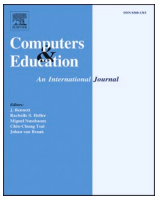
Computers & Education 167 (2021) 104175

Contents lists available at ScienceDirect

Computers & Education

journal homepage: http://www.elsevier.com/locate/compedu



Authenticity, engagement, and performance in online high school courses: Insights from micro-interactional data

Jennifer Darling-Aduana

Georgia State University, College of Education & Human Development, 30 Pryor Street SW, Atlanta, GA, 30303, USA

ARTICLE INFO

Keywords:
Distance education and online learning
Pedagogical issues
Secondary education

ABSTRACT

Increased reliance on online course-taking in high school may have a negative impact on educational opportunity gaps by exposing students belonging to marginalized subgroups to a *pedagogy of poverty*. Employing a mixed method design, I documented the prevalence and form of authentic work in asynchronous, vendor-developed online courses used primarily for credit-recovery in a large, urban school district serving a predominately low-income, minoritized student body. Rather than facilitating higher-order thinking and providing real-world relevance, lessons most often required students to recite information. However, through qualitative analysis, I identified that some types of authentic work were feasible within the asynchronous online course structure through the integration of student-directed research, writing-based tasks, and interactive assignments. The identification of only small magnitude associations with achievement indicates that authentic work should be viewed less in terms of closing the achievement gap and more as necessary to remedy the opportunity gap. Findings identify instructional practices likely to improve the authenticity, rigor, and relevance of students' educational experiences online.

1. Introduction

School districts in the United States are increasingly relying on for-profit companies to provide educational services (Boninger et al., 2017; Burch & Good, 2014). One such strategy, online education, has been adopted rapidly with approximately 14 percent of public school students at the secondary level enrolled in online courses prior to the current COVID-19 crisis (Gemin et al., 2015). In many school districts, the students targeted for online courses are disproportionately students at risk of course failure and those belonging to lower-income and minoritized subgroups (Heinrich et al., 2019; Powell et al., 2015). Yet, the often dry, direct-lecture heavy instruction observed in the most popular online course systems appears designed to meet seat-time requirements without supporting deep, or even sometimes surface-level, learning (Heinrich et al., 2019). This disparate educational access to quality student learning experiences has the potential to perpetuate and even magnify existing educational opportunity gaps (Ahn & McEachin, 2017; Heinrich et al., 2019; Heppen et al., 2017). To minimize negative equity implications, a greater understanding of the underlying processes and student experiences within these programs is essential, as problematic content can impact millions of students. Yet, for similar reasons, advances in online curriculum and instruction could contribute to vast improvements in learning opportunities, particularly as an increasingly large proportion of students are likely to receive coursework virtually in the coming school years (McDonald, 2020).

E-mail address: jdarlingaduana@gsu.edu.

https://doi.org/10.1016/j.compedu.2021.104175
Received 16 August 2020; Received in revised form 29 January 2021; Accepted 23 February 2021
Available online 27 February 2021
0360-1315/Published by Elsevier Ltd.

Towards these ends, the purpose of this study is to examine how students interact with and learn from asynchronous online course content and instructional tasks with a focus on implications for students belonging to marginalized subgroups and at risk of not graduating from high school. I focused my examination on authentic work due to concerns raised by Haberman (2010) and others that students from predominately low-income backgrounds and assigned to lower academic tracks are exposed to a *pedagogy of poverty* focused primarily on remembering and recitation versus critical thinking and application (Anyon, 1980; Darling-Hammond, 2001; Golann, 2015; Howard, 2019; Oakes, 2005) with negative ramifications for student engagement, learning, and well-being (Apple, 2019; Au, 2012). Specifically, I defined and measured authentic online work for this study as technology-supported instructional activities that facilitate student-directed learning by encouraging higher-order thinking and real-world application (Marks, 2000; Newmann et al., 1996).

Applying a sequential mixed method design (Creswell & Clark, 2018), I examined the following research questions.

1. What is the prevalence and form of authentic work in popular online courses designed by one of the largest online course vendors in the United States?
2. To what extent do students demonstrate different levels of engaged behaviors and achievement when working on online course lessons that ask them to engage in authentic work?
3. By how much do students respond differentially to authentic work based on prior year achievement?

To answer these research questions, I quantified the extent to which course videos, assignments, practice problems, and assessments facilitated authentic work across ELA, math, social studies, sciences, and elective courses. I accomplished this using a research-based rubric that asked raters to evaluate the extent to which lessons provided opportunities for higher-order thinking and real-world relevance, two primary components of authentic work (Marks, 2000; Newmann et al., 1996). Rubric ratings were then linked to data collected from every login made by 4791 students attending 45 high schools and alternative educational settings serving high school students in a large, urban district. Data were limited to students who enrolled in online courses during the 2016-17 and 2017-18 school years. Credit recovery was the most common reason for online course enrollment in the district studied. I operationalized engaged behaviors as the lesson-level active and idle time logged in the online courses system and measured achievement using the end-of-lesson assessment score. These data allowed me to examine micro-interactional processes between students and the online course interface developed by a vendor who has provided courses to over 16,000 schools, including eight of the 10 largest districts in the United States.

This is the first study of which I am aware to systematically examine the content of asynchronous online secondary courses in this manner. As such, the documentation of modal lesson authenticity, rigor, and relevance represents a contribution. Further, findings deepen understanding of the factors influencing student behaviors, learning, and outcomes in an instructional medium with rapid adoption across the United States that disproportionately influences the educational experiences of students belonging to marginalized groups and attending lower-resourced schools (Darling-Aduana et al., 2020; Powell et al., 2015). In doing so, I extend current debates on the equity implications, merits, and challenges of facilitating authentic work into the online instructional space (i.e., Kirschner et al., 2006). Expanding understanding of these processes within online platforms is of even greater importance in the current COVID-19 era of remote learning. By examining students' experiences in online courses at this level of detail, I also aim to identify potential levers for improving curricular content, design, and delivery in online classroom settings. Attention to these practices in both online and face-to-face classrooms can support the achievement and well-being of high school students who have been marginalized and alienated by current institutional structures (Au, 2012; Dee & Penner, 2017; González, Moll, & Amanti, 2006; Gutstein, 2016; Yair, 2000).

2. Authentic work in face-to-face and online settings

Authentic work allows students to create their own meaning and connect course material to their lives (Apple, 2019; Newmann et al., 1996). Students with access to these enriching experiences demonstrate improved academic engagement and performance, including higher rates of critical thinking, thoughtfulness, and in-depth knowledge (Apple, 2019; Dee & Penner, 2017; Marks, 2000; Newmann et al., 1996). Specific to technology-based tasks, students learn more when completing activities that prompt them to reflect, think critically, provide explanations for work, and generate (versus receiving) knowledge (Cavus et al., 2007; Crippen & Earl, 2007; Dinov et al., 2008; Hong et al., 2020; Means et al., 2009; Wang et al., 2020). There remains a need to establish the prevalence and form of authentic work in online courses, including the extent to which authentic work is feasible in fully asynchronous settings, and how students completing online courses predominately for credit-recovery respond to authentic work when present.

Despite the academic and psychological benefits of authentic work, access is provided disproportionately to students from more affluent backgrounds, while students from less affluent backgrounds are more often taught in a manner that emphasizes rote memorization and respect for authority (Anyon, 1980; Darling-Hammond, 2001; Golann, 2015; Howard, 2019; Oakes, 2005). Thus, students from less affluent backgrounds are often subjected to what Haberman (2010) termed a *pedagogy of poverty*, which includes teachers "giving information, asking questions, giving directions, making assignments, monitoring seatwork, reviewing assignments, giving tests, reviewing tests, assigning homework, reviewing homework, settling disputes, punishing noncompliance, marking papers, and giving grades" (p. 82). The result of these educational processes is often a sense of alienation among students, both from school and society more broadly (Au, 2012). Although these patterns appear across schools, they are also replicated within schools, where students enrolled in more advanced courses and with higher prior levels of achievement tend to be exposed to more authentic work (Gamoran & Nystrand, 1992; Oakes, 2005). As students often demonstrate initial resistance when exposed to authentic work due to

less defined metrics for success and the need for new skill acquisition (Gamoran & Nystrand, 1992; Hoffman & Ritchie, 1997), initial attempts at increasing authentic work may appear to reinforce beliefs that certain students may not benefit from this type of instruction (i.e., Zheng et al., 2020).

In fact, some researchers caution that an emphasis on authentic work may inadvertently harm students belonging to marginalized subgroups. Kirschner et al. (2006) argued that students are only likely to successfully complete the more autonomous, open-ended assignments that facilitate authentic work with sufficiently high prior knowledge and self-regulatory skills. Up until that point, students benefit from more structured, direct-instruction (Kirschner et al., 2006; Roll et al., 2018). Researchers also have been unable to identify a relationship between achievement and techniques associated with authentic work, such as inquiry-based teaching and intellectual demand (Jerrim et al., 2019; Joyce et al., 2018). The lack of an identifiable effect with achievement may be because inquiry-based activities are more likely to be associated with shorter-term attitudinal shifts that take time to translate into knowledge gains (Roll et al., 2018). Nonetheless, these studies raise concerns that increasing expectations of students to complete authentic work may maintain rather than minimize existing educational disparities.

2.1. Hypothesis

Despite the importance of authentic work and the prevalence of online, high school course-taking, there is a dearth of information on the micro-interactional processes between students and the online course content that form the building blocks of knowledge acquisition in online instructional settings. This paper contributes to understanding of how students shift their behavioral patterns based on exposure to online curricular material and instructional tasks that reflect various levels of authenticity. Based on prior literature, I expected to observe higher rates of behavioral engagement, captured through an increase in active time and decrease in idle time logged in the online course-taking system, when students were exposed to higher rates of authentic work. Increased behavioral engagement might translate into higher achievement if students had prior experience with authentic work that allowed them to successfully complete the more intellectually challenging tasks (Gamoran & Nystrand, 1992; Hoffman & Ritchie, 1997; Kirschner et al., 2006). Thus, I anticipated observing heterogeneous associations with student performance conditional on students' prior year GPA, which I used as a proxy for prior exposure to authentic work based on the work of Oakes (2005) and others on tracking. More specifically, I hypothesized that students who were more academically successful in the past, as measured by their cumulative GPA, would be better able to translate access to authentic work into improved learning due to greater familiarity and comfort with authentic tasks.

3. Methods

The study incorporated an exploratory sequential mixed method design (Creswell & Clark, 2018), whereby initial, inductive qualitative analysis informed subsequent quantitative analysis, after which each strand was integrated at the interpretation-level and used to clarify, expand, and assess the validity of emergent findings (Creswell & Clark, 2018). I used a research-based protocol to code the authenticity of approximately 200 h of online course content. A descriptive analysis of the qualitative themes and patterns

Table 1
Student demographic and academic characteristics by online course-taking status (2016-17 through 2017-18 school years).

	All High School Students	All Online Course-Taker	Analytic Sample
Female	0.494	0.455	0.440
	(0.500)	(0.498)	(0.496)
Black	0.612	0.683	0.707
	(0.487)	(0.465)	(0.455)
Hispanic	0.211	0.212	0.204
	(0.408)	(0.409)	(0.403)
White	0.098	0.069	0.060
	(0.298)	(0.253)	(0.237)
English Language Learner (ELL)	0.157	0.122	0.114
	(0.364)	(0.328)	(0.317)
Free/Reduced Price Lunch Eligible (FRL)	0.757	0.795	0.814
	(0.429)	(0.403)	(0.389)
Special Education Eligible (SPED)	0.226	0.235	0.271
	(0.418)	(0.424)	(0.444)
Prior Year GPA	1.926	1.444	1.273
	(1.046)	(0.842)	(0.806)
Failed Course(s) in Prior Year	0.591	0.835	0.873
	(0.492)	(0.371)	(0.333)
Percent Absent	0.222	0.300	0.338
	(0.231)	0.239	(0.244)
Mean Pre-test Score (if pre-tested)	–	56.028	57.779
		(30.203)	(30.340)
Number of Student-Year Observations	26,428	7110	5337

Standard deviations in parenthesis.

observed within lessons was then corroborated and supplemented with statistical analysis that leveraged student-by-course fixed effects to identify within-student variation in performance and behaviors when exposed to different levels of higher-order thinking and real-world relevance within the same course. This method allowed for the exclusion of endogenous variation in student achievement and engagement associated with student and course-specific information that remained constant during the semester when students enrolled in the course. Additional information on the data, setting, and empirical analysis is provided in the following sections.

3.1. Data and setting

The quantitative component of this study relied on administrative data provided by a large, urban school district in the Midwest of the United States. Specifically, the school district is classified as *urban emergent* indicating that it is located in a large (but not major) city that serves a predominately low-income, minoritized student population and operates with fiscal constraints and resource scarcity due to systemic funding and educational inequities (Milner, 2012). The study focused on data collected from students attending the 45 high schools and alternative educational settings serving high school students within the district. As shown in Table 1, 76 percent of high school students in the district qualified for free or reduced-price lunch, 61 percent identified as Black, and 21 percent identified as Hispanic during the study period. Student achievement and sociodemographic data were provided for the 2016-17 and 2017-18 school years for all ninth through twelfth-grade students. Among students enrolled in online courses, I also had access to information on course-taking behaviors including how many sessions the students logged and any assessment scores associated with that login. These course-taking data were collected by the online course vendor, who granted permission within their contract with the district for its data to be shared with our research team.

The research team collected additional data by reviewing and coding lessons for the presence of curricular content and instructional activities that facilitated higher-order thinking and real-world relevance according to the Online Authentic Work Rubric (see Appendix A). We accessed the course videos, assignments, and assessments associated with each lesson using a district-provided administrative login to the online course system, which allowed us to watch, read, download, and interact with the same content students received as part of each online lesson. Rubric ratings were then entered in Qualtrics.

I limited this data collection, and thus the subsequent analysis, to the 10 courses in which the most students enrolled over the two-year study period based on course-taking data provided by the online course vendor. In total, these courses represented 60 percent of all online course enrollments in the district during the study period. Courses included algebra 1, career planning and development, citizenship, ninth-grade English/language arts (ELA), personal finance, healthy living, geometry, U.S. history, world history, and physical science. The analytic sample limited to these ten courses contained 134,272 student-lesson observations, representing 4791 unique students (with 5337 student-by-year observations). As shown in Table 1, student characteristics in the analytic sample mirrored those observed in the larger population of online course-takers. However, I identified lower rates of prior academic performance and attendance among online course-takers compared to the general student population, which was expected since most students enrolled in online courses within the district to retake a course they previously failed.

The Online Authentic Work Rubric used for data collection included two subscales. The higher-order thinking subscale was designed to measure the extent to which students were asked to think deeply and critically about course content, often requiring students to generate new knowledge. The real-world relevance subscale was created to identify the extent to which course content was presented in an applied context or had the potential to be applicable to students' lives, interests, and/or aspirations. I trained three raters using the rubric. I also assigned a primary rater for each course who rated every lesson in the course. Others coded a few lessons from that course to establish interrater reliability, but only the primary rater's scores were used for analysis to minimize concerns regarding variability in ratings associated with the rater. In total, 48 lessons (12 percent) were coded by more than one rater. Raters assigned a value within one-point of each other on the four-point scale in 92 percent of cases on the training sample and 94 percent of other lessons that were double coded to determine continued consistency in ratings. The average Cohen's kappa when excluding lessons used for training and calibration was 0.35. When including the training sample, I estimated an average Cohen's kappa of 0.22. Both estimates indicate a moderate level of interrater reliability (McHugh, 2012).

After coding, I use item response theory (IRT) rating scale models to place the extent to which higher-order thinking and real-world relevance were present in each lesson on standardized, continuous scales. The Cronbach's alpha for the higher-order thinking scale was 0.82, while the Cronbach's alpha for the real-world relevance scale was 0.74. Conventions in the social sciences identified the internal consistency of the real-world relevance scale as acceptable and the internal consistency of the higher-order thinking scale as good (DeVellis, 2016).

Each course contained anywhere from 20 to 60 lessons with the typical lesson containing an approximately 20-min lecture where students were presented with slides that were read and expounded upon by an instructor asynchronously. Students saw this prerecorded instructor, who was hired and trained by the online course vendor, on the top right-hand side of the screen. The content was also developed entirely by the course provider, although the asynchronous instructors had some leeway to add examples or share relevant personal experiences that were then approved by the vendor prior to distribution.

Lecture slides often contained bullet-pointed information. Other times the slides provided interactive tools, such as a statistics calculator, or showed instructors writing and talking through a practice math problem. After the lecture, students completed an activity, most often multiple choice or true/false practice problems. Other times, students completed a worksheet that required the review and analysis of outside resources or the answering of open-ended short answer questions. Every lesson ended with an assessment of student knowledge acquisition, which contained entirely, or predominately, forced-response questions.

3.2. Descriptive analysis

To demonstrate the prevalence and form of authentic work in the online courses examined, I presented in section 4.1 descriptive statistics of rubric findings to illustrate course content and instructional delivery. All quantitative analysis described in this and subsequent sections were conducting using STATA. I systematically examined patterns in lessons demonstrating high, medium, and low rates of higher-order thinking and real-world relevance along the dimensions summarized in the qualitative coding tree using NVivo (see Appendix B). I supplemented this analysis with vignettes of course lessons to provide context, clarify classroom practices, and assist in the interpretation of the higher-order thinking and real-world relevance scales. I ensured the reliability and validity of these qualitative findings by establishing saturation, triangulating findings across courses, and searching for alternative interpretations (Huberman & Miles, 2002).

3.3. Statistical analysis

Next, I used a student-by-course (α_{ic}) fixed effects model to examine whether exposure to higher-order thinking or real-world relevance were separately or jointly associated with variation in student course-taking behaviors and achievement. When estimating OLS regression without fixed effects, selection bias would occur if students more likely to perform higher (or be more engaged) were systematically more likely to be assigned to courses with higher (or lower) rates of authentic instruction. The student-by-course fixed effect strategy allowed for the removal of all variation that was constant within a student and course during the semester that a student completed that course. This means that attributes such as a students' innate ability, prior year course failure, and general orientations to learning no longer bias estimates as they cannot or are unlikely to change over a single semester (Wooldridge, 2013).

I estimated the equation leveraging variation across lessons (l), as specified in the following equation, for each student (i) and course (c).

$$y_{icl} = \beta_0 + \beta_1 authentic\_work_{cl} + \alpha_{ic} + \boldsymbol{X_{icl}\beta} + \boldsymbol{A_{cl}\beta} + \delta_{cl} + \varepsilon_1 \quad (1)$$

Dependent variables (y_{icl}) included the students' first score on the end-of-lesson assessment as well as the natural logarithm of active and idle time. The average score earned on students' first end-of-lesson assessment attempt was 85 percent with scores ranging from zero to 100 percent. These assessments consistent entirely, or predominately, of multiple choice questions and were scored out of 100 based on the percent correct. Students spent a median of 2.19 active hours and a median of 0.39 idle hours per lesson engaged with the course system, excluding time devoted to watching the lecture video. Active time measured the time logged into a course where the student was completing course content as measured by the online course system. For instance, progressing through assessment questions or completing a virtual lab were considered active engagement, whereas navigating through already completed lessons or viewing screens not contributing to course progress would be logged as idle. Thus, active time, versus the total course duration, provides an imperfect but nevertheless helpful measure of students' behavioral engagement.

I ran three models for each dependent variable, one each where I inputted the higher-order thinking and real-world relevance subscales individually as the main treatment variable ($authentic\_work_{cl}$) and then one model where I inputted both treatment variables together. When examining the role of higher-order thinking and real-world relevance in the same equation, I also included an interaction between the two terms. The interaction represented differences in students' responses that only occurred when both higher-order thinking and real-world relevance were present at higher (or lower) levels in the same lesson.

I also controlled for student-by-lesson level ($\boldsymbol{X_{icl}}$) covariates including whether the student took the pretest for a given lesson and, if they did so, his or her pretest score. This accounted for prior knowledge students might have on lesson content, where available, to more fully isolate learning due specifically to the lesson being examined. When predicting active and idle time, I also included a variable indicating whether the student retook the lesson, since it was not possible to isolate the active time spent on only the first time a student completed the lesson. At the lesson level ($\boldsymbol{A_{cl}}$), I controlled for the total time raters estimated it would take to complete a given lesson and the order the lesson appeared in the course. These measures controlled for differences in base course length (without controlling for differences in the types of activities students were asked to do during the lesson) and any differences in student achievement or behavior associated with increasing familiarity with the lesson format as they progressed through the course. I also used teacher fixed effects (δ_{cl}) to condition on differences in the online teacher quality identified through descriptive analyses. Information on which teacher taught each lesson was collected at the same time raters watched lesson lectures to evaluate the level of authentic work using the Online Authentic Work Rubric. The inclusion of teacher fixed effects accounted for differences in instructional delivery when different teachers taught different lessons in the same course to better isolate student responses to instructional tasks. All models used student-by-course level, robust, clustered standard errors to account for correlated error terms.

Next, I examined potential heterogeneous treatment effects by splitting the dataset into four quantiles based on students' prior year GPA and running the model separately for each quantile. This allowed me to determine by how much students responded differentially to authentic work based on prior achievement based on the assumption (and prior research conducted in this setting) that students in these different quantiles likely varied in their engagement patterns (i.e., Heinrich et al., 2019; Darling-Aduana, 2019). For instance, students with higher GPAs were more likely to complete course content in longer sessions, log a higher-proportion of active-to-idle time, and log sessions outside of the school day more often than students with lower GPAs (Heinrich et al., 2019). I also wanted to examine variation by prior year GPA as a proxy for the likely level of authenticity students were exposed to in prior year courses, based on research that students enrolled in more advanced courses and with higher levels of achievement were more likely to be exposed to authentic work (Gamoran & Nystrand, 1992; Oakes, 2005). Refer to Appendix C for a description of additional sensitivity tests

conducted to examine the robustness of model specifications and the extent to which model assumptions were met.

4. Results

4.1. Authentic work in online courses

I summarized the proportion of online lessons that asked students to engage in each level of Bloom's taxonomy (Bloom et al., 1984) in Fig. 1. Tasks generally increase (from left to right) in the extent to which they require students to engage in higher-order thinking. Bloom's taxonomy is meant to assist teachers in designing lessons, where an ideal lesson would include tasks across the entire taxonomy. This allows tasks requiring greater higher-order thinking to be scaffolded on more direct forms of instruction. The typical online lesson required students to recite lecture information, demonstrate understanding by answering problems, and apply what they learned to new contexts. About half of all lessons (54 percent) required students to think critically (labeled *think*). Generally, application is considered a higher-level of thinking. However, in the case of most online lessons observed, the application was surface-level and thus did not always require critical thinking. For instance, students might be asked to apply new mathematical knowledge to a word problem. While the situation may be new, the process might not be, limiting the amount of critical thinking required. Fewer than one-quarter of lessons required students to evaluate, synthesize, or create. These measures are related but separate from the authentic work subscales. Since the higher-order thinking and real-world relevance scales were standardized, it is important to place subsequent findings in the context of the modal lesson to which students were exposed.

I also identified wide variation in the presence of higher-order thinking and real-world relevance within as well as across courses, as shown in Fig. 2. When comparing courses, citizenship and career development lessons were rated consistently above average on higher-order thinking, while the United States history class was rated consistently below average. The personal finance course provided a high level of higher-order thinking in many (but not all) lessons. The algebra one and career development courses were rated the lowest on the real-world scale, while the personal finance and physical science courses had some extremely highly rated lessons on real-world relevance.

The most common reason lessons were rated low on both higher-order thinking and real-world was an overemphasis on direct instruction that provided only surface-level information or focused on abstract versus applied motivations for understanding lesson content. Several lessons across courses epitomized this type of one-dimensional instruction. One lesson regarding the Abolitionist Movement presented a minimum level of information, such as listing jobs of enslaved workers and providing an overview of the Quakers, skimming over content that might make students feel uncomfortable, and not requiring them to think deeply or analyze content. An ELA lesson on how to use a variety of sentence types to add interest to writing provided only one opportunity for students to edit a passage on their own. Instead, most tasks consisted of watching a lecture or practicing each sentence skill separately by answering multiple choice questions. In a citizenship lesson, the instructor listed traits of good citizens as, "responsible, helpful, curious, honest, altruistic, humble, knowledgeable, respectful," without any acknowledgement of different interpretations or realizations of these qualities. The presentation of absolute, vague personality traits avoided addressing the complexities of citizenship, such as when embodying one ideal might contradict another, which could have encouraged higher-order thinking or provided a measure of real-world relevance.

Rated higher on at least one dimension, the following descriptions provide examples of lessons with a range of higher-order thinking and real-world relevance scores. These examples demonstrate the types of authentic work that are feasible within an asynchronous, online setting and help to differentiate between the two authentic work components. For instance, the following notes from a ninth-grade ELA lesson were rated highly on higher-order thinking but below average on real-world relevance.

This lesson took a step back from the Odyssey to cultivate skills such as using a dictionary, identifying synonyms and antonyms in a

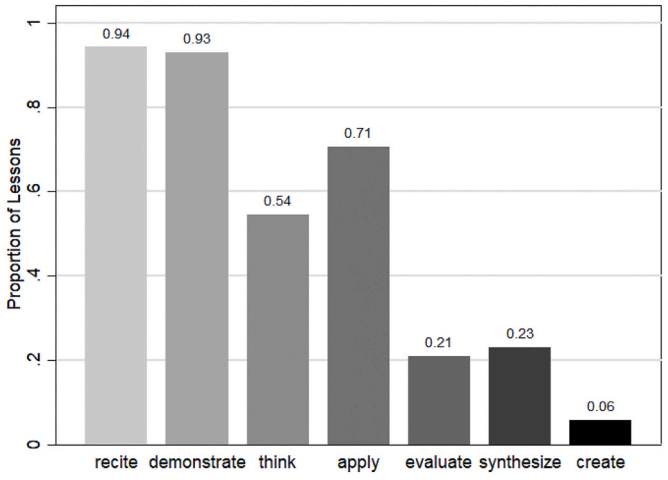


Fig. 1. Presence of tasks requiring various levels of thinking by online lesson.

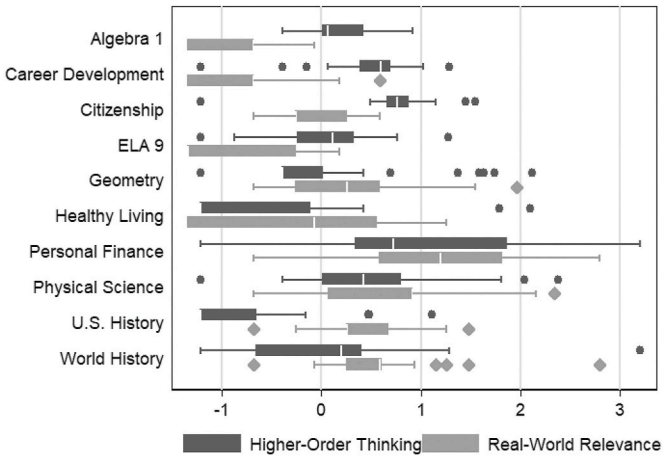


Fig. 2. Higher-order thinking and real-world relevance by course.

thesaurus, and choosing correct homophones. The instructor modeled how to do each skill, and the student practiced directly following the introduction of each skill. Questions asked students to identify facts about words (syllables, origin) from a dictionary, use context clues to choose the best definition of words with multiple definitions, and choose the best synonym for a word in context.

The above lesson introduced students to a range of skills, allowing them to select from multiple methods to answer problems with more than one right answer. These tasks required students to gather their own information, facilitating student-generated (versus instructor-directed) knowledge. However, this lesson was not rated highly on the real-world relevance scale, because the instructor failed to ground the skills learned in a meaningful context outside of academic work.

In contrast, the following U.S. History lesson provided high levels of real-world relevance with only average requirements for higher-order thinking.

During a lesson on industrialization, the instructor reviewed the concept of mass production, using machines that mechanized the shoemaking process as an example. She then talked about the dangerous conditions for workers by showing old newspaper headlines that highlighted factory accidents and presenting a graph of deaths from industrial accidents. She began the section on child labor by mentioning that child labor still exists today, and many of our consumer goods are the result of child labor in developing countries.

The instructor established a meaningful context by highlighting case studies, integrating archival documents, and linking historical content to modern life. Completion of the lesson did not require students to create work product with meaning outside an academic context, but there were opportunities for extended writing about social problems that reinforced the meaningful application of the content introduced. Despite these strengths, the lesson did not require much higher-order thinking. Instead, students could earn credit for the lesson by only remembering and reciting lecture content.

While there were examples of lessons rated favorably in only one of the two authentic work components, the lessons rated most highly in one area tended to be rated equally high in the other. These patterns provide some suggestive evidence that integrating real-world relevance might help facilitate higher-order thinking and vice versa. Many lessons in the personal finance course exemplified this intersection. For instance, in a lesson that received the highest possible score on both higher-order thinking and real-world relevance, the instructor outlined how to research careers.

The instructor provided detail about where to research jobs, including counselors, libraries, and the US Bureau of Labor Statistics. In choosing a career, he said you should know about employee benefits, comparing job offers, and lifetime income. A practice question asked students to identify Julio's best job offer out of three based on location, salary, benefits, and average monthly rent. In the lesson summary, the instructor compared the career of a doctor and mechanic, considering interests and skills, education, job outlook and income, and potential lifetime income. The assignment asked students to research two careers of their choice and show the results in a multimedia presentation.

This lesson was notable for the open-ended nature of assignments, the integration of a research-based activity that allowed student-generated knowledge, and the emphasis on multiple "right" choices depending on a student's situation and preferences. These attributes required higher-order thinking to complete the lesson successfully, but the same elements – such as the student-directed nature of the assignment and the element of choice – also personalized content in a way that enhanced real-world relevance.

A lesson in the World History course also earned a perfect score on both scales but accomplished this using writing.

The instructor asked students to write an informative essay on any turning point in history. She split the lesson up into three parts: research, draft, and revise. The next section walked through the prewriting process: identify a topic, research, and organize an outline, before going over examples of reputable sources. She then described how to create an outline in three steps (introduction, body, conclusion). She showed an example of a strong conclusion and then went over the revision process, looking at two strategies: a clear topic and effective language.

This lesson shared many of the strengths identified in the personal finance lesson presented above. Students were required to complete an intensive, multi-step, student-directed activity on a topic of the students' choosing. The World History lesson further

supported higher-order thinking by scaffolding expectations, new skill introduction, and review into the lecture component to improve the probability that students could complete what might be a challenging task in a way that reinforced the skills and knowledge acquisition processes the lesson was designed to teach.

This last example of a lesson rated highly on both dimensions describes a virtual lab that students were expected to complete as part of the Physical Science course.

The instructor showed various cabbage juice solutions and explained that the goal of the lesson was to investigate pH. A student guide for the lab was provided, working through 18 detailed steps of the procedure and providing blank data tables for students to fill in. The instructor then introduced serial dilutions and how to use a pH meter. The virtual lab consisted of dragging materials into beakers on the screen and recording the results shown digitally. Students measured the pH of different substances and then common household items. Finally, students completed a detailed lab report to illustrate their conclusions. The rubric and guide for writing this were detailed and structured.

The three types of assignments highlighted as facilitating high ratings on both scales (research, writing, and interactive activities) were common components in the most authentic lessons. As highlighted when discussing the research and writing-based tasks, the virtual lab also required interactivity, application, and critical thinking. Specifically, the emphasis on student-generated knowledge through data collection and analysis, instead of solely remembering and reciting information from an instructor-driven lecture, provided the framework for higher-order thinking and a meaningful context with real-world applicability.

4.2. Students' learning and behavioral responses to authentic online work

Having established what opportunities for authentic work looked like in modal and exceptional online lessons, I next examined how students performed differently in response to the level of higher-order thinking and real-world relevance observed by exploring variation across lessons within the same course. When completing the end-of-lesson assessments of lessons that were one standard deviation higher on the higher-order thinking scale, students scored one percent lower than they scored on other assessments within the same course, as seen in the first column of Table 2. This corresponded to no change in the active time logged and a four percent increase in the amount of idle time logged for each one standard deviation more higher-order thinking observed. For a student who logged the median number of idle hours on a lesson (0.39), a four percent increase in idle time would increase the total idle time to 0.47 h. These results remained consistent when controlling for real-world relevance, shown in column three.

However, associations between real-world relevance and student outcomes and behaviors shifted when accounting for the level of high-order thinking present. Without controlling for higher-order thinking, students exposed to one standard deviation more real-world relevance within a given course scored no higher on the end-of-lesson assessment and logged no additional active or idle time. Conditioning on the presence of higher-order thinking, as shown in column three, students scored half a percent higher on the end-of-lesson assessment. Estimates were robust to alternative specifications (refer to Appendix C for more information).[1]

Prior research indicated that students would likely respond differently to authentic work based on prior academic experiences. I was also concerned that assignment to online courses might differentially disadvantage students with lower rates of prior academic achievement who were less likely to have developed the prerequisite self-regulatory skills and academic knowledge needed to fully access and learn in an online instructional environment, particularly when asked to engage in authentic tasks. Using the model that included higher-order thinking, real-world relevance, and an interaction between the two variables, I examined estimates by prior GPA quantiles to represent prior academic experiences. This examination showed relatively consistent associations between access to authentic work and end-of-lesson assessment scores, as shown in the top half of Table 3.

When examining time logged, students in the highest quartile of prior year GPAs appeared less likely to log additional idle time and no more likely to log active time when exposed to higher rates of higher-order thinking. This might be because students with higher prior academic performance were more accustomed to and more likely to know how to successfully complete assignments requiring higher-order thinking. I also observed an outlier in the percent of active time students logged among students exposed to higher rates of higher-order thinking among students with prior year GPAs in the second quartile. Students in this quartile may not have developed the self-regulation to force themselves to spend more time on assignments requiring higher-order thinking compared to students with higher prior year GPAs. The lack of a similar drop in active time logged among students in the first quartile complicates the interpretation of this coefficient. However, descriptive statistics showed a u-shaped relationship between the amount of active time logged and students' prior year GPA, with students with extremely low and extremely high prior year GPAs logging the least amount of active time. Likely students with higher rates of prior academic achievement completed the courses more quickly because of better preparation, while students with the lowest rates of prior academic achievement might be more likely to disengage in an environment with little external monitoring. Thus, the lack of a negative coefficient on higher-order thinking when predicting active time logged among students in the first quartile might be a result of low overall engagement, while students in the second quartile might be particularly susceptible to lesson differences in authentic work. This could explain why students in the second quartile were more likely to

[1] Most estimates were qualitatively similar when examined individually by subject apart from the following exceptions. Notably, the positive association between the end-of-lesson assessment score and the interaction term persisted across subject areas. However, the association between the end-of-lesson assessment score and real-world relevance varied by subject, with a significant positive association observed in math and a significant negative association observed in social studies and science. When examining active time by subject, I identified a statistically significant positive association with higher-order thinking in math only and a statistically significant positive association with real-world relevance in science only. Results are available upon request.

Table 2
Lesson outcomes as a function of variation in higher-order thinking and real-world relevance.

	(1)	(2)	(3)
	Thinking	Relevance	Interaction
Dependent Variable: First Score (0–100%) (N = 128,463)			
Higher-Order Thinking	−0.007***		−0.010***
	(0.001)		(0.001)
Real-World Relevance		−0.000	0.004***
		(0.001)	(0.001)
Higher-Order Thinking			0.001*
* Real-World Relevance			(0.001)
Dependent Variable: Log of Active Time (N = 134,272)			
Higher-Order Thinking	−0.009		−0.029
	(0.018)		(0.022)
Real-World Relevance		0.022	0.032
		(0.020)	(0.023)
Higher-Order Thinking			0.011
* Real-World Relevance			(0.013)
Dependent Variable: Log of Idle Time (N = 134,272)			
Higher-Order Thinking	0.032**		0.038**
	(0.013)		(0.016)
Real-World Relevance		0.009	−0.008
		(0.014)	(0.016)
Higher-Order Thinking			−0.005
* Real-World Relevance			(0.010)

Each column within each dependent variable section represents a different model.

Table 3
Examination by prior year GPA.

	Lowest Quartile			Highest Quartile
	Q1	Q2	Q3	Q4
Dependent Variable: First Score (0–100%)				
Higher-Order Thinking	−0.008***	−0.010***	−0.008***	−0.010***
	(0.002)	(0.002)	(0.002)	(0.002)
Real-World Relevance	0.004*	0.003	0.001	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Higher-Order Thinking	0.001	0.002	0.001	−0.001
* Real-World Relevance	(0.002)	(0.001)	(0.001)	(0.001)
Dependent Variable: Log of Active Time				
Higher-Order Thinking	−0.032	−0.057*	0.020	−0.003
	(0.028)	(0.033)	(0.029)	(0.020)
Real-World Relevance	0.038	0.069**	0.029	0.018
	(0.029)	(0.034)	(0.029)	(0.020)
Higher-Order Thinking	0.029	0.014	−0.025	−0.009
* Real-World Relevance	(0.018)	(0.020)	(0.017)	(0.012)
Dependent Variable: Log of Idle Time				
Higher-Order Thinking	0.048**	0.034	0.054***	0.005
	(0.021)	(0.022)	(0.021)	(0.021)
Real-World Relevance	0.015	0.038	0.010	−0.020
	(0.022)	(0.023)	(0.021)	(0.020)
Higher-Order Thinking	−0.012	−0.014	0.010	0.023**
* Real-World Relevance	(0.014)	(0.013)	(0.012)	(0.012)
N	25,289	24,996	26,023	25,971

Each column within each dependent variable section represents a different model.

disengage (logging less active time) when the lesson required higher rates of higher-order thinking and reengage (logging more active time) when the lesson provided more real-world relevance.

4.3. Limitations and directions for future research

Claims of generalizability are strengthened by the examination of content developed by one of the largest online course vendors in the United States. That said, the use of courses by only one vendor prevents drawing conclusions regarding the level of authenticity (and subsequent student responses to levels of authenticity not observed in this study) in online courses developed by other vendors. Further, it is possible that the predominately remember and recite type, multiple choice assessment items might not have captured the greater depth of learning achieved when exposed to lessons that required more higher-order thinking. However, the examination of differences in active and idle time showing that the typical student logged more idle time but no more active time for each one standard deviation increase in higher-order thinking does not appear to support this hypothesis.

At the same time, there are limitations to the active and idle time measures. Specifically, the online course system that determined whether time logged was active or idle appeared to overestimate active time compared to observers physically in the classroom environment. Further, as many instructional tasks required by the online course system did not necessitate critical thinking, active and idle time provide only a proxy for behavioral engagement and not emotional or cognitive engagement. I was also limited in my ability to triangulate these active and idle time measures with interviews (or other types of first-person student reports), which would have helped clarify and address discrepancies in engagement identified between classroom observations and the active logged by the online course system. It is also possible that any benefits associated with exposure to authentic work may occur over time instead of within a single lesson (Roll et al., 2018), but determining that was outside the scope of this study. Lastly, I was unable to examine components of authentic work, such as communication and interaction with peers, that were not facilitated by the online course system evaluated.

Despite these limitations, this study makes an important contribution by delving into the black box of vendor-developed online curriculum and instruction and highlighting patterns in high school student engagement and learning based on exposure to authentic online work. Future research on this topic might benefit from more objective and nuanced sources such as eye tracking and mouse click data. Additionally, while this study attempted to capture measures of behavioral and cognitive engagement, emotional engagement remains a critical component of students' educational experience with implications for subsequent achievement in both online and face-to-face settings, which future studies should explore further (Marks, 2000; Sun & Rueda, 2012), potentially by leveraging some of the data collection methods highlighted above.

5. Discussion

Existing research tells us that students are more engaged and learn better when exposed to authentic work, which includes opportunities for higher-order thinking with real-world applications (Dee & Penner, 2017; Marks, 2000). Yet, students belonging to marginalized subgroups are systematically less likely to have access to this type of instruction (Darling-Hammond, 2001; Golann, 2015; Haberman, 2010; Howard, 2019; Oakes, 2005). The increased standardization and profit-driven motives of vendors developing online high school courses, which are often targeted for use among marginalized populations, raise concerns regarding the extent to which the incursion of online courses into the secondary education sector may magnify existing inequities in access to quality, authentic learning experiences (Boninger et al., 2017). This study represents the first large-scale attempt to document the prevalence and form of authentic work in popular online courses designed by one of the largest online course vendors in the United States. Subsequent analyses demonstrated differences in student engagement patterns and performance in lessons with more opportunities for authentic work within the same course in aggregate and by prior academic performance.

Consistent with concerns that online course-taking may exemplify the *pedagogy of poverty* for the predominately low-income, minoritized student population enrolled in online courses in this study, the lessons observed were most likely to require students to recite information and demonstrate understanding on multiple choice or true/false questions versus provide opportunities for authentic work (Haberman, 2010). This means that the widespread integration of these and similar online courses, particularly among schools serving marginalized populations, might exacerbate existing educational opportunity gaps conditional on the quality of educational opportunities available in alternative, face-to-face instructional settings. However, findings from this study also identify instructional practices likely to improve access to authentic work in online instructional settings including research, writing, and interactive activities (Cavus et al., 2007; Crippen & Earl, 2007; Dinov et al., 2008). The student-directed nature of these activities enables students to focus on topics that they deem personally relevant in a manner that allows for the integration of students' multiple funds of knowledge (González et al., 2006; Silseth, 2018).

While there is room to improve access to authentic work in online courses, the most transformative restructuring of learning for students requires more collaborative, interactive, customizable instructional environments not available in the most popular, asynchronous online course systems currently on the market (Hong et al., 2020; Kim et al., 2014; Laurillard et al., 2013; Wang et al., 2020). Specifically, Laurillard et al. (2013) recommend that online learning platforms must have systems in place to match learning designs to student need and allow for customization and flexibility of learning designs to maximize the educational experiences of students. This has been accomplished in other online courses by allowing students to customize learning goals, encouraging reflection and other self-monitoring activities, and fostering collaboration with classmates (Kim et al., 2014; Zhu & Bonk, 2019). These instructional elements have the added benefit of supporting students in developing the self-regulatory skills that are necessary to successfully engage with authentic work (Kim et al., 2014; Zhu & Bonk, 2019), which has applicability well beyond performance in a single online course.

Small magnitude associations between variation in authentic work and student outcomes demonstrate that integrating these

instructional activities on their own are unlikely to have a profound effect on reducing any sort of educational achievement gap. Instead, the most important implications of this research are in establishing the extent to which these type of widespread, standardized online course systems may be contributing to an expanding opportunity gap for students most at risk of not graduating from high school - the student group most likely to be assigned to online, high school courses (Powell et al., 2015) - and students attending low resourced schools who are most likely to be swayed out of necessity by promises of increased efficiency (Darling-Aduana et al., 2020).

It is also important to determine in a given school, program, and student context the possible merits of any increased expectations of higher-order thinking considering what seems to be a need for more support and attention for students to learn from these types of activities (Kirschner et al., 2006; Zheng et al., 2020). This is not to say that opportunities for higher-order thinking should not be integrated but rather that its integration may require larger systemic changes. For instance, courses may need to provide more intensive opportunities for student-teacher interactions to support the scaffolding and alignment of tasks requiring higher-order thinking to overcome the initial resistance and confusion students experience when exposed to these tasks for the first time (Hoffman & Ritchie, 1997; Kirschner et al., 2006; Patall et al., 2018). For these reasons, students may benefit more from the integration of higher-order thinking in face-to-face or other types of synchronous instructional settings. The demonstration of real-world relevance appears a more easily implementable leverage point, yet there is a ceiling on the extent to which this can be achieved without also requiring additional higher-order thinking (González et al., 2006; Hiebert et al., 2005).

5.1. Conclusion

While students belonging to historically marginalized groups may have the most to gain by access to authentic work (Au, 2012; Dee & Penner, 2017; González et al., 2006; Gutstein, 2016; Yair, 2000), exposure in a single online course in high school may be too little, too late to have a profound effect on their academic trajectories. Thus, the possible benefits identified in this and prior research must be weighed and considered within larger institutional structures and constraints to prevent the implementation of strategies to foster authentic work from harming students. For instance, higher instructional expectations may prevent some students from meeting minimum competencies required to obtain needed educational credentials or other prerequisites for post-secondary education, labor market, or personal success. Access to authentic work is a powerful piece, but still only one part, of dismantling systemic inequalities in educational opportunities.

Funding

This research was supported by a grant from the American Educational Research Association which receives funds for its "AERA Grants Program" from the National Science Foundation under NSF award NSF-DRL #1749275. Opinions reflect those of the author and do not necessarily reflect those AERA or NSF.

Credit author statement

Jennifer Darling-Aduana: Conceptualization, Methodology, Analysis, Writing, and Editing.

Compliance with ethic standards

This study was reviewed and approved by the University of Wisconsin-Madison Education and Social/Behavioral Science IRB, protocol #2014-1239-CP005.

Availability of data and materials: The data that support the findings of this study are available from the school district partner but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the school district partner.

Code availability

Code is available upon request from the author.

Declaration of competing interest

The author declares that they have no conflict of interest.

Acknowledgements

I thank the American Educational Research Association, National Science Foundation, and William T. Grant Foundation for generous funding of this research; Mason Shero, Hillary Bendert, and Jacqueline Aboulafia for their excellent research assistance; Carolyn Heinrich, Gary Henry, Jason Grissom, Mark Warschauer, and Lauren Margulieux for their invaluable feedback; Vendor researchers for their support in accessing data from the online instructional system; District administrators and educators for their assistance in providing student record data and support of the project data analysis and fieldwork; and Vanderbilt University and the Wisconsin Evaluation Collaborative at the Wisconsin Center for Education Research, University of Wisconsin-Madison, for ongoing

support of this initiative.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2021.104175.

References

Ahn, J., & McEachin, A. (2017). Student enrollment patterns and achievement in Ohio's online charter schools. *Educational Researcher, 46*(1), 44–57. https://doi.org/10.3102/0013189X17692999

Anyon, J. (1980). Social class and the hidden curriculum of work. *Journal of Education*, 67–92.

Apple, M. W. (2019). *Ideology and curriculum* (4th ed.). New York, NY: Routledge.

Au, W. (2012). *Critical curriculum studies: Education, consciousness, and the politics of knowing*. New York, NY: Routledge.

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1984). Bloom taxonomy of educational objectives. In *Allyn and Bacon*. Pearson Education.

Boninger, F., Molnar, A., & Murray, K. (2017). *Asleep at the switch: Schoolhouse commercialism, student privacy, and the failure of policymaking*. Boulder, CO: National Education Policy Center.

Burch, P., & Good, A. G. (2014). *Equal scrutiny: Privatization and accountability in digital education*. Cambridge, MA: Harvard Education Press.

Cavus, N., Uzonboylu, H., & Ibrahim, D. (2007). Assessing the success rate of students using a learning management system together with a collaborative tool in web-based teaching of programming languages. *Journal of Educational Computing Research, 36*(3), 301–321. https://doi.org/10.2190/T728-G676-4N18-6871

Creswell, J. W., & Clark, V. L. P. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Crippen, K. J., & Earl, B. L. (2007). The impact of web-based worked examples and self-explanation on performance, problem solving, and self-efficacy. *Computers & Education, 49*(3), 809–821. https://doi.org/10.1016/j.compedu.2005.11.018

Darling-Hammond, L. (2001). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.

Darling-Aduana, J. (2019). Behavioral engagement patterns among high school students enrolled in online courses. *AERA Open, 5*(4). https://doi.org/10.1177/2332858419887736

Darling-Aduana, J., Good, A. G., & Geraghty, E. (2020). *The culture of power online: An analysis of cultural responsiveness and relevance in online high school courses*. Urban Education. https://doi.org/10.1177/0042085920972169

Dee, T. S., & Penner, E. K. (2017). The causal effects of cultural relevance: Evidence from an ethnic studies curriculum. *American Educational Research Journal, 56*(1), 127–166. https://doi.org/10.3102/0002831216677002

DeVellis, R. F. (2016). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage Publications.

Dinov, I. D., Sanchez, J., & Christou, N. (2008). Pedagogical utilization and assessment of the statistic online computational resource in introductory probability and statistics courses. *Computers & Education, 50*(1), 284–300. https://doi.org/10.1016/j.compedu.2006.06.003

Gamoran, A., & Nystrand, M. (1992). Taking students seriously. In F. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 40–61). New York, NY: Teachers College Press.

Gemin, B., Paper, L., Vashaw, L., & Watson, J. (2015). *Keeping pace with k-12 digital learning: An annual review of policy and practice*. Durango, CO: Evergreen Education Group.

Golann, J. W. (2015). The paradox of success at a no excuses school. *Sociology of Education, 88*(2), 103–119, 10:1177/00380407.

González, N., Moll, L. C., & Amanti, C. (Eds.). (2006). *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Routledge.

Gutstein, E. R. (2016). "Our issues, our people – math as our weapon": Critical mathematics in a Chicago neighborhood high school. *Journal for Research in Mathematics Education, 47*(5), 454–504. https://doi.org/10.5951/jresmatheduc.47.5.0454

Haberman, M. (2010). The pedagogy of poverty versus good teaching. *Phi Delta Kappan, 92*(2), 81–87.

Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., & Michelman, V. (2017). The struggle to pass algebra: Online vs. face-to-face I for at-risk urban students. *Journal of Research on Educational Effectiveness, 10*(2), 272–296. https://doi.org/10.1080/19345747.2016.1168500

Heinrich, C. J., Darling-Aduana, J., Good, A. G., & Cheng, H. (2019). A look inside online educational settings in high school: Promise and pitfalls for improving educational opportunities and outcomes. *American Educational Research Journal, 56*(6), 2147–2188. https://doi.org/10.3102/0002831219838776

Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H., Smith, M., & Gallimore, R. (2005). Mathematics teaching in the United States today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis, 27*(2), 111–132. https://doi.org/10.3102/01623737027002111

Hoffman, B., & Ritchie, D. (1997). Using multimedia to overcome the problems with problem based learning. *Instructional Science, 25*, 97–115. https://doi.org/10.1023/A:1002967414942

Hong, H., Ma, L., Lin, P., & Lee, K. Y. (2020). Advancing third graders' reading comprehension through collaborative knowledge building: A comparative study in Taiwan. *Computers & Education*. https://doi.org/10.1016/j.compedu.2020.103962

Howard, T. C. (2019). *Why race and culture matter in schools: Closing the achievement gap in America's classrooms*. New York, NY: Teachers College Press.

Huberman, M., & Miles, M. B. (2002). *The qualitative researcher's companion*. Thousand Oaks, CA: Sage Publications.

Jerrim, J., Oliver, M., & Sims, S. (2019). The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learning and Instruction, 61*, 35–44. https://doi.org/10.1016/j.learninstruc.2020.101310

Joyce, J., Gitomer, D. H., & Iaconangelo, C. J. (2018). Classroom assignments as measures of teaching quality. *Learning and Instruction, 54*, 48–61. https://doi.org/10.1016/j.learninstruc.2017.08.001

Kim, R., Olfman, L., Ryan, T., & Eryilmaz, E. (2014). Leveraging a personalized system to improve self-directed learning in online educational environments. *Computers & Education, 70*, 150–160. https://doi.org/10.1016/j.compedu.2013.08.006

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Laurillard, D., Charlton, P., Craft, B., Dimakopoulos, D., Ljubojevic, D., Magoulas, G., Masterman, E., Pujadas, R., Whitley, E. A., & Whittlestone, K. (2013). A constructionist learning environment for teachers to model learning designs. *Journal of Computer Assisted Learning, 29*(1), 15–30. https://doi.org/10.1111/j.1365-2729.2011.00458.x

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal, 37*(1), 153–184. https://doi.org/10.3102/00028312037001153

McDonald, K. (2020). *May 11). Four K-12 education models that may gain popularity during COVID-19*. Forbes.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, D.C.: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.

Milner IV, H. R. (2012). But what is urban education? *Urban Education, 47*(3), 556–561, 0.1177/0042085912447516.

Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education, 104*(4), 280–312.

Oakes, J. (2005). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

Patall, E. A., Hooper, S., Vasquez, A. C., Pituch, K. A., & Steingut, R. R. (2018). Science class is too hard: Perceived difficulty, disengagement, and the role of teacher autonomy support from a daily diary perspective. *Learning and Instruction, 58*, 220–231. https://doi.org/10.1016/j.learninstruc.2018.07.004

Powell, A., Roberts, V., & Patrick, S. (2015). *Using online learning for credit recovery: Getting Back on track to graduation*. Vienna, VA: International Association for K-12 Online Learning (iNACOL).

Roll, I., Butler, D., Yee, N., Welsh, A., Perez, S., Briseno, A., & Bonn, D. (2018). Understanding the impact of guiding inquiry: The relationship between directive support, student attributes, and transfer of knowledge, attitudes, and behaviours in inquiry learning. *Instructional Science, 46*(1), 77–104. https://doi.org/10.1007/s11251-017-9437-x

Silseth, K. (2018). Students' everyday knowledge and experiences as resources in educational dialogues. *Instructional Science, 46*(2), 291–313. https://doi.org/10.1007/s11251-017-9429-x

Sun, J. C.-Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology, 43*(2), 191–204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

Wang, C., Fang, T., & Gu, Y. (2020). Learning performance and behavioral patterns of online collaborative learning: Impact of cognitive load and affordances of different multimedia. *Computers & Education, 143*. https://doi.org/10.1016/j.compedu.2019.103683

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach* (5th ed.). Mason, OH: South-Western CENGAGE Learning.

Yair, G. (2000). Educational battlefields in America: The tug-of-war over students' engagement with instruction. *Sociology of Education, 73*(4), 247–269. https://doi.org/10.2307/2673233

Zheng, B., Lin, C. H., & Kwon, J. B. (2020). The impact of learner-, instructor-, and course-level factors on online learning. *Computers & Education, 150*. https://doi.org/10.1016/j.compedu.2020.103851

Zhu, M., & Bonk, C. J. (2019). Designing MOOCs to facilitate participant self-monitoring for self-directed learning. *Online Learning, 23*(4), 106–134. https://doi.org/10.24059/olj.v23i4.2037