

Elementary Teachers' Perceptions of Automated Feedback and Automated Scoring:
Transforming the Teaching and Learning of Writing Using Automated Writing Evaluation

Joshua Wilson, Ph.D.^a (ORCID#: 0000-0002-7192-3510)

Cristina Ahrendt^b

Emily A. Fudge^c

Alexandria Raiche^d

Gaysha Beard, Ed.D.^e

Charles MacArthur, Ph.D.^f (ORCID#: 0000-0002-7202-5536)

Author Note

^aJoshua Wilson, School of Education, University of Delaware; ^bCristina Ahrendt, School of Education, University of Delaware; ^cEmily Fudge, School of Education, University of Delaware; ^dAlexandria Raiche, School of Education, University of Delaware; ^eGaysha Beard, Red Clay Consolidated School District, Delaware; ^fCharles MacArthur, School of Education, University of Delaware.

This research was supported in part by Grant R305H170046 from the Institute of Education Sciences, U.S. Department of Education, to the University of Delaware. The opinions expressed in this paper are those of the authors and do not represent the views of the Institute, or the U.S. Department of Education, and no official endorsement by these agencies should be inferred.

Declarations of interest: none.

Correspondence concerning this article should be addressed to Joshua Wilson, Ph.D., University of Delaware, School of Education, 213E Willard Hall Education Building, Newark, DE, 19716, United States. Tel: +13028312955. Email: joshwils@udel.edu.

Citation:

Wilson, J., Ahrendt, C., Fudge, E., Raiche, A., Beard, G., & MacArthur, C. A. (2021). Elementary teachers' perceptions of automated feedback and automated scoring: Transforming the teaching and learning of writing using automated writing evaluation. *Computers & Education*, 168, 104208. <https://doi.org/10.1016/j.compedu.2021.104208>

Abstract

The present study used a focus group methodology to qualitatively explore elementary writing teachers' attitudes and experiences using an automated writing evaluation (AWE) system called MI Write as part of a districtwide implementation of MI Write in Grades 3-5 in 14 elementary schools. We used activity theory as a theoretical framework to answer the following research questions: In what ways do teachers perceive AWE to mediate their own, and their students', activity in the elementary writing classroom? How, and in what ways, do teachers perceive AWE to transform the teaching and learning of writing? Using an inductive coding method, four themes emerged that highlight the complexities of using AWE to support instruction: (1) AWE both assists and challenges teachers; (2) AWE influences students' writing motivation and independence in complex ways; (3) AWE functionality may create new instructional challenges; (4) AWE implementation and efficacy is influenced by broader aspects of the instructional activity system. Findings have implications for informing district and school stakeholders interested in the use of AWE to support writing outcomes advancing research on AWE, for guiding future research, and for guiding the design of AWE systems.

Keywords: Automated writing evaluation; automated feedback; automated scoring; teacher perceptions; writing instruction.

**Elementary Teachers' Perceptions of Automated Feedback and Automated Scoring:
Transforming the Teaching and Learning of Writing Using Automated Writing Evaluation**

1. Introduction

Despite the importance of writing in K-12 and postsecondary settings (National Commission on Writing for America's Families, Schools, and Colleges, 2003, 2004, 2005), most US students lack sufficient writing skills (National Center for Education Statistics, 2012; Persky, Daane, & Jin, 2002). In researching ways to improve writing instruction, researchers and educators have become interested in technology-based interventions (e.g., Morphy & Graham, 2012) because of their scalability and the success of technology-based interventions for improving student learning in other subject areas, such as mathematics and science (see Hillmayr, Zierwald, Reinhold, Hofer, & Reiss, 2020; Zhu, Liu, & Lee, 2020). Technology-based interventions are also of interest because of their potential to address some of the ubiquitous barriers facing writing instruction, such as the time costs of evaluating writing (Warschauer & Grimes, 2008) and the difficulty of providing students with timely, effective feedback (Parr & Timperley, 2010).

One promising technology-based writing intervention is *automated writing evaluation* (AWE). AWE is software that provides students with immediate automated feedback, typically coupled with automated essay scoring, intended to help students revise and improve their writing (Grimes & Warschauer, 2010; Wilson & Roscoe, 2020). AWE is intended to support the teaching and learning of writing by reducing teachers' grading load, enabling teachers to provide more frequent opportunities for writing practice and greater amounts of feedback on higher-level writing skills (Kellogg, Whiteford, & Quinlan, 2010; Wilson & Czik, 2016). Research shows that AWE helps students improve their writing quality (Franzke, Kintsch, Caccamise, Johnson, & Dooley, 2005; Palermo & Thomson, 2018; Palermo & Wilson, 2020; Stevenson & Phakiti, 2014), their

ability to effectively revise their writing (Knight et al., 2020; Wang et al., 2020; Wilson, 2017; Wilson et al., 2014), their writing motivation (Grimes & Warschauer, 2010; Wilson & Czik, 2016), their attitudes towards writing (Roscoe, Allen, Johnson, & McNamara, 2018), and their writing self-efficacy (Wilson & Roscoe, 2020).

Nevertheless, while efficacy studies have reported favorable results, there remain concerns regarding AWE implementation in classroom settings, such as AWE will come replace the teacher as a feedback agent and that students who use AWE will produce formulaic writing that emphasizes length and complexity over quality (Conference on College Composition and Communication, 2014; Ericsson & Haswell, 2006; Herrington & Moran, 2001; National Council of Teachers of English, 2013; Perelman, 2014). Unfortunately, it is difficult to affirm or alleviate such concerns because the majority of AWE research focuses on: (a) documenting the reliability of the underlying scoring algorithms rather than documenting ways that AWE may support instruction (e.g., Shermis, 2014; Shermis & Burstein, 2013); (b) testing the effects of AWE as a stand-alone technology rather than testing the use of AWE when integrated with teacher instruction (c.f., Palermo & Thomson, 2018); (c) evaluating AWE with students in secondary and post-secondary setting (e.g., Koltovskaia, 2020; Link, Mehrzad, & Rahimi, 2020; and see Stevenson & Phakiti, 2014), as well as in English as a Foreign Language or English as a Second Language contexts (Koltovskaia, 2020; Li et al., 2019; Link et al., 2020; Ranalli, Link, & Chukharev-Hudilainen, 2017); and (d) evaluating efficacy with relatively little attention paid to documenting teachers' perceptions of AWE (c.f., Palermo & Wilson, 2020; Wilson et al., 2021).

It is important to investigate teacher perceptions because they have been shown to influence whether an educational technology system, such as AWE, is utilized well, or at all, for classroom instruction (Backfisch et al., 2021; Ertmer, 1999; Tondeur, van Braak, Ertmer, & Ottenbreit-

Leftwich, 2017). Indeed, teacher perceptions of the usefulness of educational technology comprise an important dimension when measuring teachers' technological-pedagogical content knowledge (Saubern, Urbach, Koehler, & Phillips, 2020). Thus, teacher perceptions of AWE present an important source of evidence—evidence of social validity (Lyst, Gabriel, O'Shaughnessy, Meyers, & Meyers, 2005)—when evaluating the efficacy of AWE.

To our knowledge, prior research on teachers' perceptions of AWE has exclusively focused on secondary teachers; no prior studies have examined elementary teachers' perceptions of AWE. Secondary teachers generally perceive AWE to assist with reducing grading, increasing student motivation, and simplifying classroom management during writing instruction (Klobucar et al. 2013; Palermo & Thomson, 2018; Roscoe & McNamara, 2013; Stevenson, 2016; Warschauer & Grimes, 2008). Secondary teachers have criticized the accuracy and validity of automated scoring and the ability of AWE to make teaching writing more enjoyable (Grimes & Warschauer, 2010). However, it is unclear whether findings generalize to elementary teachers. Thus, the present study explored elementary teachers' perceptions of an AWE system called *MI Write* for supporting writing instruction in Grades 3-5. We were specifically interested in ways that AWE might mediate the activity of teachers and students in the elementary writing classroom, and ways that AWE might transform the teaching and learning of writing.

1.1 *MI Write*: Automated Writing Evaluation Based on Natural Language Processing

MI Write (www.miwrite.net) is a web-based AWE system developed and marketed by Measurement Incorporated. *MI Write* is an interactive learning environment that facilitates a number of pedagogically-supported interactions among system, student, teacher, and peers (see Graham, Hebert, & Harris, 2015; Graham, McKeown, Kiuahara, & Harris, 2012). Prior to 2019, *MI Write* was named *PEG Writing*. Though the present study was conducted prior to 2019, to

facilitate readers identifying the relevant, existing AWE system, we continue to refer to the AWE system as MI Write.

In MI Write, system-student interactions center on students submitting their writing to MI Write for immediate evaluation and feedback. Each time a student submits a new revision, MI Write provides a new round of evaluation and feedback. MI Write's automated feedback system is undergirded by its automated essay scoring (AES) system known as *Project Essay Grade* (PEG; Page, 2003), a system developed using natural language processing techniques and machine learning algorithms (i.e., support vector regression).

MI Write uses PEG to evaluate the quality of students' writing across six dimensions, or "Traits," of writing quality—the Six Trait scoring model is a widely-adopted analytic scoring system intended to link assessment and instruction (Coe, Hanita, Nishioka, & Smiley, 2011). MI Write provides students with an AES score (range = 1–5) for each of the following traits: development of ideas, organization, style, word choice, sentence fluency, and conventions. In addition, students receive an overall score to characterize their writing performance holistically; the overall score (range = 6–30) is formed as the sum of the six trait scores. MI Write provides students with trait-specific automated feedback to help students evaluate their writing (i.e., meta-cognitive prompts, such as "Does your writing have a clear conclusion?") as well as suggestions for improvement (e.g., "Many of your sentences are very simple and begin in the same way. Add introductions or descriptive phrases to add more information and create more interesting sentences."). MI Write also provides students with in-line feedback regarding potential spelling and grammar errors, as well as links to MI Write's interactive skill-building lessons.

In MI Write, system-teacher interactions begin with teachers assigning writing activities. MI Write's automated scoring and feedback system is designed to be prompt-independent. This

means that teachers have the option of assigning pre-packaged writing prompts and teacher-created, curriculum-specific writing prompts. With any writing assignment teachers can require students to use one of MI Write's electronic graphic organizers. Teachers may also adjust a host of advanced options such as specifying the amount of time students have to complete the task or the maximum allowable number of drafts or require peer review based on groups that the teacher designates randomly or deliberately. Teachers also interact with MI Write by reviewing its automated scores and feedback for students' writing. Teachers view students' score reports and may provide supplemental feedback through in-line or summary comments. Because MI Write is a prompt-independent scoring system, it does not evaluate prompts for accuracy of content, only writing quality, making it important for teachers to also view and provide feedback on student writing. Finally, teachers may utilize MI Write's suite of reporting functions to monitor classroom or individual performance, and consequently adjust their instruction.

Student-teacher interactions within MI Write are enabled by its instant messaging function. Students can leave a comment for their teacher to request help, ask a clarifying question, or engage in other relevant dialogue. Teachers can either respond to or initiate these dialogues. In addition, teachers and students interact through the supplemental feedback that teachers provide.

Student-peer interactions within MI Write consist of peer review. MI Write facilitates either identifiable or anonymous peer review, which is decided by the teacher on a prompt-to-prompt basis. At the time this study was conducted, MI Write's peer review utilized the "Two Stars and A Wish" format (Webb & Jones, 2009), in which peers identify two things the author did well and one thing the author could improve. The peer review functionality has since been updated.

2. Theoretical Framework

The present study adopted *activity theory* (Engeström, 1987, 2000; and see Zheng, Kim, Lai, & Hwang, 2020 for a recent application of activity theory with educational technology) as a lens for understanding how the adoption of AWE may influence teachers (teaching) and students (learning). An extension of Vygotsky's (1978) notions of socially-mediated learning via the zone of proximal development, activity theory considers learning, and mental and psychological development, to result from tool-mediated activity in pursuit of a goal (Engeström, 1987; Jonassen, 2000; Russell, 1995). Tools, therefore, are a central focus within activity theory. Tools are historically- and culturally-situated (Russell, 1995), and include cultural tools (language, signs, symbols) as well as physical/technological tools (e.g., pencil and paper, word processor, AWE). Tools are considered mediators because they are used by a subject to interact with and alter the environment to achieve [an objective](#). Activity theory posits that collective goal-oriented activity is organized in an *activity system*, which has the following interacting components (Engeström, 2000; Jonassen, 2000):

- *subject(s)*—the individual(s) engaged in goal-directed activity;
- *objective*—the goal towards which a subject's activity is aimed; an objective typically entails the production of artifacts, be they physical, mental, or symbolic (Jonassen, 2000);
- *mediating tools*—the means by which subjects interact with and alter their environment to achieve their goal; mediating tools shape the way that people act and think (Jonassen, 2000);
- *rules*—the system of stated or normative regulations that shape activity within a system (Jonassen, 2000);
- *community*—the individuals who participate in, guide the activity of, or benefit from and consume the artifacts produced by an activity system (Jonassen, 2000; Whipp et al., 2005);

- *division of labor*—the way in which activity is divided among different subjects within the activity system (Jonassen, 2000).

When viewed through the lens of activity theory, writing instruction may be viewed as an activity system. The subject, a teacher, may be assumed to hold the objective of helping students learn to write proficiently—we recognize that no single objective characterizes the intentional activity of all teachers within this activity system, and that individual teachers may have multiple objectives (e.g., to have students learn content, to have students pass a state test, to have students behave appropriately etc.), but for sake of illustration we assume the aforementioned objective is broadly applicable to any elementary teacher engaged in writing instruction. To achieve his or her objective, the teacher will engage in activity, and this activity involves the use of certain tools, such as language, curricula, pedagogical techniques (Graham et al., 2012), as well as educational technology tools, such as word processors (Little, Clark, Tani, & Connor, 2018), intelligent tutoring systems (McNamara, Crossley, & Roscoe, 2013) or AWE systems (Graham et al., 2015; Lee, 2017). These tools mediate (i.e., transform) the teacher's activity (i.e., instruction), as well as the student's activity (i.e., learning). Thus, from the perspective of activity theory, a teacher's choice of tools may have profound effects on instruction and learning.

Although the heart of the writing-instruction activity system is a teacher who utilizes tools to mediate his or her activity in pursuit of the [objective](#) of improving students' writing proficiency, other elements of the activity system influence a teacher's activity. First, a teacher must act according to certain rules. In the case of writing instruction, one critical set of rules is that which governs the amount of time devoted to writing instruction and what curriculum teachers should use. [Rules also relate to the influence of accountability assessments on curriculum and instruction, leading teachers to emphasize or de-emphasize aspects of the curriculum \(i.e., washback; Au,](#)

2007). In addition, rules governing the conventional instantiation of different genres in written form (see Bhatia, 2002) influence how those genres are taught and assessed in school. These rules are outside of teachers' control, but influence teachers' instructional activity.

In addition, a teacher acts in accord with the beliefs and values, and available knowledge, of the community(ies) with which he or she identifies. For example, there are different beliefs about how students best learn to write and those beliefs manifest in different pedagogical methods, such as the process writing approach (Graham & Sandmel, 2011) or strategy instruction (Graham et al., 2012). Further, writing communities influence beliefs and establish, formally or informally, rules regarding what "counts" as high-quality writing within that community (Graham, 2018), thereby influencing assessment standards and feedback given to students to help them meet those standards. The values and beliefs of the communities with which a teacher identifies will arguably influence which educational technology tools a teacher employs in the classroom (Ertmer, 1999; Roscoe, Cooke, Branaghan, & Craig, 2018; Tondeur et al., 2017), and how those tools are used to in the classroom (Backfisch et al., 2021).

Finally, a teacher's activity is influenced by the broader division of labor within the activity system. In the case of writing instruction in elementary school, divisions of labor may be created externally such as when schools departmentalize writing instruction. Divisions of labor may also be created internally: a teacher may gravitate to certain tools, such as AWE, if the tools support a teacher's desired division of labor, such as the division between teaching and grading.

3. Purpose of the Present Study and Research Questions

Given the increasing prevalence of AWE in elementary schools, as well as the dearth of research on AWE implementation and teacher perceptions of AWE at the elementary level, the present study utilized focus group methodology to qualitatively explore teachers' perceptions of

AWE as a tool to support writing instruction. Specifically, we adopted activity theory as a lens for answering the following research questions:

RQ1: In what ways do teachers perceive AWE to mediate their own, and their students', activity in the elementary writing classroom?

RQ2: How, and in what ways, do teachers perceive AWE to transform the teaching and learning of writing?

4. Methods

4.1 Context, Data Source, and Participants

The current study was conducted as part of a research-practice partnership between a university and a school district in the mid-Atlantic region of the United States. The school district adopted MI Write for Grades 3-5 in all 14 of its elementary schools beginning January 2016. The present study reports results of focus groups conducted with teachers in Spring 2018, the Spring of the district's first full year of district-wide implementation of MI Write.

The school district serves approximately 17,500 students, 42% of whom are White, 27% Hispanic or Latino, 21% Black, 6% Asian, approximately 4% Multi-Racial, and <.05% American Indian or Native Hawaiian. Low-income students constitute 28% of the student body. English learners and students with disabilities comprise 15% and 16% of the student body, respectively. The percent of students demonstrating proficiency in English language arts as measured by the Smarter Balanced ELA test was 46% in Grade 3, 51% in Grade 4, and 56% in Grade 5.

Focus groups were conducted at seven elementary schools that represented a diversity of locale (urban, suburban, town), socio-economic status (Title 1 versus non-Title 1), and student demographics. Within each school, focus groups were conducted with grade level teams with two exceptions due to scheduling conflicts. In those cases, focus groups were conducted with all ELA

teachers in Grades 3-5. In all, 17 focus groups were conducted with a total of 90 teachers who served approximately 3,000 students. Focus group participants were mainly White women who taught in the district for 3+ years. Names and other demographic characteristics were not collected to preserve anonymity.

4.2 Measures

Focus groups were conducted using a semi-structured protocol developed in collaboration with the school district (see Appendix). Questions probed the following areas: (a) the frequency and manner that teachers used MI Write that year, (b) the procedures and routines teachers developed for integrating MI Write into instruction, (c) the barriers teachers and students faced while using MI Write, (d) perceived benefits of MI Write for students and teachers, (e) limitations of MI Write, and (e) supports that teachers would find helpful for using MI Write effectively. All focus groups ($n = 17$) were audio recorded and then transcribed verbatim into Word documents. The 17 anonymized transcripts were the basis of qualitative analysis reported in the present study.

4.3 Qualitative Data Analysis

Qualitative data analysis followed a general inductive approach (Thomas, 2006). First, four members of the research team independently read and reread focus group transcripts to generate an initial impression of the data. Impressions were shared at research team meetings to develop a mutual understanding and create initial categories and themes for coding. This first stage relied on open coding. Second, the research team repeated this process multiple times over the course of several weeks, each time with a new sample of transcripts; existing categories and themes were revised and refined. Links between existing categories and themes were identified and categories consolidated. This stage of analysis relied on both open and axial coding. Upon completing this stage of analysis, four themes emerged. Third, the trustworthiness of the thematic coding was

assessed in two ways: (a) through weekly discussions and debriefings to ensure consensus and consistency of interpretation, and (b) formal inter-coder reliability checks. We conducted seven rounds of reliability checks with each round involving five quotations that were independently coded by three researchers. Average reliability, as measured by percent exact agreement between the three raters, was very high ($M = 95.28\%$, $SD = 6.21\%$, range = 87-100%). Finally, having ensured the coding system was reliable, we independently reviewed the transcripts to select illustrative quotations to convey the essence of the four themes.

5. Results

Each of the four themes that emerged from the qualitative analysis is discussed in turn.

5.1 Theme 1: AWE Both Assists and Challenges Teachers

5.1.1 Assistance. Teachers spoke positively of the benefits of MI Write as a tool for helping achieve their pedagogical objectives. “MI Write is a good partner with the classroom teacher,” said one teacher. Teachers primarily credited this benefit to MI Write’s ability to increase the amount and immediacy of feedback. Teachers viewed MI Write’s feedback as a helpful source of suggestions for students, other than what the teacher could provide, an “additive to teacher instruction,” as described by one teacher. Common praise was, “MI Write can cover many more aspects and give feedback...quicker than I could do for all my students.” These comments underscore teachers’ perceptions of how MI Write positively mediated divisions of labor within their instructional activity system.

Moreover, teachers concurred that a unique advantage of MI Write as a tool to support writing instruction was that it “teaches them that you’re not done the first time you write something. It’s okay to have to go back.” Teachers agreed that MI Write helped them teach students that writing is a process and revising is important. In this way, MI Write served to

reinforce community beliefs (i.e., that writing is a process) and rules that teachers sought to enforce (i.e., repeatedly revise and improve your writing).

In addition, teachers perceived MI Write to be a helpful tool for transforming their own understanding and beliefs about students' needs. Many teachers expressed that prior to using MI Write, they struggled to diagnose and prioritize students' needs; as one teacher put it, she wondered, "Where am I going to start?" In contrast, teachers praised MI Write for helping identify a starting place when conferencing with individual students. One teacher summarized the benefit this way: "It helped me know where they were and what they can do." In turn this feature assisted teachers with prioritizing feedback on specific skills:

When they come and they say, "Oh, look at my scores!" I'm like, "Well, look at this. You could do better at that; now go back and add some more sentences, or strengthen, or make your sentences longer or put more adjectives in it."

Teachers appeared to perceive MI Write as authoritative with respect to its ability to reinforce assessment norms and rules, and thus its scoring and feedback facilitated teachers' own analysis and assessment activity. In turn, this further aided teachers in achieving a desirable division of labor. One teacher said, "The immediate feedback is so important cause it's impossible to grade like 20 essays and then give them back within a week." Another teacher said, "I told them I'm not checking spelling and grammar, because you can do that on MI Write...I am conferencing with you because 'Do you have facts? Do you have details? Do you have transitions?...Is it well organized? Is it flowing?'" A third teacher said, "[MI Write] gets all the little things out of the way like the spelling errors and little things. Like 'Ok this is the little stuff and now we can focus on the bigger stuff.'" Overall, MI Write was seen as decreasing teachers grading time—"When they do conference with us, we don't spend an hour trying to fix all their mistakes"—and this enabled teachers to prioritize their feedback on students' instructional needs.

Finally, teachers perceived MI Write's automated feedback and scoring as assisting teachers with communicating with other community members, such as parents. Several teachers expressed that "the reports were very helpful for conferences with parents." One teacher used MI Write's data during a parent-teacher conference as evidence of a student's level of effort:

I sat down with a parent. "Look, they had a week to do this. They made one revision and resubmitted and they were done....They could have had as many revisions as they needed." So it was nice for showing to parents and holding the kids accountable.

Teachers appreciated that MI Write helped present objective evidence of student effort and performance to parents. [In this way, MI Write's mediated teachers' activity with respect one of the communities that influenced their activity system.](#)

5.1.2 Challenges. Teachers simultaneously perceived MI Write's automated feedback as creating new instructional challenges. First, MI Write's feedback was perceived as misaligned to teachers' instruction. One teacher stated, "I don't feel like the things that they were checking in MI Write aligned with what we were teaching in writing." Another teacher shared, "We are having trouble having that mesh with MI Write, to get direct feedback on the actual skill we taught for the day." Teachers desired to use MI Write to support daily instruction, but teachers struggled to do so given the decontextualized nature of MI Write's automated feedback.

In addition, MI Write's automated scoring created instructional challenges for teachers. First, teachers shared that, since their curriculum did not use the Six Trait scoring system, they struggled to align the scoring in MI Write with curriculum- and district-specific scoring rubrics:

Where [MI Write's] rubric was just argumentative, then ours was a speech, which was different. Or we are doing a science a fiction narrative, but it's going to score it just on the functions of a narrative speech, not a science fiction narrative.

[MI Write reinforced a different set of rules than those the teachers wished to prioritize.](#)

Teachers also voiced concerns about the accuracy of the automated scoring and its effect on

students, again struggling with MI Write's enforcement (and reinforcement) of a different set of assessment rules than those that typically governed their activity system. The automated scoring was at times perceived as too stringent: "Yeah, I kind of curve the PEG score...I always curve it because I feel like PEG is a little harsh." Other times the scoring was regarded as too lenient: "If I felt like it was inflated in one area or the other, I would adjust that for myself. It would be nice if it were more closely aligned." For these reasons, teachers concurred that the score was strictly for formative use and not for summative, grading purposes: "We don't use the PEG score at all. I tell [students] to use it as they're revising and editing."

Of greater concern to teachers was that students would internalize the wrong message about writing quality (i.e., the wrong rules). The concern stemmed from teachers' perception that MI Write's scores "go by the number of words." For instance, one teacher stated: "The more they write, even if it's not great writing, the score goes up...or quality writing that is not as long is scored lower." Another teacher shared: "They are like, 'Oooo! Look, I'm at 358 words.' But are they good words? Are they quality words? Do they really say what you want to say?" Consequently, teachers were concerned that MI Write's scores would lead students to value quantity over quality.

Finally, teachers struggled to adapt to MI Write's consistent and unwavering feedback and scoring with their own feedback and grading methods that accounted for individual differences and progress within the curriculum. That is, MI Write's automated feedback and scoring does not adjust its expectations for performance based on student ability, or timing within the instructional cycle or school year, as a teacher would. One teacher explained the difference like this:

You're picking and choosing what you're focusing on with different kids. Some kids you're focusing on sentence structure and that's the most important thing, you need that first. Some kids you're really focusing on like your voice or your sentence fluency or your sentence structure. So it's hard when technology scores it one specific way.

Yet another teacher described the difference in feedback like this:

Like I know that I don't start with making sure that your paragraph is indented at the beginning of the year, like their narrative is just making sure that they have the key features of a narrative. Once we get to the opinion then I'm like alright you have to indent the first paragraph.

In contrast to the flexible, graduated, and individualized manner of teacher feedback and grading, MI Write applied the same standard for all students **and reinforced those standards (i.e., rules)** regardless of whether teachers expected students to have mastered the skill at that point in the year.

5.2 Theme 2: AWE Influences Students' Writing Motivation and Independence in Complex Ways

5.2.1 Writing Motivation. Teachers concurred in their perception that MI Write **mediated students' activity via its** positive effect on students' writing motivation. The ability to receive an immediate score and to see progress was perceived to increase students' writing motivation. Teachers frequently made comments like, "They love seeing their score grow; that's hugely motivating," and "The kids want to improve. So, they get their scores and it's a motivation for them to go back and revise or edit because they want their score to go up." Indeed, a common refrain was, "Students were excited to see and try to improve their scores." One teacher described the phenomenon like this:

Yeah, I like the score because the kids can see—"Wow! If I changed this little thing, or added a paragraph here or dialogue there, it made my score go up. Now I'm excited to go back and make some more [changes]. I'm ready to edit and revise."

Another teacher spoke of the benefits this way:

They really enjoy having the opportunity to just do better, so like they'll—having them revise go back and check their scores and actually see them seeing their own progress—I feel like that was a huge benefit because, you know, going back to our traditional way of grading papers, I mean—take a paper, edit it up, mark it up, hand it back to them—but that

wasn't necessarily something that motivated them to do better...MI Write has a great increase everyone will feel motivated to do more and add more so I think that was definitely a great benefit for my students and I just like it for myself

In sum, generally teachers generally perceived that, by means of students being more motivated and putting forth greater effort, MI Write aided teachers in achieving their pedagogical objectives.

5.2.2 Independence. Teachers also perceived that MI Write helped students complete more of the writing process independently. One teacher said that with MI Write, students “take ownership of their writing.” Rather than waiting for a teacher to provide assistance, as was common with traditional instruction, teachers perceived that students were able to write, revise, and edit with greater independence. Teachers agreed that this increased independence was a desirable outcome and an objective of theirs, and that increased student independence had positive effects on teachers’ division of labor. As the following quotations reveal, teachers not only viewed MI Write as a useful instructional tool, but they also viewed it as a useful classroom management tool:

Having the students on their Chromebook listening to the lessons [in MI Write] working through a computer program instead of looking at a piece of paper with red marks all over it and just rewriting, rewriting, rewriting...it definitely makes it easier for the teacher to have the one-to-one conferences while the rest of the class is working and writing...it's also helping me manage the classroom.

Another teacher commented that:

I think it gives them feedback when I'm conferencing with kids about their writing and they're sitting here with their computer with me one-on-one. All the other kids are out there needing my attention. But now they don't need my attention because they can look to their score to tell them what to fix...That's been really helpful with me for conferencing.

Indeed, many teachers commented that students’ desire to use MI Write, coupled with increases in students’ independence, facilitated a positive classroom environment: “It’s just a peaceful time...that’s probably my most quiet moment of the day.” For these reasons, one teacher referred to MI Write as the students’ “little teacher that keeps them busy in a productive way.”

5.2.3 Complexities related to motivation. It was not always the case that teachers perceive MI Write to positively mediate students' writing motivation. Students who received low scores, despite their best effort, would often feel frustrated. As one teacher explained, "They got really frustrated with the revision because they want that good score, but they didn't know how to put in the work to get that good score." This was true even for students who received high scores:

I have an amazing writer, I could never teach her anything; you know, she's just a gifted writer, and try as she might, I would read it and give her a 100% but MI Write won't. She'll make revisions and it would lower her grade and it's just, at some point we had a conversation: "Like you're a good writer, just walk away from it."

In other cases, teachers perceived that students focused their motivation on the wrong goal, aiming at improving *scores*, rather than improving *writing*. This led students to adopt bad habits like plagiarism or adding in unnecessary details instead of "applying...the skills taught as much in class." Thus, MI Write was perceived to influence students' motivation in complex, and not always positive, ways.

5.2.4 Complexities related to independence. It also was not always the case that teachers perceived MI Write to increase student independence. Teachers reported that some students struggled to work independently through the writing process because MI Write's feedback was perceived as misaligned to the target community of users: it was not "very kid friendly," as one teacher said. Another teacher stated that "Over half the class didn't understand a lot of the words being used...it wasn't anything they could do independently." Consequently, this complicated the activity system because students would not make productive revisions and instead make basic low-level edits with the hopes that their score would improve: "[students would] just look at the number and add a period here, and then just go back and look at that number." Also, teachers reported spending a great deal of time trying to explain the feedback to students. As one teacher said, "They can't process the feedback and know what to do to make their score better without an adult sitting

with them.” Thus, when MI Write’s automated feedback was not appropriately targeted to students’ zone of proximal development, MI Write did not function as a tool that supported teachers’ pedagogical **objectives** or facilitated a productive and desirable division of labor.

5.3 Theme 3: AWE Functionality May Create New Instructional Challenges

5.3.1 Writing with style is difficult. Teachers discussed MI Write’s limitation to recognize formatting and other stylistic elements as an important aspect of writing instruction. Teachers expressed worries about being unable to fully teach the required curriculum due to these limitations. For instance, a common concern was that MI Write would remove paragraph formatting and show text as a solid block: “They can’t indent. Like some of that formatting that we teach them in the classroom on paper they can’t [apply].” Another issue pertained to specific kinds of formatting for dialogue, such as the use of italic font. One teacher explained, “Like as an example in our reading anthology there’s the italicized print and we talk about how that is the character’s thoughts, inner thoughts, but they can’t do that in MI Write when they are writing a narrative.” Although MI Write does allow students to indent and italicize their writing in the “formatting” (i.e., publishing stage) of the writing process, teachers expressed desires that this functionality be available in the drafting stage to connect instruction, practice, and assessment more clearly.

More concerning to teachers, though, was MI Write’s inability to recognize stylistic elements that added voice to students’ writing, particularly to their narrative writing. For instance, “If students used interjections like ‘WOW!’ and they write it in all caps, it tells them it’s not a real word.” Another teacher recounted the following issue pertaining to onomatopoeia:

Our first narrative was about setting and I know a lot of the kids wanted to add sound effects with what they would hear that the frogs are doing. And the sound effects MI Write would count as a non-word, so they were trying to add dialogue or sensory details and MI Write wasn’t allowing it.

Teachers shared that other issues arose when students tried to integrate specific names, characters, and animals into their writing (i.e., proper nouns): “Sometimes the dictionary knowledge-base wasn’t always the best because they would write about animals and things like that and MI Write wouldn’t recognize the word, and MI Write would be like you spelled it wrong.”

In sum, MI Write’s limitations with respect to accepting certain formatting during the drafting stage and recognizing and accepting onomatopoeia and creative or content-specific vocabulary, created a misalignment between what teachers expected of students versus what MI Write accepted, *between the rules governing the technology and those influencing teachers*. Consequently, teachers expressed concerns that MI Write may work in opposition to their *pedagogical objectives* by limiting voice and style in student writing.

5.3.2 Barriers for struggling writers. Teachers also shared concerns that MI Write’s “good-faith-effort” detection algorithms complicated instruction for struggling writers. These algorithms are designed to detect essays with very high percentages of misspelled words, too much repetition, or insufficient text and prevent such essays from being scored by the system. The intention of such algorithms is to allow only “good faith” attempts to be scored and prevent the gaming of the system (e.g., by typing a bunch of gibberish or copying and pasting a paragraph multiple times). However, teachers reported that for struggling writers, and some English learners and students with disabilities, their good faith attempts at producing high quality writing were flagged by these detection algorithms and students were denied opportunities to receive a score and feedback, *limiting the writing activity of certain members (i.e. students) within the community*.

One teacher shared:

We have some kids who wouldn’t write a lot, would write one sentence and [MI Write] wouldn’t accept it, so we would have to try to figure out something or we have to do the

spelling for them, um, if they're just doing how they know how to spell it. Because if they have multiple spelling errors it just says too many errors.

Another teacher lamented that “We had some kids that couldn't even submit their work because they were so low and they were making so many mistakes that [MI Write] wouldn't even accept their work.”

Issues such as these not only inhibited students' [activity](#) and writing motivation, but they also caused teachers to expend more labor and effort. Many teachers acknowledged that they ended up spelling so that the first draft would be accepted by the system. Other teachers, when faced with the prospect of taking dictation for individual students, simply circumvented the use of MI Write altogether during the drafting stage, relying instead on other composing tools. Several teachers reported using GoogleDocs in these instances: “Our students could type it in Google, and they could use the text to speech and then they would copy what was produced via Google and then put it into MI Write.” When the core functionality of MI Write, its automated scoring and feedback, were made inaccessible to struggling writers due to a stringent application of good faith effort detection algorithms, teachers perceived MI Write to hinder their pedagogical [objectives](#) and increase their labor.

5.4 Theme 4: AWE Implementation and Efficacy is Influenced by Broader Aspects of the Instructional Activity System

5.4.1 Rules related to time for writing instruction. The implementation of MI Write is situated within a broader activity system that includes rules regulating the amount of time that should be devoted to writing instruction. Although the district required writing instruction daily for 45min, individual schools did not always adhere to this requirement. Across schools, teachers reported master schedules that included writing instruction ranging from 45min once a week, to

20min daily, to 45min daily. As a result, there was significant variability in teachers' implementation of writing instruction and MI Write. As one teacher shared:

Our schedule is the issue. We just didn't have the time to do anything beyond like, until we got done what we had to do as a district requirement, there was nothing left to do other stuff, you know, MI Write or any other kind of publishing.

Teachers in schools with limited time for writing often expressed a desire to use MI Write to a greater extent and to learn more about its features, but they lamented the lack of time available to do so: "I feel like the biggest obstacle was just finding the time to teach writing. And that has nothing to do with MI Write." Rules dictating the allotment of instructional time for writing influenced teachers' decisions about their division of labor and what tools they used. When faced with time constraints, teachers often focused on meeting curricular requirements rather than investing time in learning to use, and teaching students to productively use, MI Write.

5.4.2 Community supports for implementation. Teachers recognized the value of community for productively implementing MI Write. Teachers spoke of the value of working with each other to learn and integrate MI Write with instruction: "My wonderful colleague helped show me things," said one teacher, echoing a common sentiment among the teachers. Indeed, teachers used the focus groups as an opportunity to learn from each other about different features of MI Write, and how to troubleshoot or integrate MI Write with instruction. This phenomenon exemplified the value teachers placed in implementing MI Write as part of a community of educators.

Teachers also recognized that implementing MI Write was an evolving process within their classroom, grade level, school, and district communities. Indeed, many teachers expressed how their implementation of MI Write gradually grew within and across years; effective implementation was recognized as a gradual and evolving process, not a one-time event. Teachers

discussed how they benefitted from having time to explore and use MI Write from one year to the next. One teacher shared: “Now that we got the hang of it, I feel a little bit, I feel 50 times better than I did last year.” A teacher from a different school echoed this sentiment: “Since we’ve used it for some time now, I’m comfortable adding prompts and figuring out how to share them amongst the school and accessing district ones.” Many teachers expressed [objectives](#) for expanding their use of MI Write the following school year either by incorporating new functions or troubleshooting issues they encountered that year.

A final community factor that teachers perceived to influence their productive implementation of AWE was student exposure to MI Write. One fifth-grade teacher happily commented that “They have been using it since third grade, so we have the advantage where they are used to logging into MI Write. I personally didn’t have to set up how to get onto MI Write because they already knew how.” Ongoing and continued use of MI Write was perceived to benefit students and teachers alike. As students became more proficient with MI Write, teachers recognized that they focused less of their time on teaching students to use MI Write and more of their time on teaching writing, [the desired focus of their labor](#).

6. Discussion

This qualitative analysis of focus groups conducted with elementary teachers who used an AWE program called MI Write identified four major themes: (1) AWE both assists and challenges teachers; (2) AWE influences students’ writing motivation and independence in complex ways; (3) AWE functionality may create new instructional challenges; and (4) AWE implementation and efficacy is influenced by broader aspects of the instructional activity system.

Interpreted considering Activity Theory, study findings indicate that AWE mediates teachers’ instructional activity in complex ways. Many ways were facilitative, helping teachers to:

(a) provide students with greater amounts of feedback and more efficiently, (b) emphasize that writing is a process requiring revision, (c) identify students' needs and prioritize their feedback, and (d) help students be more motivated and independent. In these ways, AWE appears to help teachers achieve a more desirable division of labor and more effectively achieve their pedagogical objectives.

Findings also indicate that AWE mediates teachers' instructional activity in potentially counterproductive ways. For instance, teachers perceived MI Write to (a) lead students to overly focus on the automated scores and enact maladaptive writing habits, (b) cause certain students to experience decreased writing motivation, (c) create more work for teachers when students did not understand the automated feedback, (d) create challenges arising from its standardized scoring and feedback and certain formatting features. When these phenomena were present, AWE created a less desirable division of labor, reinforced a misaligned set of assessment and writing rules, and potentially undermined teachers' pedagogical objectives.

However, the extent to which AWE mediates writing activity in facilitative or unhelpful ways is affected by additional aspects of the activity system, namely the rules and community a teacher is subject to and participates in. For instance, teachers commented that rules governing the availability of instructional time for writing significantly affected their implementation of MI Write. With scarce instructional time, it is understandable that teachers would perceive the time required to teach students to use MI Write as an up-front cost to be avoided, rather than an investment that may pay dividends later. Teachers also expressed that community supports and continued usage of AWE over time may mitigate the challenges and maximize the benefits of implementing AWE.

In sum, AWE mediates teachers' instructional activity in complex ways that may support or hinder teachers' pedagogical **objectives**. AWE is neither wholly beneficial nor wholly detrimental, neither savior nor scourge. Whether AWE acts as one or the other is dependent on complex relationships among elements of the instructional activity system, including the affordances and limitations of the AWE system, the subject (the teacher), the needs of the students, and the rules and community influencing the instructional system.

6.1 Implications

Study findings have important implications for practice, research, and the design of AWE systems. First, districts and schools should recognize that AWE is not a magic bullet, but a pedagogical tool that **mediates** the instructional environment in new and sometimes challenging ways. Moreover, districts and schools should understand that aspects of the broader instructional activity system may either constrain the affordances of AWE, such as the curriculum or a school's master schedule; or act as supports, such as the ability for teachers to work in communities of practice (see Desimone & Garet, 2015). Finally, districts and schools should recognize that using AWE to improve writing outcomes requires aligning the use of AWE with elements of the broader instructional activity system, such as scheduling, curriculum, and professional development. This recommendation is consistent with findings from analyses of the effectiveness of digital tools in other subject areas, such as mathematics and science: effect sizes on student learning are larger when teachers receive explicit training and support for using digital tools to support instruction (Hillmayr et al., 2020) and there is strong organizational support including a "shared vision for technology-enriched learning" within a school (Olszewski & Crompton, 2020, p. 7).

Second, findings suggest that research on AWE's effectiveness should place greater emphasis on documenting the instructional context. Although AWE *generally* has a positive effect

(Grimes & Warschauer, 2010; Palermo & Thomson, 2018; Palermo & Wilson, 2020; Stevenson & Phakiti, 2014; Warschauer & Grimes, 2008), findings of the present study suggest that AWE may not have the same effect for all students. Thus, future AWE research should move beyond the basic question as to *whether* AWE is effective and should instead address questions regarding *for whom* and *under what conditions* AWE is effective.

Finally, study findings provide important guidance for AWE developers. Findings illustrate how certain features that are intended to solve one problem (e.g., gaming the system) may create new problems for teachers and students (e.g., struggling writers being denied access to feedback). In addition, developers should be sensitive to the role that AWE plays in reinforcing rules, particularly those related to assessment standards given the authority with which AWE appears to be perceived by teachers and students. AWE developers should focus on teachers and their needs and evaluate how different AWE functionalities and assessment features (e.g., the use of Six Traits versus other criteria) affect instruction and provide or withhold access to certain community members, such as English learners and student with disabilities. This approach to AWE development is consistent with a human-systems engineering approach (see Roscoe et al., 2018).

6.2 Limitations

First, focus groups were conducted with elementary teachers within a single school district in the Mid-Atlantic region. While the school district was diverse and quite large (14 elementary schools), the sample was not nationally representative, but one of convenience, which may limit generalizability of findings. Second, focus group methodology relies on teacher self-report. Teachers may have felt social pressure to exaggerate or withhold certain information. Although this limitation cannot be ruled out, teachers' self-reported experiences were consistent with those reported in prior research (Grimes & Warschauer, 2008; Warschauer & Grimes, 2010) and with

behavioral measures of AWE implementation that show significant variability in usage and effect (Attali, 2004; Foltz et al., 2013; Roscoe & McNamara, 2013). Finally, the present study examined the topic of AWE through the lens of one AWE system, MI Write. AWE systems have common functionality, but systems differ with respect to scoring and feedback algorithms and their functionality (e.g., supporting peer review). Thus, future research should continue to qualitatively explore elementary teachers' experiences using other AWE systems to evaluate the generalizability of the four themes identified in the current study.

References

- Attali, Y. (2004, April). *Exploring the feedback and revision features of Criterion*. Paper presented at the National Council of Measurement in Education (NCME), San Diego, CA.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267. <https://doi.org/10.3102/0013189X07306523>
- Backfisch, I., Lachner, A., Stürmer, K., & Scheiter, K. (2021). Variability of teachers' technology integration in the classroom: A matter of utility! *Computers & Education*, 166, 104159. <https://doi.org/10.1016/j.compedu.2021.104159>
- Bhatia, V. K. (2002). Applied genre analysis: A multi-perspective model. *Ibérica, Revista de la Asociación Europea de Lenguas para Fines Específicos*, (4), 3-19. Retrieved from <http://www.aelfe.org/documents/text4-Bhatia.pdf>.
- Coe, M., Hanita, M., Nishioka, V., & Smiley, R. (2011). *An investigation of the impact of the 6+1 Trait Writing model on grade 5 student writing achievement* (NCEE 2012–4010). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Conference on College Composition and Communication. (2014). *Writing assessment: A position statement*. Retrieved from <http://www.ncte.org/cccc/resources/positions/writingassessment>.
- Desimone, L.M., & Garet, M.S. (2015). Best practices in teacher's professional development in the United States. *Psychology, Society, & Education*, 7, 252-263.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.

- Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, *43*, 960-974.
- Ericsson, P. F., & Haswell, R. J. (2006). *Machine scoring of student essays: Truth and consequences*. Logan, UT: Utah State University.
- Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: strategies for technology integration. *Educational Technology Research and Development*, *47*, 47–61.
- Foltz, P. W., Streeter, L. A., Lochbaum, K. E., & Landauer, T. K. (2013). *Implementation and applications of the Intelligent Essay Assessor*. In M. Shermis, & J. Burstein (Eds.), *Handbook of Automated Essay Evaluation* (pp. 68-88). New York, NY: Routledge.
- Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary Street®: computer support for comprehension and writing. *Journal of Educational Computing Research*, *33*, 53-80.
- Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist*, *53*(4), 258–279. DOI: 10.1080/00461520.2018.1481406
- Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *Elementary School Journal*, *115*, 523-547.
- Graham, S., McKeown, D., Kiuahara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*, 879-896.
- Graham, S., & Sandmel, K. (2011). The process writing approach: A meta-analysis. *The Journal of Educational Research*, *104*, 396-407.

- Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *Journal of Technology, Learning, and Assessment*, 8(6), 4–43.
- Herrington, A., & Moran, C. (2001). What happens when machines read our students' writing? *College English*, 63(4), 480-499.
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context specific meta-analysis. *Computers & Education*, 153. <https://doi.org/10.1016/j.compedu.2020.103897>
- Kellogg, R. T., Whiteford, A. P., & Quinlan, T. (2010). Does automated feedback help students learn to write? *Journal of Educational Computing Research*, 42(2), 173-196.
- Klobucar, A., Elliot, N., Deess, P., Rudniy, O., & Joshi, K. (2013). Automated scoring in context: Rapid assessment for placed students. *Assessing Writing*, 18, 62-84.
- Jonassen, D. H. (2000). Revisiting activity theory as a framework for designing student-centered learning environments. In D. H. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments* (pp. 89-121). Mahwah, NJ: Lawrence Erlbaum.
- Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sándor, Á., Kitto, K., Liu, M., Vijay Mogarkar, R., & Buckingham Shum, S. (2020). Acawriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, 12. Retrieved from https://www.jowr.org/articles/vol12_x/JoWR_2020_volx_nrx_Knight_et_al.pdf
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by *Grammarly*: A multiple case study. *Assessing Writing*, 44. <https://doi.org/10.1016/j.asw.2020.100450>

- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Singapore: Springer.
- Li, R., Meng, Z., Tian, M. Zhang, Z., Ni, C., & Xiao, W. (2019). Examining EFL learners' individual antecedents on the adoption of the automated writing evaluation in China. *Computer Assisted Language Learning, 32*, 784-804.
- Link, S., Mehrzad, M., & Rahimi, M. (2020). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2020.1743323>
- Little, C. W., Clark, J. C., Tani, N. E., & Connor, C. M. (2018). Improving writing skills through technology-based instruction: A meta-analysis. *Review of Education, 6*, 183-201.
- Lyst, A. M., Gabriel, G., O'Shaughnessy, T. E., Meyers, J., & Meyers, B. (2005). Social validity: Perceptions of check and connect with early literacy support. *Journal of School Psychology, 43*, 197-218.
- McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods, 45*, 499-515.
- Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: a meta-analysis of research findings. *Reading and Writing, 25*, 641-678.
- National Center for Education Statistics (2012). *The Nation's Report Card: Writing 2011* (NCES 2012-470). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.
- National Council of Teachers of English. (2013). *NCTE position statement on machine scoring*. Retrieved from http://www.ncte.org/positions/statements/machine_scoring.

National Commission on Writing for America's Families, Schools, and Colleges. (2003). *The neglected "R": The need for a writing revolution*. Iowa City, IA: The College Board

National Commission on Writing for America's Families, Schools, and Colleges. (2004).

Writing: A ticket to work...or a ticket out. A survey of business leaders. Iowa City, IA: The College Board.

National Commission on Writing for America's Families, Schools, and Colleges. (2005).

Writing: A powerful message from state government. Iowa City, IA: The College Board.

Olszewski, B., & Crompton, C. (2020). Educational technology conditions to support the development of digital age skills. *Computers & Education, 150*.

<https://doi.org/10.1016/j.compedu.2020.103849>

Page, E. B. (2003). Project Essay Grade: PEG. In M. D. Shermis, & J. C. Burstein (Eds.),

Automated essay scoring: A cross-disciplinary perspective (pp. 43-54). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Palermo, C., & Thomson, M. M. (2018). Teacher implementation of self-regulated strategy development with an automated writing evaluation system: Effects on the argumentative writing performance of middle school students. *Contemporary Educational Psychology, 54*, 255-270.

Palermo, C., & Wilson, J. (2020). Implementing automated writing evaluation in different instructional contexts: A mixed-methods study. *Journal of Writing Research 12*(1), 63-

108. <https://doi.org/10.17239/jowr-2020.12.01.04>

Parr, J. M., & Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing, 15*, 68-85.

- Perelman, L. (2014). When “the state of the art” is counting words. *Assessing Writing*, 21, 104-111.
- Persky, H. R., Daane, M. C., & Jin, Y. (2002). *The Nation's Report Card: Writing 2002*. (NCES 2003-529). National Center for Education Statistics, Institute of Education Sciences. U. S. Department for Education, Washington, D. C.
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37, 8-25.
- Roscoe, R. D., Allen, L. K., Johnson, A. C., & McNamara, D. S. (2018). *Automated writing instruction and feedback: Instructional mode, attitudes, and revising*. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2089–2093. Retrieved from <https://journals.sagepub.com/doi/10.1177/1541931218621471>
- Roscoe, R. D., Cooke, N. J., Branaghan, R. J., & Craig, S. D. (2018). Human systems engineering and educational technology. In R. D. Roscoe, S. D. Craig, & I. Douglas (Eds.), *End-user considerations in educational technology design* (pp. 1-34). Hershey, PA: IGI Global.
- Roscoe, R. D., & McNamara, D. S. (2013). Writing Pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology*, 105, 1010-1025.
- Russell, D. R. (1995). Activity theory and its implications for writing instruction. In J. Petraglia (Ed.), *Reconceiving writing, rethinking writing instruction* (pp. 51-78). Hillsdale, NJ: Erlbaum.

- Saubern, R., Urbach, D., Koehler, M., & Phillips, M. (2020). Describing increasing proficiency in teachers' knowledge of the effective use of digital technology. *Computers & Education, 147*. <https://doi.org/10.1016/j.compedu.2019.103784>
- Shermis, M. D. (2014). State-of-the-art automated essay scoring: Competition, results, and future directions from a United States demonstration. *Assessing Writing, 20*, 53-76.
- Shermis, M. D., & Burstein, J. C. (Eds.). (2013). *Handbook of automated essay evaluation*. New York, NY: Routledge.
- Stevenson, M. (2016). A critical interpretative synthesis: The integration of Automated Writing Evaluation into classroom writing instruction. *Computers and Composition, 42*, 1-16.
- Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing, 19*, 51-65.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*, 237-246.
- Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Education Technology Research and Development, 65*, 555-575.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, E. L., Matsumura, L. C., Correnti, R., Litman, D., Zhang, H., Howe, E., Magooda, A., & Quintana, R. (2020). *eRevis(ing)*: Students' revision of text evidence use in an automated writing evaluation system. *Assessing Writing, 44*.
<https://doi.org/10.1016/j.asw.2020.100449>

- Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies*, 3, 22–36.
- Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 16, 165-184.
- Whipp, J. L., Eckman, E. W., & van den Kieboom, L. (2005). Using sociocultural theory to guide teacher use and integration of instructional technology in two professional development schools. *Journal of Computing in Teacher Education*, 22, 37-43.
- Wilson, J. (2017). Associated effects of automated essay evaluation software on growth in writing quality for students with and without disabilities. *Reading and Writing*, 30, 691-718. <https://doi.org/10.1007/s11145-016-9695-z>
- Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English language arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers and Education*, 100, 94-109. <https://doi.org/10.1016/j.compedu.2016.05.004>
- Wilson, J., Huang, Y., Palermo, C., Beard, G., & MacArthur, C. A. (2021). Automated feedback and automated scoring in the elementary grades: Usage, attitudes, and associations with writing outcomes in a districtwide implementation of MI Write. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-020-00236-w>
- Wilson, J., Olinghouse N. G., & Andrada, G. N. (2014). Does automated feedback improve writing quality? *Learning Disabilities: A Contemporary Journal*, 12, 93-118. <https://eric.ed.gov/?id=EJ1039856>
- Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, 58, 87-125. <https://doi.org/10.1177%2F0735633119830764>

Zhen, X., Kim, H., Lai, W., & Hwang, G. (2020). Cognitive regulations in ICT-supported flipped classroom interactions: An activity theory perspective. *British Journal of Educational Technology*, 51. <https://doi-org.udel.idm.oclc.org/10.1111/bjet.12763>

Zhu, M., Liu, O. L., & Lee, H. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education*, 143. <https://doi.org/10.1016/j.compedu.2019.103668>

Appendix – Focus Group Protocol

1. How long is your writing block? Your writing block is the time when you teach writing. We define teaching writing as instructing students in the craft and structure of writing, not using writing to support learning.
2. Did you use MI Write this year with your students? How much would you say you used it?
3. Here are a few specific questions about how you used MI Write this year.
 - a. How often did your students go through the entire writing process, starting with their first draft in MI Write?
 - b. How often did they write their first draft by hand and then use MI Write to revise, edit, and publish it?
 - c. How often did they use MI Write just for editing and publishing, for essay clean-up and to get a score?
4. What were the main barriers to you and your students to using MI Write?
 - a. Queries to follow-up with if these don't come up in teachers' responses:
 - i. Keyboarding skills
 - ii. Reading skills – maybe the feedback is too complex and they can't read it
 - iii. Lack of alignment of the MI Write scoring system with district system
 - iv. Not enough time in the day
 - v. Lack of teacher training in using MI Write
5. What procedures and routines did you develop for integrating MI Write into writing instruction this year?
6. Did you feel like MI Write integrated well with your writing instruction? Why or why not?
 - a. Queries:
 - i. Did you use any of MI Write's planning tools? If so, how did they fit with other planning instruction?
 - ii. Did you teach students to evaluate their writing? What criteria or rubrics did you use? And how did it work?
 - iii. How did you integrate the MI Write scores with your grading?
7. What are the benefits/effects of using MI Write for your students?
 - a. Queries:
 - i. What do you think the effect of MI Write is on students' writing motivation and self-confidence as writers?
 - ii. What do you think the effect of MI Write is on students' writing performance? Is it helping them become better writers?
8. What are the benefits/effects of using MI Write for you as a teacher?
 - a. Queries:
 - i. Were you able to assign more writing?
 - ii. Assist you with your grading?
 - iii. Allow you to provide more feedback?
 - iv. Allow you to focus on higher-level writing concerns?
 - v. Assign more writing?
 - vi. Encourage revisions?
9. What were the limitations of MI Write as you used it this year?
 - a. Queries
 - i. Did you think the scoring was accurate?
 - ii. Did students understand the scores and feedback?
 - iii. Were there technical issues within MI Write?
10. What kinds of supports would be helpful to you and your students next year when you use MI Write?
11. Anything else you would like to talk about regarding MI Write or using MI Write during writing instruction?