in the Common Core State Standards), and the kinds of open-ended inquiry characteristic of action civics. Action civics projects are not designed to teach information that can be found in textbooks (Levinson, 2012); correspondingly, tests based on textbooks will not be sensitive to

the kinds of learning youth can achieve. Second are the limitations of existing rubrics for summative and formative assessment of action civics. In a review, conducted by the authors, of 13 youth presentation rubrics from civics organizations and school districts across the country, seven closely resemble public-speaking or debate rubrics, including items to assess voice projection and posture. One instrument dedicated 18 of 26 questions to the design of presentation slides. Those rubrics that did assess the content of the presentation tended to utilize checklists ("was it there or not?"), with little or no reflection of differences in quality. Only four of the rubrics provided a Likert-type scale to assess quality, alas none included a description for individual score categories—leading to a lack of differentiation of what makes a score a 3 versus a 2 or 4. None of these tools provided information on their reliability and validity, leaving in question whether these rubrics can produce accurate and consistent scores—also posing a limitation in their use in research and assessment of youth learning. Given the limitations in existing rubrics, it appears that youth are not receiving constructive feedback on their policy presentations and educators are limited in their ability to assess the quality of youth learning.

With regard to OST programs, although they have more flexibility about how they assess youth learning, they often face pressure to provide outcome data. One relatively simple and inexpensive way to do this is by administering self-report survey instruments designed to measure civic development constructs, such as civic agency or intentions to vote in the future (Flanagan et al. 2007). However, self-report surveys have a number of known limitations that make them less than ideal for the purpose of estimating youth learning (see Howard, 1980). Such instruments are subject to a range of construct-irrelevant influences, the most serious of which are social desirability bias (Krumpal, 2013) and idiosyncrasies in how individuals respond to survey items in general (see Messick, 1991). As a practical matter, scores on these instruments often have a high floor at pretest, limiting their capacity to detect change. Further compounding matters, traditional methods of validation of self-report instruments may be largely incapable of detecting serious flaws in their design (Maul, 2017). Although assessment should never dictate the boundaries or imagination of youth voice projects, it can be helpful in giving feedback to youth or reporting learning to educators and stakeholders.

For youth voice programming the concept of "authentic assessment" offers a useful framework as it mirrors mature civics practices. According to Wiggins (1990), authentic assessment aims to evaluate youth performances or products as they "rehearse the complex ambiguities of the 'qame' of adult and professional life" (p. 1). What is more, when youth master a real-world skill

related to an issue they care about, Levinson (2012) argued, "they are much more likely to retain what they have learned and apply it to other contexts than when they merely try to memorize facts or the steps of a skill for a school-based test" (p. 69). Mastery is key, but, in the case of youth voice presentations, it is complicated to assess. We developed the MYPA to be an authentic assessment of youths' ability to perform policy arguments; the MYPA provides a more complete assessment picture by examining youths' problem identification, research, policy development, and presentation abilities across a range of topics (Kirshner et al., 2020).

The Measure of Youth Policy Arguments

The MYPA is an observation protocol designed to assess the quality of youth policy presentations to external stakeholders, which often are the culminating activity for action civics and YPAR projects. We were motivated to develop the MYPA in part from our review of youth policy presentations, where we found that youth presentations did not address the root causes of their focal problem, employed questionable or misaligned research methods, proposed incomplete or off-base policy solutions, and did not take the opportunity to enlist the help of their audience in enacting social change. Additionally, the MYPA has potential applications as a summative assessment for adults looking to evaluate youth presentations and programs hoping to assess learning outcomes.

Intended Uses

The MYPA is intended to be used by educators in youth voice programs in and out of schools. The tool was designed to be used primarily for formative purposes, in order to help middle and high school youth prepare high-quality policy arguments. It can be used as a guide to help youth construct a policy presentation or to provide youth with feedback as they practice said presentation. As a summative assessment tool, the MYPA might be used to examine youth learning outcomes either at the classroom or program level. At the classroom level, the MYPA can be used for grading group policy arguments. At the program level, the MYPA might be used to assess youths' critical thinking skills or ability to engage in civic action; this may be useful for programs looking to provide data on youths' learning for administrators or funders. Finally, the MYPA might be used for scoring youth voice and action civics competitions, where judges can use it to compare the quality of teams' presentations.

There are certain conditions where the use of the MYPA is not recommended: It was not designed for use in high-stakes testing. We argue that the goals of high-stakes testing are incompatible with the MYPA. The MYPA is an authentic assessment of youths' ability to engage in policy argumentation, not a generalizable assessment of knowledge.

Moreover, we caution potential users against the schoolification of the MYPA--that is, by treating it as a checklist for a grade rather than to support quality social action (Rubin, et al., 2017). Classrooms often focus on products like tests, essays, and speeches; in these contexts, youth voice can become an exercise instead of a tool to enact change. Recently, for example, we observed a civics education day where young folks used a trifold to identify issues, root causes, and tactics for change in one of four assigned topic areas. Though this structure provided a helpful one-stop visual that made it easy to locate information across presentations, this generic product was devoid of passion and persuasion. Any assessment tool, the MYPA included, risks becoming empty of meaning when schoolified; our aim is for the MYPA to be a tool that supports youth passion and engagement in leadership and social change.

Scoring

The MYPA aims to assess the quality of youth presentations organized according to six constructs: presentation and delivery, problem identification, research methods, policy proposal, collaboration, and responsiveness to questions. The MYPA consists of 25 total items. Two openended items ask the rater to identify the focal problem and proposed policy discussed in the youths' presentation; these are intended to guide the rater in evaluating other elements of the presentation and are not scored. The other 23 items require the evaluator to rate presentation quality on specific criteria (e.g., relevance to speakers or extent to which evidence is convincing), and have two, three or four scoring categories (depending on the item) ranging from lowest to highest quality.

When used for summative purposes, the scores on individual items can be weighted to prioritize higher-order constructs, such as critical analysis and quality of response to questions. These items require youth to demonstrate critical thinking skills. This is particularly true of responses to questions, where youth cannot rely on a script, but must demonstrate their knowledge and potentially defend against counter arguments. A copy of the MYPA can be found at https://transformativestudentvoice.net/curriculum/.

It should be noted that there is a minimal training required of raters using the MYPA. The MYPA scorecard is organized so that raters can work through categories of quality from left to right. This allows the rater to understand all that should be in a presentation for it to receive the strongest possible score. Additionally, the constructs are ordered in the manner that they are most likely to appear in a presentation. Raters are also asked to attend to the language of the rating categories: in particular, youth should receive ratings based on what was actually demonstrated and verbalized, not what could be inferred. Furthermore, a rater should be attentive to their biases—in piloting we found that raters often scored youth higher on topics that spoke to the rater's interests. As such, raters should be attentive to the presentation itself and provide a score based on what youth discuss. What is more, we encourage raters to be critical of youth presentations. We have found that youth appreciate receiving critical feedback as they are looking to improve their argumentation skills. Finally, we encourage adaptations of the MYPA based on the reader's specific needs. The reader might consider alternative scoring structures or the addition of items that speak to specific youth-learning outcomes. For example, we developed an alternate version of the MYPA with more items related to youths' responses to questions that might be appropriate for situations in which more time is provided for Q&A with adult decision makers or judges.

Development of the MYPA

In the development of the MYPA we drew on construct-centered assessment (Wilson, 2005) and research-practice partnerships (RPP; Coburn & Penuel, 2016). We began with the specification of constructs, each with different levels of quality, represented in construct maps. The initial version of constructs were informed by a thorough review of scholarly literature on policy arguments and a review of youth policy presentations. For this latter process we invited seven research team members, eight local educators, two community organizers, and three local politicians to view videos of youth policy arguments in order to identify potential constructs. We also utilized the principles of co-design (Penuel et al., 2007) with experienced educators and frontline users and youth, with iterative cycles of development, piloting, and feedback. Feedback during this time led to several iterations of changes to item language and efforts to shorten the protocol to make it more user friendly. Changes were also made to clarify the gradations of quality in calls to action and how youth drew on evidence. For more detail on how the MYPA was developed in the context of RPP see Kirshner et al. (2020).

Evidence of Reliability

A first piece of evidence in support of the validity argument (see Bell et al., 2012) for the adequacy and appropriateness of the MYPA for its intended uses is that trained raters exhibit sufficient convergence in their ratings so as to be confident that at least the significant majority of variance in scores is not due to the idiosyncrasies of the individual rater—that is, that interrater reliability (IRR) is sufficiently high. To help establish this, researchers attended and scored youth policy presentations at two action civics events. The pseudonyms of the Western Contest and the Midwestern Capstone will be used to maintain the anonymity of these events and their organizers. The Western Contest is organized by the Student Civic Leadership (SCL) program that is housed within the school district of an urban city in the western United States. Eighteen high school teams participated in the Contest, which was held in May and reflects work completed throughout the academic year. During the Contest high-school-aged youth describe a community or local issue, research they undertook to better understand the issue, and a proposed policy meant to address the problem. Each team received 5 minutes to present and 3 minutes to answer questions from a panel of judges. At this event two raters scored four youth presentations (these ratings were distinct from the scores of the judges at this event).

The Midwestern Capstone is a year-end showcase of YPAR. The Capstone is organized by a collaborative that includes a state university, a local education agency, and multiple local school districts of an urban Midwestern city. A total of 15 high schools participated in this collaborative and 16 teams (one school had two youth teams) participated in the Capstone. During the Capstone youth had 5 minutes to share the results of their YPAR projects with school and community leaders. Our involvement with the Capstone was as a result of their project manager reaching out to us with an interest in using the MYPA to assess the quality of their youth's presentations and gain insights on learning outcomes. We came to a mutual agreement that allowed our researchers to attend the event, so we might evaluate IRR using the MYPA, and provide the collaborative with feedback on their youths' presentations. Two raters scored all 16 presentations. This study was approved by the institutional review board of the home institutions of the first, second, and third authors.

Raters

Two raters viewed and scored all presentations associated with this study. Both raters were doctoral students and graduate research assistants associated with this research project. Rater 1 is a White male and a former teacher with experience leading action civics projects. Rater 2 is an African American female, a former school administrator, and executive director of a nonprofit

youth organization for Black girls. Both raters participated in norming activities with the first author, though at separate points in time. As part of norming, the first author and the raters discussed the constructs associated with the MYPA and viewed multiple youth policy presentation videos from a collection of prior youth voice events. First, they would view a video and collectively score on the MYPA, allowing time to discuss why certain ratings were most appropriate. Eventually, they would watch and score videos separately and later come together to compare ratings and discuss disparate scores. Norming continued until the first author and the rater could achieve an overall 80% agreement on the MYPA. These norming activities took approximately 3 hours. The two raters did not participate in norming activities with each other.

Procedure

The raters attended both the Western Contest and Midwestern Capstone, at which they watched presentations live and scored in the moment. This required the raters to score during the presentation and to use the limited transition times to complete their assessment, which approximates typical conditions for raters at a youth voice event. Raters scored presentations independently using an online version of the MYPA. Given the rapid pace of the environment and the fact that scores were stored online, the raters did not debrief or otherwise discuss their scores during these events. Microsoft Excel was used to assess the percentage of agreement between raters.

Results

Percentage agreement was used as a measure of IRR, as it provided a method to assess by item, construct, and the MYPA overall. Percent agreement examines the number of instances upon which both raters provided identical ratings divided by the total number of ratings (McHugh, 2012). In the case of the MYPA, the overall percent agreement for a single observation would be the number of times the raters agreed on their assessment divided by 23—for the 23 scored items on the MYPA. At the Western Contest, IRR ranged from 73.91% and 78.26%, with an average percent agreement of 75% across all MYPA items and all four presentations. For the Midwestern Capstone, IRR ranged from 69.57% and 95.65% with an average agreement of 83.24% across all MYPA items and all 16 presentations. Across all 20 presentations the raters achieved an average IRR of 80.87%.

Across all presentations, raters had the highest percentage of agreement on the constructs of Response to Questions (100%), Collaboration (86.67%), and Presentations and Delivery (84.17%). The MYPA constructs where raters had the lowest percentage of agreement were

Research Methods (78.75%), Policy Proposal (76%) and Problem identification (75%). Lower IRR on these three constructs might be related to their associated items having three or four categories of quality, requiring greater attention to differentiation, and these constructs containing two of the three items that are more difficult to score. In terms of individual items, the raters had perfect agreement (100%) on items 2, 22, and 25. The raters had a percentage of agreement of 80% or higher on 15 items and a percentage of agreement of 90% on 10 items. Raters had the lowest percentage of agreement on items 11 (60%), 4 (55%), and 15 (50%).

Evidence for Use

In order to argue for the validity of the MYPA for its intended purposes, we draw on data from user feedback, program implementation, and a mini-case study of a young person trained in policy argumentation using the MYPA. Our hope in providing this preliminary evidence is to show that the MYPA produces appropriate results in authentic action civics and YPAR contexts.

Raters

As part of the judging process for the Western Contest a group of eight judges used the MYPA to assess the quality of youth policy presentations. Prior to the contest, the judges received a 2-hour training on how to use the MYPA. These judges were educators, community activists, and elected officials. At the conclusion of the contest, we solicited feedback from the judges, specifically asking if their scores on the MYPA aligned with their perception of presentation quality. As one team member facilitated the conversation, a second team member took notes of the judges' comments. Generally, the judges expressed that the MYPA accurately captured their assessment of a presentation's quality and that the MYPA items represented elements of high-quality policy presentations.

Programs

Following their experience with the MYPA, both programs described previously adopted the MYPA for formative and summative purposes. The SCL adopted the MYPA for scoring youth presentations in the Contest and redeveloped their curriculum to align with MYPA constructs, in a form of backwards mapping. Similarly, the collaborative that organizes the Midwestern Showcase adopted the MYPA, but for the purposes of assessing youth learning outcomes. The collective used the MYPA to assess if school teams reached desired competencies related to YPAR and argumentation. The collective used these results to redesign their curriculum and

attend to areas where teams scored lower on the MYPA. Based on this initial evidence, we contend that MYPA might be appropriate both for assessing the quality of youth presentations and for assessing learning outcomes associated with action civics and YPAR programs.

Mini-Case Study

In addition to evidence that judges felt scores on the MYPA aligned with the quality of youth presentations and that programs adopted the MYPA to assess youth learning, we share a third type of evidence from our ongoing research with SCL. Specifically, this case study presents how youth versed in the MYPA applied their learning to an authentic issue unrelated to any school class or program: defunding the police in their school district.

The political context of the Black Lives Matter movement during the Spring of 2020 provided a stage for youth to demonstrate their civic analysis and action skills, especially in policy argumentation. One young person in particular, Nina (who was a youth leader in SCL and obtained training using the MYPA for policy argumentation), voluntarily took on a leadership role in advocating for the removal of police officers from local schools. In the late spring of 2020, on the steps of a large local high school, Nina delivered a powerful speech demanding that the school district end its contract with the police department. Her delivery was closely aligned with the MYPA, as she shared a clear message of "counselors not cops" and supported her argument with powerful evidence, emotional lived experience, and shrewd critiques of unjust social systems. She ended with a clear call to action for the school district and the city as a whole: "Defund the police and finally start to put actions behind your words. Students of color deserve more than broken promises, handcuffs, and juvenile records." Though Nina entered the SCL program with high levels of interest in issues of equity and justice, her ability to organize an argument and her belief in herself to publicly deliver it grew strongly through her experiences with the MYPA.

Implications

Based on these preliminary findings, we submit that the MYPA advances the youth development field by supporting the planning, delivery, and assessment of authentic and dynamic youth voice presentations. What is more, the MYPA foregrounds higher-order constructs not present in existing presentation rubrics, better equipping youth to develop high quality policy presentations grounded in logic, personal testimony, and research in ways that engage and persuade adult decisionmakers.

We have seen that action civics showcases often do not meet the needs of youth for an authentic 21st century learning experience. The trifold—a staid, static, and schoolified product—is a tangible artifact of this disconnect. But the trifold is just the beginning. Our analysis of existing rubrics spotlighted tools that lead to surface-level feedback for presenters. Alas, poor rubrics, as teachers know, make for poor roadmaps for success. Moving beyond the trifold requires breaking new ground. The MYPA does that by including constructs like problem identification, research methods, policy proposal, and response to audience questions. Further, the MYPA includes items that require critical thinking and analytic skills, such as using a value frame that resonates with their audience and identifying a clear timeline and responsibilities for implementing the proposed policy. Thus, the MYPA can be an asset for youth in their current policy advocacy; additionally, it aids youth in building knowledge and skills useful for future instances of legislative advocacy.

Supporting Youth Presenters

The primary goal of the MYPA is to provide youth with formative feedback. In studying action civics presentations across the country, we found that the feedback youth typically receive about their policy presentations tended to be surface level. This type of feedback seems like a slight when youth worked on their projects for the better part of a year and are addressing issues with implications for their lives. Criticism-free celebrations verge on tokenization, in which the words of youth aren't fully valued on the same level as adults (Conner et al., 2016). By using the MYPA, youth voice educators can provide youth with meaningful feedback.

To this end, the MYPA can be used by adult mentors to provide youth with feedback throughout the presentation development process. First, as youth prepare to develop their presentation, the sections of the MYPA can help them understand what constitute elements of high-quality presentations, such as the importance of hooking the audience early and being sure to have a call to action for the audience. The MYPA can then be used to provide feedback on trial presentations. This type of feedback is vital as it inspires youth to produce their best work.

Adult Allies

Educators and community activists often lack resources to help youth prepare for policy presentations (Kirshner et al., 2020). Though the MYPA was not designed to be a curricular tool, the structure of the MYPA can help educators in planning youth voice curriculum. For

example, educators might use the MYPA in backwards mapping, as was done by the SCL program. If educators begin by framing the constructs of the MYPA as desired end competencies, then one could backwards map the knowledge and skills youth need to reach said competencies. This then provides some direction concerning the types of lessons that need to be designed.

As a summative tool, the MYPA can be used to assess youth learning, as was done by the Midwest Collective. In an educational landscape that is assessment oriented, the ability to accurately measure youth learning is paramount (Levinson, 2011). Though the MYPA was not created for high-stakes testing purposes, it can help to document youth learning and identify areas of strength and weakness.

Furthermore, we believe that the MYPA constructs (problem identification, research methods, and policy proposal), and items within, align directly with the action civics and YPAR cycle. As such, the MYPA is well suited to assessing youth learning in these types of programs. What is more, the MYPA might provide a more authentic assessment of learning associated with action civics and YPAR as it can measure youths' ability to engage in civics activities such as preparing and giving policy presentations to decision makers.

Limitation and Future Research

Although the preliminary evidence from this study indicates that the MYPA is appropriate for assessing the quality of youth policy arguments, additional research is required to further establish validity. Though raters demonstrated good IRR on most of the MYPA constructs and items, the affiliation of the raters to this research team is a limitation. Future evaluations of IRR should be conducted with a group of raters external to this research team.

Additionally, the sample size of both raters and presentations was low and raters had challenges with IRR on certain items. This might be a function of these items being more subjective in nature, but they nonetheless pose a potential limitation to the use of the MYPA. To this end, additional research is needed to assess the language and scoring of these items, particularly 4, 11, and 15, to ensure they can be scored consistently from one rater to the next.

Another potential limitation to our argument for the use of the MYPA is that both programs with which we partnered used youth voice curriculum based in either action civics or YPAR. The

usability of the MYPA might be linked to scoring presentations associated with these types of programs. As such, research is needed to evaluate the validity and reliability of the MYPA with youth voice presentations not based in action civics or YPAR. Additional research is needed to assess the validity and reliability of the MYPA in other educational and community settings.

Conclusion

As youth claim their voice in public spheres to argue for issues impacting their lives, tools are needed to aid youth in preparing for and to document learning outcomes associated with policy arguments. We argue that the MYPA can help youth prepare for policy arguments and for adults to provide youth with critical feedback. Though additional research is needed, we hope this article inspires further discussion on how educators can foster, support, and assess authentic youth civic development.

Author Note

This work was supported by the William and Flora Hewlett Foundation under Grant 3248; This work was supported by the Spencer Foundation under Grant 201600010.

Correspondence concerning this article should be addressed to Carlos P. Hipolito-Delgado, University of Colorado Denver, Campus Box 106, P.O. Box 173364, Denver, CO 80217-3364. Email: carlos.hipolito@ucdenver.edu.

References

- Akiva, T., & Petrokubi, J. (2016). Growing with youth: A lifewide and lifelong perspective on youth–adult partnership in youth programs. *Children and Youth Services Review, 69*, 248-258.
- Andolina, M. W., & Conklin, H. G. (2018). Speaking with confidence and listening with empathy: The impact of *Project Soapbox* on high school students. *Theory & Research in Social Education,* 46(3), 374-409. https://doi.org/10.1080/00933104.2018.1435324
- Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, 17(2-3), 62-87.
- Blevins, B., LeCompte, K., & Wells, S. (2016). Innovations in civic education: Developing civic agency through action civics. *Theory and Research in Social Education, 44*(3), 344-384. https://doi.org/10.1080/00933104.2016.1203853

- Cabrera, N. L., Milem, J. F., Jaquette, O., & Marx, R. W. (2014). Missing the (student achievement) forest for all the (political) trees: Empiricism and the Mexican American studies controversy in Tucson.

 *American Educational Research Journal, 51(6), 1084-1118.

 https://doi.org/10.3102/0002831214553705
- Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, *45*(1), 48-54. https://doi.org/10.3102/0013189X16631750
- Conner, J. O., & Rosen, S. (2016). *Contemporary youth activism: Advancing social justice in the United States.* Praeger.
- Conner, J. O., Ober, C. N., & Brown, A. S. (2016). The politics of paternalism: Adult and youth perspectives on youth voice in public policy. *Teacher College Record*, *118*(8), 1-48.
- Felsenthal, E. (2019, December). *Time's editor in chief on why Greta Thunberg is the person of the year*. Time Magazine. https://time.com/person-of-the-year-2019-greta-thunberg-choice/
- Flanagan, C., Syvertsen, A., & Stout, M. (2007). Civic measurement models: Tapping adolescents' civic engagement. CIRCLE Working Paper 55. Center for Information and Research on Civic Learning and Engagement. https://files.eric.ed.gov/fulltext/ED497602.pdf
- Gould, J., Jamieson, K. H., Levine, P., McConnell, T., Smith, D. B. (Eds.). (2011). Guardian of democracy:

 The civic mission of schools. Leonore Annenberg Institute for Civics of the Annenberg Public

 Policy Center at the University of Pennsylvania.

 https://www.socialstudies.org/newsandadvocacy/newreportguardianofdemocracythecivicmissionofschools
- Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review*, *4*(1), 93-106.
- Kirshner, B., Zion, S., DiGiacomo, D., & Logan, G. (2020). The measure of youth policy arguments: An approach to supporting democratic participation and student voice. *Democracy and Education,* 28(2), Article 4. https://democracyeducationjournal.org/home/vol28/iss2/4
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47*(4), 2025-2047.
- Levinson, M. (2011). Democracy, accountability, and education. *Theory and Research in Education, 9*(2), 125-144. https://doi.org/10.1177/1477878511409622
- Levinson, M. (2012). No citizen left behind. Harvard University Press.
- Maul, A. (2017). Rethinking traditional methods of survey validation. *Measurement: Interdisciplinary Research and Perspectives, 15,* 51-69.
- Messick, S. (1991). Psychology and methodology of response styles, in R. E. Snow and D. E. Wiley (Eds.) *Improving inquiry in social science,* (pp. 161-200). Erlbaum.

- McHugh M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276-282. https://doi.org/10.11613/BM.2012.031
- NBC News. (2018, March 24). *Emma Gonzalez leads emotional moment of silence during rally* [Video file]. https://www.youtube.com/watch?v=4llJIChK2Do
- Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning*, 2(1), 51-74. https://doi.org/10.1142/S1793206807000300
- Pope, A., Stolte, L., & Cohen, A. K. (2011). Closing the civic engagement gap: The potential of action civics. *Social Education*, *75*(5), 262-268.
- Rubin, B. C., Ayala, J., & Zaal, M. (2017). Authenticity, aims and authority: Navigating youth participatory action research in the classroom. *Curriculum Inquiry, 47*(2), 175-194. https://doi.org/10.1080/03626784.2017.1298967
- Torre, M. E., & Fine, M. (2006). Researching and resisting: Democratic policy research by and for youth. In S. Ginwright, P. Noguera, & J. Cammarota (Eds.), *Beyond resistance: Youth activism and community change: New democratic possibilities for policy and practice for America's youth* (pp 269-285). Routledge.
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation,* 2(2), 1-4. https://doi.org/10.7275/ffb1-mm19
- Wilson, M. (2005). Constructing measures: An item response theory approach. Routledge Academic.
- Witt, E. (2018, February 19). How the survivors of Parkland began the Never Again movement. *The New Yorker*. https://www.newyorker.com/news/news-desk/how-the-survivors-of-parkland-began-the-never-again-movement
- Wu, H.-C. J., Kornbluh, M., Weiss, J., & Roddy, L. (2016). Measuring and understanding authentic youth engagement: The youth-adult partnership rubric. *Afterschool Matters*._
 https://eric.ed.gov/?id=EJ1095954