Project Topeka: Word from the Wise (Practitioners) to the Wise (Policymakers, Changemakers, Entrepreneurs) on the Promise of an AI-driven Instructional Tool

By Viki M. Young
August 2023

Digital Promise®
Accelerating Innovation in Education

Suggested Citation

Young, V. M. (2023). *Project Topeka: Word from the wise (practitioners) to the wise (policymakers, changemakers, entrepreneurs) on the promise of an AI-driven instructional tool.* Digital Promise. https://digitalpromise.org/wp-content/uploads/2023/09/Topeka-implications.pdf

Acknowledgments

The research team (Hillary Greene Nolan, Mai Chou Vang, and Viki Young) thanks Jess Alanis, Karen Cator, Merijke Coenraad, Briza Diaz, Vanessa Peters Hinton, Megan Pattenhouse, Kristal Brister Philyaw, Teresa Solorzano, Stefani Pautz Stephenson, Jeff Wayman, and Josh Weisgrau for their partnership on Project Topeka.

We are grateful to the many teachers who implemented Project Topeka and shared their experiences through various research activities. In particular, we acknowledge and honor the wisdom and expertise of Kimberly Artis, Rachel Baker, Melissa Castner, Charles Frey, Christy Gibbs, Aida Hadzovic, Elizabeth Hancock, Terry Janssen, Juvy Mojares, Cynthia Orton, Victoria Salcedo, Mary Beth Skerjanec, Millicent Twumasi, Denée Tyler, Adriana Vargas, and Theresa Wampler, who engaged with the Project Topeka team in-depth convenings around instructional AI use, writing instruction, and research findings. The research series amplifies their insights.

This material is based upon work supported by the Bill & Melinda Gates Foundation and Gates Ventures. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Bill & Melinda Gates Foundation or Gates Ventures.

Contact Information

Email: vyoung@digitalpromise.org

Digital Promise:
Washington, DC:
1001 Connecticut Avenue NW, Suite 935
Washington, DC 20036

Redwood City, CA:
702 Marshall Street, Suite 340
Redwood City, CA 94063

Website: https://digitalpromise.org/

Project Topeka: Word from the Wise (Practitioners) to the Wise (Policymakers, Changemakers, Entrepreneurs) on the Promise of an AI-driven Instructional Tool

> *[With AI, there will be a] shift in that the teacher is no longer the guru or owner of knowledge — [our role] shifts to facilitator. You're the nurturer of students and inspiration for learners. You're the lead learner, who shows students how to use a tool — we are the manipulator of the tools to the craft of learning. You still need to have interpersonal skills and ability to intercede and lead students without micromanaging every nuance.*
>
> *–ELA Teacher, Project Topeka convening, June 2022*

Every day, a new story about the promise of artificial intelligence (AI) tools in education or the concerns they provoke is pushed out to education professionals and the general public. The tools offer the potential for greater efficiency and proficiency in instruction, including saving teachers time, providing students with more equitable access to opportunities beyond their assigned teachers and the school walls, and differentiating to individual student needs. They also raise concerns about the teachers' role as we know it, the teaching profession writ large, the skills and knowledge students need to learn about AI and how to use it, and the skills and knowledge that AI might render unnecessary for students to learn.

Most importantly, what does equity mean with AI tools in teaching and learning? While AI tools may have the potential to improve equity in access, they also enter an educational system that is currently inequitable in both access and outcomes. The influence may indeed run the other way—that the inequities in the system can shape how AI tools are used. (See *Delivering on the Promise of Digital Equity* [Weaver, 2022] and *Breaking with the Past: Embracing Digital Transformation in Education* [Brizard, 2023] for discussion and strategies on instructional technology as it intersects with equity.)

Project Topeka was an early AI-driven tool designed to support teachers in argumentative writing instruction and to provide students with differentiated and timely feedback. Teachers'

perspectives based on their expertise and experiences using Project Topeka provide crucial insights central to our developing understanding of AI tools in instruction.

Digital Promise staff leading Project Topeka programming and research are releasing various findings and thought pieces to inform and further the fields of argumentative writing instruction and teacher practice with AI tools. Key topics focus on an updated description of middle school English language arts (ELA) teachers' typical writing instruction; the development of teachers' practices alongside an automated essay scoring (AES) tool; examination of scoring differences between teachers and the AES tool and explanations about why; and an in-depth look at how expert writing teachers evaluate and why they elaborate on the argumentative writing rubric used by an AES tool.

The rest of the paper series treats the findings from Project Topeka in more depth. This overview shares high-level implications drawing from the research on teachers' argumentative writing instruction and how they engaged with Project Topeka.

Project Topeka Overview

Project Topeka[1] aimed to solve enduring challenges in argumentative writing instruction—the need for more timely feedback on student writing, insufficient teacher time to provide such feedback, and uneven teacher capacity to teach argumentative writing.

Project Topeka Features. Project Topeka offered teachers of grades 6–9 ELA across the country an automated essay scoring tool that provided students with individualized line-level feedback on argumentative essays responding to six different prompts. Each prompt included aligned information sources offering multiple perspectives on the prompt topics. The tool scored student writing on a 4-point scale for each of the following four dimensions: Claim and Focus; Support and Evidence; Organization; and Language and Style. The tool also provided a total score that was the sum of those four scores (possible total scores of 4–16).

Students could activate "signal checks," which generated updated automated feedback and scores, after they revised their writing and could request signal checks at intervals they determined. Teachers had access to online curricular resources, including the writing rubric, lesson plans, graphic organizers, presentation slides, and other guidance such as a student-facing guide to using signal checks. An online community with discussion topics initiated by

[1] The Bill & Melinda Gates Foundation funded Digital Promise to develop (including overseeing development partners), pilot, and research Project Topeka in multiple waves from December 2019 through December 2022.

teacher leaders using Project Topeka provided professional learning opportunities for teachers to share teaching strategies and tips, as well as moderated professional development.

Pilot implementation waves. The program was in the pilot phase and available for teachers to use in the classroom for three different waves: winter 2020 (as the COVID-19 pandemic hit); fall semester 2020; and school year 2021–22. We also convened 16 teacher leaders in person in summer 2022 and virtually during fall semester 2022 to dive deeper into their teaching practices with AI as an instructional tool. Alongside these implementation waves, we conducted exploratory, mixed methods research on the nature of implementation and emerging outcomes for teachers and students. We conducted teacher interviews, pre- and post-implementation teacher surveys, and hierarchical modelling of student writing scores. These data provided input for planning each successive implementation wave.

Pandemic context. After the initial shock of the COVID-19 pandemic, teachers already registered to use Project Topeka surprisingly persisted in their implementation through spring 2020, largely relieved that the tool was ideally suited to support the teachers' urgent pivot to online learning. In the subsequent implementation wave in fall 2020, while the suitability of the tool for hybrid learning still held, disruptions and community stress from the pandemic continued and districts narrowed in on teaching essential standards through 2020–21. Reports of learning loss and exacerbated inequity in learning opportunities and society as a whole underpinned calls for radically different instruction and indeed education systems. In this context, Project Topeka—with a choice of relevant topics and associated information sources that varied in reading levels, media, perspectives; relatively easy adaptability to online learning; and direct feedback and scoring that addressed teachers' workloads—presented an option that might partially meet the needs of the moment.

What Do Project Topeka Teachers' Experiences Teach Us?

As mentioned, we are developing more in-depth papers on specific topics and lessons for different audiences. Here, we crystallize a few high-level implications from teachers' perspectives and data analysis that program and edtech developers, policymakers, and education systems leaders would be wise to heed. We have much to learn from the wisdom of teachers to ground the challenges AI tools can help solve and the contexts within which they will operate in the near future.

Students initially struggling the most with argumentative writing showed greater score gain, on average, than did other students, sometimes by mediating the teachers' relationship with students.

Students initially struggling the most with argumentative writing showed greater score gain, on average, than did other students.

Student score analysis over successive waves consistently showed that students beginning with the lowest scores for any given dimension or in total score made the largest score gains using Project Topeka. For example, in the last implementation wave (September 2021–June 2022), students grew an average 2.2 points,[2] with those scoring 4, 5, or 6 at baseline exhibiting higher-than-average growth.[3] Even though we do not know the extent to which this finding reflects unevenness in the level of difficulty represented at each step of the scale (e.g., moving one point at the higher end of the scale may be more difficult than moving one point at the lower end of the scale), it is clear that those who might struggle most with argumentative writing have the most to gain from the tool. At the same time, those beginning with some proficiency in argumentative writing did not, on their own, advance significantly to higher levels as scored by the tool. (See *Automated Essay Scoring in Middle School Writing: Understanding Key Predictors of Students' Growth and Comparing AI-Generated and Teacher-Generated Scores and Feedback* [Greene Nolan & Vang, 2023] for the full results of student writing scores under Project Topeka.)

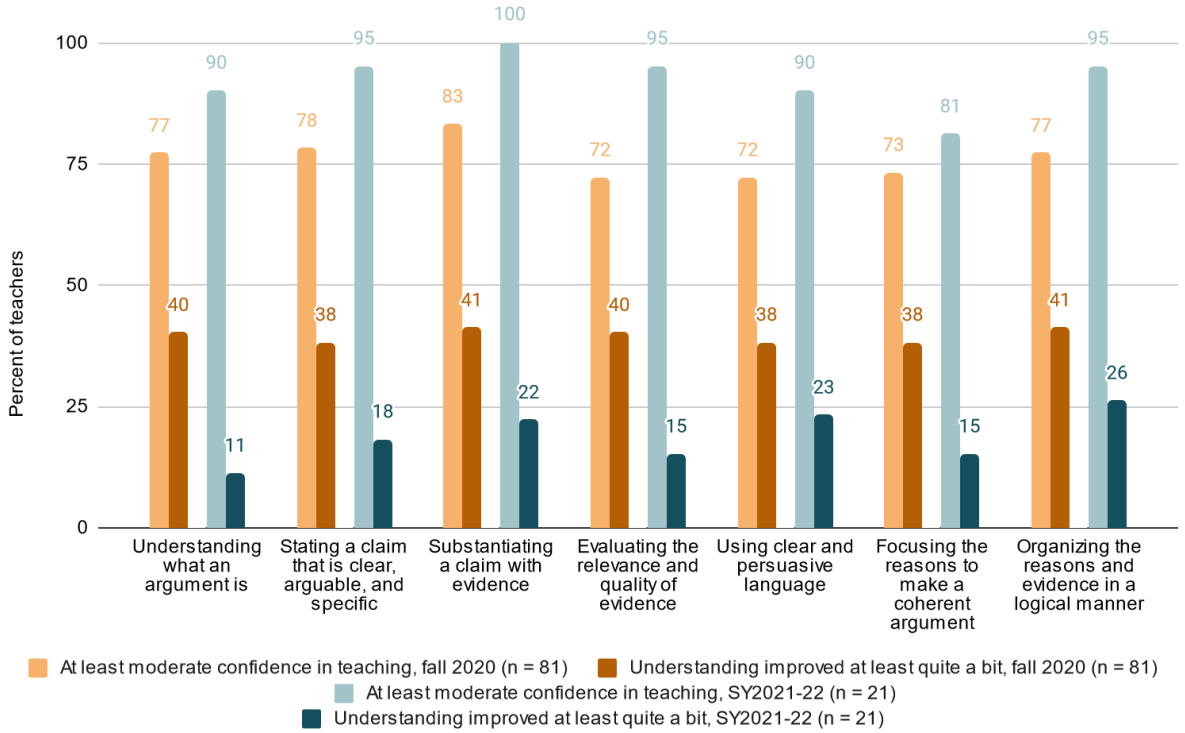
Teachers provided several possible reasons for why these students might have benefited more from Project Topeka (as implemented). Students received the automated feedback when they requested it, and teachers reported that the tool-provided feedback and kept the students going when the teacher could not have provided feedback as quickly. Teachers also reported that because the tool saved them time in providing feedback and students could work independently for longer, teachers had the time to conference with individual students, usually choosing to do so with the students who needed more help. Perhaps a less obvious revelation about why Project Topeka might have benefited students struggling with writing more was teachers' reports that some students received feedback more willingly from the tool than from the teacher. Some students perceived the tool to be more objective than the teacher and it sidestepped any trust issues in teachers' and students' relationships.

[2] Possible total score of 4 to 16, based on possible 1 to 4 points on each of four argumentative writing dimensions.

[3] We predicted students' growth on their sum scores (the total across 4 dimensions, max 4 points in each dimension) as a function of their baseline sum score, using 2-level models with students nested in teachers, while controlling for grade (8th vs. 7th), time (winter vs. fall), and prompt topic.

The professional guidance and instructional materials enveloping the AI tool supported teachers in learning argumentative writing. Even though teachers consistently expressed high confidence in teaching argumentative writing, a notable percentage reported increased understanding of the seven dimensions we asked about, more so in the third wave in fall 2020 (Exhibit 1).

Exhibit 1. Teacher Reported Confidence in Teaching and Improved Understanding of Argumentative Writing Dimensions



Note: Confidence ratings based on a 4-point scale, where 1 = Not at all confident, 2 = Somewhat confident, 3 = Moderately confident, and 4 = Very confident

Understanding ratings based on a 4-point scale, where 1 = Same understanding as before, 2 = Somewhat better understanding, 3 = Quite a bit better understanding, and 4 = Much better understanding

Source: Digital Promise Project Topeka teacher surveys, matched pre- and post-implementation results in two most recent implementation periods

Teachers also reported statistically significant increases in their preparation to teach argumentative writing as a result of using Project Topeka (Exhibit 2). Moreover, data indicated that compared with those more confident, teachers initially less confident in teaching

argumentative writing reported marginally better understanding of argumentative writing after implementing Project Topeka (fall 2020, mean 2.6 vs. 2.2 on 4-point scale, $p = 0.067$, $n = 81$).

Exhibit 2. Project Topeka Teachers' Reported Preparation to Teach Argumentative Writing, Before and After Implementation Across Three Implementation Waves

Implementation Wave	*n*	Pre-implementation mean	Post-implementation mean	*p*
Fall 2020	81	2.8	3.4	< 0.0001
Fall 2021–Spring 2022	21	2.7	3.4	< 0.001

Note: Scores based on a 4-point scale, where 1 = strongly disagree and 4 = strongly agree

Source: Digital Promise Project Topeka teacher surveys, matched pre- and post-implementation results in three different implementation periods

Even teachers with a high degree of expertise and experience in writing instruction and who had many of their own writing materials already developed reported that the Project Topeka resources gave them a common language they could use to teach students skills associated with argumentative writing: *"I have new vocabulary to explain how to revise writing"* (middle school ELA teacher, spring 2022). Without these professional learning scaffolds so easily accessible, teachers would have had to seek or build those bridges themselves, likely making the tool less useful and usable. (See *A Descriptive Analysis of Teachers' Writing Instruction Using Project Topeka, an Automated Essay Scoring Tool* [Vang, 2023] for a more comprehensive description of teachers' instruction using Project Topeka.)

> [T]eachers clearly articulated the necessary role they play in mediating the feedback the tool provided to students.

Teachers played a crucial role in mediating the tool-generated automated feedback that students receive. Both of the above themes lead to the most important implication, that teachers clearly articulated the necessary role they play in mediating the feedback the tool provided to students. At a basic level, the tool required a minimum amount of writing before it could give feedback and score it, which for some students, teachers needed to provide that initial explanation and motivation. Students also had difficulty interpreting the feedback independently, got frustrated at multiple attempts to respond to the feedback without seeing their score change, or wondered why they received new comments the tool had not provided before when they asked for signal checks. Because the scoring tool did not read the essay to provide holistic feedback, teachers needed to provide it and to coach students on different ways to apply the tool-generated feedback since addressing piecemeal, line-level feedback

often does not lead to a coherent, flowing essay. Without this teacher partnership with the AI tool, using the tool as a standalone substitute for the teacher meant that students could plateau and would receive little instruction in analysis and revision skills. Crucially, we began to see equity concerns emerge from differences in this instructional stance vis à vis the tool, with teachers reporting higher percentages of special education students on average adopting a substitute approach to AI tool use requiring more independent student efforts than teachers with fewer special education students. With variation in this stance towards using AI tools, we need to ask what the benefits and costs are for the most vulnerable students if they routinely receive more standalone tech-based instruction. (See *Teaching Partner, Grading Assistant or Substitute Teacher?* [Greene Nolan, 2023] and *Teaching Partner, Grading Assistant, Substitute Teacher: Three Approaches to Teaching Middle School Writing with Artificial Intelligence* [Greene Nolan et al., under review] for a deeper elaboration on these stances and their implications for equity and the teaching profession.)

The promise of applicability and scale for Project Topeka was circumscribed by contexts at multiple levels. As with any education improvement initiative, implementation is situated within embedded contexts that include (at least) the *students'* cultures, identities, prior education experiences, learning, and achievement; *teachers'* subject-specific and general teaching preparation, experiences, and expertise, and their orientation towards and expectations for students; *school* culture, climate, leadership, professional community, and resources; *district* policy, resources, and leadership; and myriad *state* policies.

The focus on argumentative writing, with emphasis on identifying and using evidence to advance a thesis or argument, has broad applicability as a skill across disciplines and is fundamental for an informed citizenry. Thus Project Topeka and other argumentative writing tools ought to have great potential in the school curriculum. The experiences teachers shared identified the realities that need to be addressed for broader applicability. At the secondary level, content coverage tends to be paramount, reinforced by state assessment regimes, which led to some teachers using Project Topeka solely as an assessment. Teachers reported having limited time for writing instruction and what time they had needed to cover argumentative writing and other genres for that grade level. Few teachers offered more than one full argumentative essay writing cycle (from prewriting through revision) and many reported that students needed each step to be broken down. Teachers explicitly taught and provided practice at each step for the class together, such that the full essay culminated as the sum of the individual pieces, often with little or no opportunities for engaging in the revision process. Teachers reported spreading these steps out through a semester or even full year, thus limiting the amount of time they could use the tool to support more writing practice. These conditions are currently outside of what a tool like Project Topeka was designed to address but

nonetheless, from teachers' perspectives the conditions shape whether and how they can use the tool.

Epilogue: Since Project Topeka

Since Project Topeka implementation concluded, GenerativeAI (GenAI) tools like ChatGPT have ignited debate about whether students should be allowed to use it to originate text and if so, what writing skills they need to learn alongside the functionality of a text-generating AI tool like ChatGPT.

Drawing on data from Project Topeka and an Inclusive Innovation pilot (an equity-centered R&D model) that brought together teams of students, families, community members, teachers, and school and district leaders to address the challenge of engaging secondary-level students in writing, we learned root causes inhibiting student engagement and writing quality that provide some food for thought.

- **Argumentative writing is a form of argumentation—that is, skill in analysis.** While an AI tool might be able to draw information from the internet and marshal it into an argument ostensibly addressing the specific writing need, the user (who would have been the writer) still needs to critically assess whether the argument is the best argument, still needs to evaluate the evidence for veracity, relevance, and strength. Most importantly—and perhaps made more difficult by the presence of existing text—the user needs to identify what is missing and different possible constructions of the problem (problem-framing) and therefore how to conceptualize the overall argument. The teachers in our sample almost universally lamented the difficulty of teaching students how to revise and how to motivate them to revise their writing. Relatively polished and grammatical text as generated by an AI tool will likely erode that motivation even more because the generated text *sounds* complete and error-free, and the quality of a student's writing might not be as refined and revision may appear to decrease the writing quality.

- **Power of suggestion is strong—indeed some may call it insidious.** A complete written piece embeds choices of what to include and what to exclude. The influence that the AI tool has on content and perspective has the possibility to be broader in that sources may span a wider range than what a user would be able to read, absorb, and use quickly. It also holds the potential to shape thinking—thus amplifying and not diminishing—the need for students to master analytic and critical thinking skills.

- **Students' desires and teachers' goals (and content standards) to increase authentic student voice in their writing** and other forms of expression appear on the surface more difficult to accomplish with an AI tool. Or that expression may be reduced to a series of commands such as "use language demonstrating passion," as opposed to students' expressing that passion in their own words. As a middle school student on an Inclusive Innovation team explained, one root cause of students' lack of engagement in writing assignments lies in the difference between "forced writing" (teacher-assigned/required writing) and "freedom writing" (about something the writer cares about and for the purpose of expressing and communicating rather than being assessed). What is freedom writing to a text-generating AI tool? What does student self-expression on societal challenges, history, pop culture mean when the expression itself is generated by a tool that suggests words and phrases not grounded in the writer's experience and understanding?

Other emerging AI tools such as intelligent tutoring systems that can interact with students in prewriting stages (e.g., prompting students to reflect on the quality of evidence they have identified) and during writing (e.g., in laying out a logical flow) may hold promise in continuing to teach students the critical skills underpinning argumentative writing, as opposed to writing for the individual.

New AI tools are entering the market rapidly and will evolve continuously, and evidence of their impact on teaching and learning will fuel debate for the foreseeable future. In that debate, we must remember the teachers and students closest to the education challenges that the AI tools are trying to solve. They have invaluable expertise and lived experience on whether and how those tools have benefits and consequences for students furthest from opportunity and whether and how those tools can unlock more equitable access to powerful learning opportunities and outcomes.

References

Brizard, J.-C. (2023, April.) *Breaking with the past: Embracing digital transformation in education.* Digital Promise. https://doi.org/10 51388/20 500 12265/176

Weaver, D. J. (2022, Dec.). *Delivering on the promise of digital equity.* Digital Promise. https://digitalpromise.org/wp-content/uploads/2022/12/Delivering-on-the-Promise-of-Digital-Equity.pdf

Project Topeka Paper Series

Greene Nolan, H. (2023, March 1) *Teaching partner, grading assistant, or substitute teacher.* Digital Promise. https://www.edsurge.com/news/2023-03-01-teaching-partner-grading-assistant-or-substitute-teacher

Greene Nolan, H., Pattenhouse, M., & Young, V.M. (under review). *Teacher partner, grading assistant, substitute teacher: Three approaches to teaching middle school writing with artificial intelligence.* Digital Promise.

Greene Nolan, H., & Vang, M. C. (2023, June 14). Can Ai help teachers with grading? Digital Promise. https://www.edsurge.com/news/2023-06-14-can-ai-help-teachers-with-grading

Greene Nolan, H., & Vang, M. C. (2023) *Automated essay scoring in middle school writing: Understanding key predictors of students' growth and comparing AI-generated and teacher generated scores and feedback.* Digital Promise. https://digitalpromise.org/wp-content/uploads/2023/09/Topeka-score-analysis.pdf

Vang, M. C. (2023). *A descriptive analysis of teachers' writing instruction using Project Topeka, an automated essay scoring tool.* Digital Promise. https://digitalpromise.org/wp-content/uploads/2023/09/Topeka-writing-instruction-description.pdf

Young, V. M. (2023, July 24). *Expertise as elaboration: Teachers' reflection on an AI tool-embedded writing rubric.* Digital Promise. https://www.edsurge.com/news/2023-07-24-expertise-as-elaboration-teachers-reflections-on-an-ai-tool-embedded-writing-rubric