

# Measuring Equity-Promoting Behaviors in Digital Teaching Simulations: A Topic Modeling Approach

Joshua Littenberg-Tobias 

Elizabeth Borneman

Justin Reich

Massachusetts Institute of Technology

*Diversity, equity, and inclusion (DEI) issues are urgent in education. We developed and evaluated a massive open online course (N = 963) with embedded equity simulations that attempted to equip educators with equity teaching practices. Applying a structural topic model (STM)—a type of natural language processing (NLP)—we examined how participants with different equity attitudes responded in simulations. Over a sequence of four simulations, the simulation behavior of participants with less equitable beliefs converged to be more similar with the simulated behavior of participants with more equitable beliefs (ES [effect size] = 1.08 SD). This finding was corroborated by overall changes in equity mindsets (ES = 0.88 SD) and changed in self-reported equity-promoting practices (ES = 0.32 SD). Digital simulations when combined with NLP offer a compelling approach to both teaching about DEI topics and formatively assessing learner behavior in large-scale learning environments.*

Keywords: *diversity, equity, inclusion, simulations, MOOCs, natural language processing, structural topic modeling*

AN increasing number of organizations are investing significant resources in diversity, equity, and inclusion (DEI) training to address issues of bias and racist behavior (Carter et al., 2020). Many U.S. K–12 schools and organizations have also adopted DEI professional education to address issues of systemic racism. However, there is very little empirical evidence that these types of interventions lead to substantive changes in educators' teaching practice (Bravo et al., 2014; Civitillo et al., 2018; Ehrke et al., 2020; Love, 2019). DEI professional education often emphasizes broad, theoretical concepts that teachers have difficult transferring into practice (Emdin, 2011).

Simulations hold promise as a pedagogical approach for professional learning on equity-based teaching practices. Simulations have been used successfully in teacher education to approximate various teaching scenarios and support the transfer learning into new situations (Dalinger et al., 2020; Dotger & Ashby, 2010; Kaufman & Ireland, 2016). Increasingly, simulations are being used to help prepare teachers for equity-based teaching (G. A. Chen, 2020; J. A. Chen et al., 2020; Cohen et al., 2020; Self & Stengel, 2020). However, because of the difficulty of analyzing large amounts of text it can be difficult to assess participants' simulation performance. Accordingly, most studies of simulation have either limited their analysis to a small number of participants (G. A. Chen, 2020; Robinson et al., 2018; Self, 2016) or have focused on participants' reactions to the simulations rather

than behavior in the simulations themselves (Bautista & Boone, 2015; J. A. Chen et al., 2020; Dalinger et al., 2020; Girod & Girod, 2006).

In our study, we describe a novel method of analyzing participants' text-based responses to simulations using natural language processing (NLP) tools. We focus on data collected from a series of four digital equity teaching simulations that were given to participants within a massive open online course (MOOC) for K–12 classroom educators on equity teaching practices ( $N = 963$ ). In the simulations, participants responded using oral or written text, and they produced more than 30,000 rows of unstructured text data. Using structural topic modeling (STM; Roberts et al., 2014)—an NLP method that does not require pre-labeled data—we analyzed participants' text responses for discrete “topics” that represented different tacks participants might take in responding to prompts from characters in a simulation.

To assess the validity of this approach, we examined whether the prevalence of topics within simulation responses was associated with participants' attitudes toward equity on surveys. Finally, we investigated whether STM can be used to detect changes in participants' behavior across different simulations. Using a reference group of “high-equity” (HE) participants—those who were in the top quartile (25%) of equity-related responses on the course presurvey—we analyzed whether the remaining participants converged with these HE participants in their simulation responses and



corroborate changes by examining whether these patterns were consistent with changes in attitudes and self-reported behaviors on surveys and ratings given by external human raters.

As attention to DEI issues become an essential component of teacher education and teacher professional learning, simulations are likely to play an increasingly important pedagogical role (Cohen et al., 2020; Self & Stengel, 2020). This study provides insight into how NLP tools can be used to formatively assess educators' text-based responses within simulations. As one of the first studies to apply NLP tools to equity-based simulations, we examine the potential of these tools for understanding participants' experiences in simulations and for ultimately developing automated supports to facilitate teacher learning on equity teaching practices.

### **Background and Context**

#### *The Role of Simulations That Links Equity and Practice in Teacher Education*

Teacher education, like many other fields, has long been criticized for examining equity issues disconnected from practice (Sleeter, 2012). Even when teacher education programs include equity issues in their curriculum, they often do not make explicit connections to specific teaching practices (Bravo et al., 2014; Kavanagh & Danielson, 2020). Once teachers enter the profession, though they might be aware of overall societal factors that limit opportunities for particular groups of students, they struggle to translate this knowledge into practice in order to disrupt inequitable practices in the day-to-day experience of students in schools (Emdin, 2011; Love, 2019; Matias et al., 2016; Milner, 2010; Schiera, 2020).

Simulations have particular affordances that make them well-suited for DEI educator professional learning. Simulations provide opportunities for what Resnick (1987) described as “bridging apprenticeships” connecting theoretical learning with classroom practice. Because simulations necessitate action, participants draw on the same cognitive, affective, and linguistic processes as they would in comparable real-world situations (Aldrich, 2009; G. A. Chen, 2020; Slater et al., 2009). At the same time, simulations often simplify the real world. Many are intentionally designed to represent some aspects of reality, but not others, directing the attention of users to specific features and patterns in the simulated environment (Aldrich, 2009; Mislavy, 2013). This allows them to serve as “approximations of practice” (Grossman et al., 2009); creating spaces for less-skilled practitioners to try out new practices and receive feedback within scaffolded, low-stakes environments (Kaufman & Ireland, 2016; Theelen et al., 2019).

Simulations have a long history in teacher education and were used during the height of the “microteaching” movement of the 1960s and 1970s in order to prompt teachers to

enact specific, discrete skills (Cohen et al., 2020). Some of these methods were critiqued for decontextualizing and deprofessionalizing teachers' practice (Kavanagh et al., 2020; Zeichner, 2012). Instead of focusing on developing discrete skills, more recent teaching simulations focus on improving teachers' “instructional vision”; their ability to notice and interpret educationally significant information through a professional teaching lens (Gibbons & Cobb, 2016; Sherin et al., 2011; Theelen et al., 2019). Simulations have been used to help educators learn how to modify their instruction based on student needs (Ferry et al., 2005; Girod & Girod, 2006); respond to student off-task behavior (Cohen et al., 2020); elicit student ideas (Bautista & Boone, 2015; Wang et al., 2020); and engage in difficult conversations with parents (G. A. Chen, 2020; Thompson et al., 2019).

Simulations that prompt educators to attend to equity issues are still an emerging area of research, but initial studies show promise for helping educators change behavior. For example, Self and Stengel (2020) used simulations to call participants' attention to the disproportionate discipline rate of Black students in schools. Novice teachers who participated in the simulations initially did not recognize the racial bias highlighted in the simulation, but through cycles of reflection and debriefing, the participants moved toward greater consciousness of racial bias. Cohen et al. (2020) used a randomized control trial to examine the effect of instructor feedback in a mixed-reality simulation on responsive classroom management. Participants who received feedback during and after the simulation experience were less likely to recommend that punitive steps be taken against student avatars who exhibited minor off-task behaviors in the simulation than participants who only reflected on the experience. These promising early findings suggest that learning experiences that include simulations as part of a learning cycle of instruction, practice, self-reflection, and expert feedback may play an important role in developing DEI training that leads to changes in teacher practice.

#### *Developing Digital Teaching Simulations on Equity Issues*

Simulations in teaching range from interactions with a live human actor to interactions with human-controlled digital puppets to fully digital simulations where users interact with a computer. Most research on teacher education simulations on DEI issues examines only simulations that include human actors (Cohen et al., 2020; Dalinger et al., 2020; Dotger & Ashby, 2010; Kannan et al., 2018). One advantage of fully automated digital simulations, without human actors, is they are inexpensive to produce and can be flexibly deployed enabling learners to access them remotely at the time and location of their choice. Particularly for in-service professional learning where time and resource constraints are a constant concern, fully digital simulations may be a practical compromise expanding access on this urgent topic to more teachers.

In this study, we focus on a series of fully digital equity teaching simulations that were embedded within an online course for K–12 educators on equity in education. The digital simulations were developed in *Teacher Moments*, a free openly licensed teaching simulation platform developed by the MIT Teaching Systems Lab (<https://teachermoments.mit.edu/>; Dutt et al., 2021; Hillaire et al., 2020; Sullivan et al., 2020; Thompson et al., 2019). These digital simulations were designed differently than other computer-based simulations where participants selected actions from a predetermined list of responses (Ferry et al., 2005; Girod & Girod, 2006). In these simulations, participants are immersed in a specific teaching situation that is represented by using text, images, and video within a web-based platform. As participants move through the simulations they encounter a number of decision points where they respond with natural oral language or text responses that provide a more authentic simulation experience than multiple choice scenarios (Hirumi et al., 2016). Engaging and reflecting on these types of interactions can help participants become adept at noticing and attending to issues of equity (Borneman, Littenberg-Tobias, & Reich, 2020).

#### *Using Natural Language Processing to Evaluate Responses From Digital Simulations*

Although studies of *Teacher Moments* scenarios have found preliminary evidence that participants find them authentic and contribute to their learning (Robinson et al., 2018; Thompson et al., 2019), at the time when the simulations used in the study were developed, the platform had no way of evaluating natural text and oral language from participants' responses. NLP uses computational tools such as machine learning to automate analysis of text responses. With NLP, researchers are able to quantitatively analyze text data for patterns that are predictive of participants' attitudes, behaviors, and practices. NLP tools have been used, for example, to analyze transcripts of classroom teaching to detect the presence of particular instructional and discourse moves or to evaluate classroom climate (Jensen et al., 2020; Ramakrishnan et al., 2019).

Many applications of NLP in education rely on supervised approaches that rely on previously labeled text responses using a priori classifications (Datta et al., 2021; Jensen et al., 2020; Ramakrishnan et al., 2019). In unsupervised approaches to computational text analysis, these categories are drawn from the data itself. The advantage of unsupervised methods is that they do not require a priori assumptions about the structure of the data or time-consuming, hand-labeling of text responses. These properties make unsupervised approaches a good fit for analyzing text responses within large-scale learning environments like MOOCs where participants may produce hundreds of thousands of lines of unstructured text.

One unsupervised approach is the STM—a form of computational text analysis that has been used as for computer-assisted interpretation of open-ended text responses (Roberts et al., 2014). Similar to other types of topic models (e.g., Latent Dirichlet Allocation), STM uses a probabilistic mixture model to estimate the probability distribution of topics within documents and words within topics (Blei et al., 2003). STM extends these models by allowing for the inclusion of covariates in the topic modeling (Roberts et al., 2014). The inclusion of covariates in topic modeling is useful in social science research because it allows researchers to examine how the prevalence and content of the topics might vary based on information about the participant, such as demographics, attitudes, or behaviors (Roberts et al., 2014).

A number of studies have used STM to examine relationships between patterns in open response texts and participant demographics (Roberts et al., 2016; Yeomans et al., 2018). For example, Sterling et al. (2019) applied STM to political tweets assessing whether liberals and conservatives had different conceptions of what constituted a “good society” based on the prevalence of topics in their tweets. Some recent studies have used STM to explore population-level changes in topic distributions within a population over time. For example, X. Chen et al. (2020) studied trends in topics within *Computers & Education* over a 40-year period, finding that certain topics, such as teacher training and social networks, increased over time while programming languages and hardware decreased. Kim et al. (2020) applied STM to self-reported inventories of religion and spirituality over a 60-year period, demonstrating changes in how the authors of the texts experienced their spirituality.

These studies suggest that STM can be a useful way to identify aggregate trends across a variety of different types of media. However, there has been limited research on how STM might perform with simulation response data, particularly with simulations on DEI issues in education. In this study, we seek to understand how this method of computational text analysis performs when used to analyze a large-scale data set of equity teaching simulation data.

#### **Theoretical Framework**

To measure equity-oriented mindsets and behaviors, we drew on the opportunity-centered teaching framework developed by Milner (2010, 2012). Opportunity-centered teaching is rooted in critical race theory and culturally responsive and sustaining teaching and pedagogy (Gay, 2002; Ladson-Billings, 1995; Paris, 2012). It posits that when schools fail to effectively serve racially, ethnically, and culturally diverse students, it is not because of innate lack of academic ability or a demonstrated lack of effort, but because of systemic barriers that limit the potential of these students. These systemic barriers are often enacted and justified by well-intentioned educators who nonetheless

possess mindsets that enable these barriers to persist. For example, many educators believe in the “myth of meritocracy,” which is the belief that all success is earned through talent and/or effort and, by extension, failure is a result of individual faults (Milner, 2012). Similarly, many educators maintain a “color-avoidant” stance toward discussing race and racism with students, believing that avoiding the topic is preferable to acknowledging it explicitly (Bonilla-Silva, 2015; Cochran-Smith, 1995; Neville et al., 2013). Changing these educator mindsets toward more equity-oriented mindsets is a first step toward dismantling racist systems and practices in schools.

In the design and analysis of the *Teacher Moments* simulations that we developed, we further drew on an adaption of Milner’s (2010, 2012) opportunity-centered teaching framework by Filback and Green (2013) that describes equity as contrast between two opposing viewpoints, where one of the viewpoints represents an equity-oriented lens (e.g., “Equality vs. Equity,” “Deficit vs. Asset,” “Avoidant vs. Aware,” and “Context-Neutral vs. Context-Centered”). For example, an educator with an “Equality” mindset believes that all students should receive the same treatment regardless of their circumstances, while an educator with an “Equity” mindset believes that students should be treated differently depending on their circumstances and needs.

Each of these viewpoints has an appropriate time and place in teaching. For example, equality should, of course, be one factor in teacher decision making. However, these competing tensions are often out of balance in schools where concerns about equality often lead to inequitable outcomes. For example, a homework policy that does not allow flexibility for students to submit assignments after the due date is “equal” because all students are held to the same standards. However, this policy is inequitable because wealthier students have more parental support for homework and fewer family obligations and, thus, are more likely to be able to submit homework on time (Sayers et al., 2020).

The four *Teacher Moments* simulations we developed were designed to elicit educator responses related to a specific mindset (see Table 1 for a description of all the simulations). For example, the *Roster Justice* simulation was designed around the “Avoidant” versus “Aware” mindset. Participants play the role of a middle school teacher trying to talk with their school principal about demographic disparities—including race, gender, and dis/ability—in class rosters for computer science and math. The principal is resistant to changing the rosters and at various points offers excuses, tries to offer “quick-fixes” that do not address the underlying problems, and moves the conversation away from race toward logistical challenges. The participant must decide after each interaction how much they want to explicitly press the issue of the demographic disparities, as well as the structural issues with the course schedules, to the principal. Furthermore, the participant must decide whether they are

willing to accept solutions that do not address the underlying discriminatory problems.

## Current Study

In this study, we applied the structural topic model, a form of unsupervised NLP, to participants’ simulation responses in order to evaluate how participants’ applied equity mindsets within these simulations. We also used these tools to study how participants’ responses to these simulations change over a series of successive simulations that were embedded within an 8-week MOOC on equity teaching practices in education. We address the following research questions:

*Research Question 1 (RQ1):* To what extent can structural topic modeling be used to detect aggregate trends in participant responses within digital simulations on equity teaching practices?

*Research Question 2 (RQ2):* To what extent can topic distributions, when compared to a reference group, be used to detect individual changes in participant responses within digital simulations?

*Research Question 2a:* To what extent did course participants change in their topic distributions over sequential simulations to be more similar to a reference group of participants with initially more equitable mindsets?

*Research Question 2b:* To what extent are these measures consistent with external ratings and changes in equity attitudes and self-reported practices on surveys?

## Method

### *Study Context*

In spring 2020, we launched an 8-week online course on equity in education on edX called “Becoming a More Equitable Educator: Mindsets and Practices.”<sup>1</sup> Although the primary audience was K–12 educators, the course was free with open enrollment and thus anyone with an interest in the topic could register for the course. The course was publicized through emails to registrants from previous education-related courses, via social media advertisements, and through individual networks. The course was explicitly marketed to a broad educator audience rather than solely at educators explicitly interested in equity. However, given the topic of the course, we expected educators who signed up for the course to be interested in addressing equity issues in schools.

A sample unit (Figure 1) starts with a video lecture that introduces one of the four equity mindsets and describes its importance to teaching. Participants then watched documentary-style case study videos of practicing teachers describing how they implemented that mindset in their teaching. The participants then completed a *Teacher Moments* digital

TABLE 1  
*Descriptions of Each Simulation and Associated Dimension of Equity*

Dimension of equity	Simulation description
<p><i>Equality vs. Equity</i>—An equality perspective argues that all students should be treated the same, whereas an equity perspective argues that students should be treated differently, and sometimes given additional resources, based on their needs.</p>	<p><i>Jeremy’s Journal</i>—Over the period of a week, participants observe the work products and social behaviors of an outgoing student, Jeremy Green. Jeremy demonstrates some understanding of the learning material, yet he seems to prioritize his social life with frequent distractions. External events come up midweek causing Jeremy to miss a day of class, and he does not have a doctor’s note, which conflicts with school policy. Teachers must decide whether they will have Jeremy take the scheduled Friday quiz alongside the rest of the class, keeping in mind his nervousness, lack of preparedness, and fear of failure.</p>
<p><i>Deficit vs. Asset</i>—A deficit perspective focuses on how educators can “fix” students by changing their behavior and attitudes, whereas an asset perspective focuses on how educators can recognize and build on students’ existing strengths.</p>	<p><i>Coach Wright</i>—Participants play an instructional coach who is observing a teacher, Ms. Porter, over an ELA class period. Ms. Porter is working to strengthen student relationships and facilitates a classroom discussion in which Jeremy Green participates without raising his hand. Participants observe an interaction with Jeremy Green and his basketball coach, Coach Wright, while Jeremy is lingering in the hallway on his bathroom break. Although Ms. Porter would have given Jeremy a referral for lingering per school rules, Coach Wright motivates Jeremy to return to his class work by drawing on his assets. As the instructional coach, participants have the chance to coach Ms. Porter in motivating Jeremy through his unique strengths and assets and must decide whether he will receive the referral.</p>
<p><i>Avoidant vs. Aware</i>—An avoidant perspective explicitly avoids mentioning or considering race in order to be racially unbiased, whereas an aware perspective acknowledges the role that race plays in students’ experiences in schools and seeks to explicitly name and disrupt system that permutate racial inequity.</p>	<p><i>Roster Justice</i>—Roster Justice gives participants an opportunity to detect inequitable patterns in school scheduling at the start of a new semester, and to practice articulating these discrepancies to their school principal, Mr. Holl. Participants navigate a dialogue with Mr. Holl who would rather not address the demographic issues due to time constraints and assumed student choice. Participants practice suggesting tangible solutions at the school policy level, advocating for racial awareness and equitable opportunity to take computer science classes for all students.</p>
<p><i>Context-Neutral vs. Context-Specific</i>—A context-neutral perspective maintains that “good teaching” is the same in all contexts and educators do need to adjust their curriculum and instruction to reflect the community students are in, whereas a context-specific perspective argues that communities and home-lives affect learning and thus these need to be integrated these into teaching.</p>	<p><i>Layers</i>—In Layers, participants reflect on how students’ communities and home-lives impact their learning, performance, and belonging in the classroom. Teachers practice building connections between classroom material and the whole student, by adapting a lesson based on new information they learn about their students.</p>

Note. ELA = English language arts.

simulation where they were asked to apply their understanding of an equity mindset within a simulated scenario environment. After completing the simulation, participants watched a documentary-style filmed debrief that featured practicing K–12 teachers who had completed the same simulation. The debrief spliced footage of a live, in-person, facilitated group discussion with cutaway scenes where case study teachers are interviewed about the decisions they made in the simulation. Each unit concludes with an assignment where participants were asked to apply the particular equity mindset in their own practices.

The online course had 7,918 registrants and of those, 5,678 (72%) clicked into the course at least once. In order to

focus on participants who meaningfully interacted with the content in the course, we restricted our analysis sample to participants who engaged in at least one simulation ( $N = 963$ , 12% of registered participants) during the course. Of those in the analysis sample, 609 (63% of analysis sample) interacted with at least two of the simulations in the course and 342 (36% of analysis sample) interacted with all four of the simulations. Many of participants in the analysis sample were located in the United States (58%) and 81% identified as fluent English speakers. The majority of participants reported working in K-12 schools (53%). The demographics of participants in the analysis sample roughly reflected the demographics of educators and those working in



FIGURE 1. Course outline and sample unit.

education-aligned industries such as teacher preparation and instructional design (Table 2).

### Instrumentation

Researcher-developed survey instruments were embedded within the course platform, which previous research has found increases response rates (van de Oudeweetering & Agirdag, 2018). We administered the survey instruments three times during the course: precourse, postsimulation, and immediately postcourse. A follow-up survey was administered to all participants in the analysis sample in October–November 2020. Response rates for the survey were generally high for an MOOC (precourse = 56%, postsimulation = 64%, postcourse = 42%, follow-up = 18%).

### Measures

**Equity Mindsets.** Our main survey instrument was a set of scales designed to measure each individual pair of educator mindsets (e.g., Equality vs. Equity, Avoidant vs. Aware). We developed a new set of scales to ensure that the survey

measures were aligned to the mindsets as they were presented in the online course. We developed an initial set of items in Fall 2019 containing more than 100 items. These items were evaluated for content validity and clarity by a panel of reviewers that included equity experts, survey researchers, and educators. Based on their feedback, we modified and reduced the items to a 64-item survey. We gave this survey to a pilot sample of K–12 educators who also gave feedback on the items ( $N = 125$ ; Anghel & Littenberg-Tobias, 2020). Based on feedback from the pilot sample and the psychometric performance of these items, we selected 26 items and administered those within a nonequity-focused MOOC for educators ( $N = 502$ ). We then evaluated the psychometric performance of the items and selected a set of 18 items that balanced internal consistency with construct representation. For the final instrument, we added another three items resulting in a 21-item instrument with four scales. To assess the concurrent validity of the scales, we included a validated measure of equity attitudes, the *Colorblind Racial Awareness—Blatant Racial Issues* (CoBRAS-BRI) scale on all administrations of the survey (Neville et al., 2000). All the scales we developed were correlated with the validated

TABLE 2  
Participant Demographics

Individual demographics		Employer, roles, and school characteristics	
Category	Percentage		Percentage
Gender		Employer	
Female	69	K–12 school	53
Male	31	Not-for-profit or NGO	15
Highest degree		University	19
High school or less	4	Other	13
Associate’s degree	3	Role (K–12 school only)	
Bachelor’s degree	32	Classroom teacher	73
Master’s or professional degree	53	Administrator	10
Doctoral degree	9	Instructional coach	12
Race		School type (United States only)	
American Indian or Alaska Native	<1	District public school	67
Asian	18	Private school	14
Black or African American	8	Charter school	15
Hispanic, Latino, or Spanish	8	School sector (Non–United States)	
More than one race	8	Public school	55
Other	4	Private school	45
White	55	Percentage economically disadvantaged	
Current location		0–10	21
In United States	58	11–25	22
Outside United States	42	26–50	19
English proficiency		>50%	38
Basic	1		
Intermediate	6		
Proficient	13		
Fluent	81		

Note. NGO = nongovernmental organization.

CoBRAS-BRI scale ( $r = 0.33\text{--}0.71$ ,  $p < 0.001$ ) suggesting that the scales were assessing similar, though not identical, constructs.

The equity mindset scales were administered four times: precourse, postsimulation, immediate postcourse, and 4-month follow-up. We did not include the Equality versus Equity mindset scale in *Jeremy’s Journal* because it was at the beginning of the course and thus participants’ mindsets were unlikely to have changed from the presurvey. Descriptive statistics on each of the survey items and scales can be found in Table 3, and Cronbach alpha statistics are in Table 4.

*Equity-Promoting Practices.* We also developed a survey instrument to measure participants’ use of equity-promoting practices (Table 3). The five-item instrument assessed participants’ self-reported use of activities such as “reflecting on how your identity influences your actions,” “identifying student strengths,” and “participating in networks on equity.” We selected these activities because of their alignment with

the content of the course. The instrument was included in the precourse, immediate postcourse, and follow-up surveys, and in all cases had Cronbach alpha statistics of greater than 0.80, indicating a high level of internal consistency (Table 4). Participants’ self-reported use of equity-promoting practices was also moderately correlated with the externally validated CoBRAS-BRI scale (0.16–0.32,  $p < 0.001$ ), which suggested evidence for concurrent validity.

#### *Structural Topic Modeling of Open-Text Simulation Data*

In this section, we describe the steps we used to prepare and estimate the STM model on the open-text data from the equity simulations in the online course. Our analysis pipeline is also depicted in Figure 2.

*Data Preprocessing.* We followed the standard methods for preprocessing data for topic modeling. We removed common stopwords (e.g., “and,” “or,” “I”), punctuation, and numbers, and then applied the SnowballC stemming function to each token. We also separated each response out by

TABLE 3  
*Descriptions Statistics for Survey Measures*

Survey measure	Presurvey			Postsimulation			Postsurvey			Follow-up survey		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
Equality-Equity Overall	4.68	1.01	534				4.88	1.03	403	5.02	0.76	170
Success in school is primarily the student's responsibility.	4.50	1.23	530				4.83	1.10	401	4.89	0.97	170
All people are born with the same opportunities to be successful.	5.08	1.34	533				5.08	1.41	403	5.34	1.11	170
Today's schools help all students equally.	4.90	1.22	528				5.10	1.16	400	5.14	1.11	169
Anyone who works hard enough can do well in school.	3.79	1.47	531				4.25	1.48	401	4.35	1.32	170
Students from poor families have the same opportunities to succeed as students from rich families.	5.15	1.24	532				5.14	1.35	400	5.37	0.99	169
Asset-Deficit Overall	5.11	0.58	536	5.23	0.57	519	5.32	0.55	402	5.26	0.56	171
Teachers should have high expectations for all students.	5.10	1.13	533	5.42	1.03	518	5.40	0.97	401	5.33	1.02	171
Every student can be successful given the right supports.	5.37	0.87	533	5.56	0.85	518	5.55	0.65	402	5.39	0.94	169
Teachers should identify all students' strengths even if they do not fit within traditional school norms.	5.49	0.73	532	5.64	0.80	516	5.62	0.65	402	5.56	0.88	171
Teachers do not need to know much about their students beyond their grades and behavior in class.	5.31	1.03	532	5.53	0.91	518	5.52	1.02	401	5.49	1.02	180
All students should be expected to follow the same traditional school norms.	4.27	1.31	530	3.95	1.30	517	4.52	1.27	401	4.72	1.12	171
It is a teacher's job to challenge all students academically.	5.12	0.96	530	5.29	0.85	518	5.29	0.93	401	5.05	1.04	170
Avoidant-Aware Overall	4.74	0.92	534	5.08	0.69	452	5.10	0.79	404	5.13	0.78	171
Teachers should consider students' race when teaching.	4.45	1.44	528	5.06	1.20	447	5.04	1.19	403	5.02	1.24	170
Students' race affects their experiences in schools.	5.01	1.11	534	5.40	0.92	451	5.36	0.90	403	5.32	1.00	170
The current school curriculum is meaningful for students from almost all backgrounds.	4.09	1.42	529	4.26	1.28	451	4.34	1.43	402	4.56	1.31	169
Students' identities affect their access to opportunities in schools.	5.02	1.11	532	5.17	0.98	448	5.31	1.01	403	5.32	0.99	171
Teachers should talk with their colleagues about how race affects students' experiences in schools.	5.12	1.01	529	5.55	0.68	448	5.47	0.77	404	5.42	0.90	171
Context-Specific and Context-Neutral Overall	5.34	0.60	534	5.61	0.50	447	5.55	0.58	403	5.44	0.71	171
Acknowledging the context in which the school is located can help students learn.	5.22	0.84	528	5.57	0.68	446	5.54	0.70	401	5.35	0.82	170
Students' surroundings affect the way they engage with the material.	5.44	0.69	531	5.70	0.54	446	5.60	0.63	403	5.50	0.80	169
Engaging with the community can help motivate students.	5.37	0.68	531	5.64	0.58	446	5.53	0.67	402	5.43	0.84	171
Communities play a big role in students' success.	5.37	0.71	531	5.52	0.70	447	5.53	0.68	403	5.46	0.82	170
Educators should include elements of students' lives outside of school in their teaching.	5.28	0.85	532	5.66	0.60	445	5.57	0.68	402	5.47	0.84	171

(continued)



TABLE 3 (CONTINUED)

Survey measure	Presurvey			Postsimulation			Postsurvey			Follow-up survey		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
Composite Equity Attitudes	4.97	0.60	527				5.22	0.56	397	5.22	0.54	170
Equity-Promoting Behaviors Overall	3.46	0.76	587				3.66	0.76	422	3.76	0.76	175
Reflect on how your own identity influences your interactions with students	3.91	0.90	586				4.10	0.88	419	4.09	0.90	175
Identify strengths for the students you work with	4.08	0.82	586				4.16	0.80	420	4.25	0.79	173
Discuss with colleagues equity issues in your school or context	3.44	1.00	585				3.52	1.03	419	3.67	1.00	174
Share resources with colleagues about equity issues	3.17	1.11	585				3.40	1.09	421	3.62	1.07	175
Participate in a network of educators formed specifically around issues of equity (online or in-person)	2.69	1.22	587				3.09	1.20	418	3.16	1.26	174

TABLE 4

*Cronbach Alpha Statistics for Survey Measures*

Survey measure	Presurvey	Postsimulation	Postsurvey	Follow-up survey
Equality-Equity	0.82		0.85	0.71
Asset-Deficit	0.56	0.62	0.61	0.52
Avoidant-Aware	0.80	0.69	0.78	0.76
Context-Neutral and Context-Specific	0.85	0.81	0.91	0.91
Equity-Promoting Behavior	0.80		0.82	0.80

sentence using standard terminal English punctuation marks to identify the end of sentences. We then followed the procedure recommended by Roberts et al. (2014) and removed any infrequent terms (e.g., those that appeared in fewer than five documents) and created a document-term matrix. We then merged in the metadata for each document including a dummy indicator for each simulation prompt and the survey scale score for the attitude addressed in the simulation. The final data set for each simulation consisted of one row per sentence from the open-text simulation responses (*Jeremy's Journal*  $N = 13,160$ ; *Coach Wright*  $N = 11,492$ ; *Roster Justice*  $N = 7,606$ ; *Layers*  $N = 7,429$ ).

*Number of Topics to Extract.* We determined the number of topics to extract by generating a series of candidate models ranging from five to 50 topics. For each candidate model, we extracted the semantic coherence and exclusivity. Semantic coherence measures how cohesive the topic is based on how often higher probability words for the topic co-occur in the same document, and exclusivity measures how often words that have high probability in one topic have a low probability in another topic (Roberts et al., 2014). We then chose the model that best balanced these two criteria. The semantic coherence and exclusivity of our candidate models and

related figures can be found in the Supplemental Appendix (available in the online version of this article).

*Structural Topic Model Estimation.* We used the *stm* package in R to estimate the STM model using the “spectral” initialization type (Roberts et al., 2019). We included the survey scale score for the associated mindset and a dummy variable indicating the simulation prompt as covariates. The survey scale score allowed us to examine whether individuals with more equity-oriented mindsets related to that attitude responded differently to the simulation prompts than those with less equity-oriented beliefs. The scale was standardized so that a 1-unit change was associated with a 1 standard deviation (*SD*) increase in the equity mindset survey scale. This allowed us to identify the topics that were more likely to occur for participants with higher endorsements of the more equitable mindset on the survey. A dummy variable for the simulation prompt was included to improve the fit of the STM model.

*Assigning and Validating Topic Labels.* Following the procedures described in Roberts et al. (2016) and X. Chen et al. (2020), we extracted the five most frequent, distinct words for each topic using the FREX metric, which is the weighted

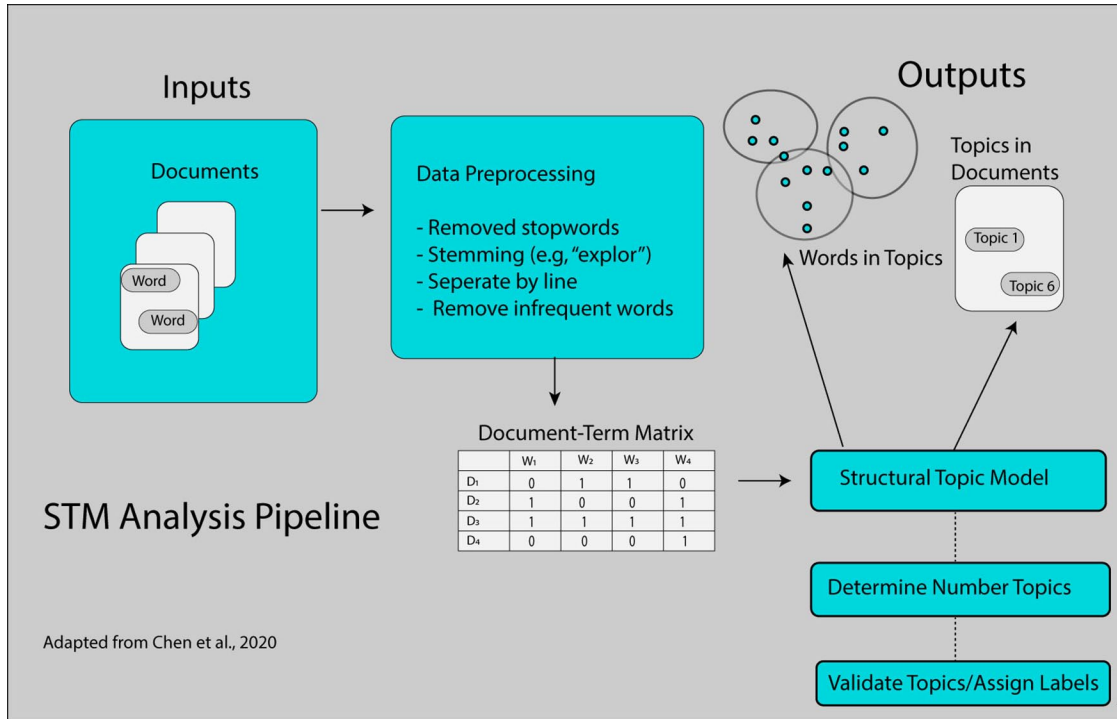


FIGURE 2. Structural topic model analysis pipeline.

harmonic mean of a word’s rank in terms of frequency and exclusivity to the topic (Roberts et al., 2016). We also identified the five most representative texts for each topic (using the *plotQuote* function in the *stm* R package). Members of the research team (one author and two research assistants) then independently generated descriptive labels for each of the topics. After generating the labels, the members of the research team discussed the label terms until a consensus was achieved for each topic.

To validate this method for assigning topic labels, we adapted a proposed method by Chang et al. (2009) to the topic labels we assigned in *Jeremy’s Journal*. We randomly selected 10% of participant responses from each *Jeremy’s Journal* prompt. We then extracted the topic that had the highest probability of appearing in that response and three other randomly selected topics associated at a lower than random chance probability with the response. Four research assistants who were not previously involved in assigning labels served as validation raters. Validation raters were shown a response (e.g., “Jeremy is not paying attention, may be distracted”) and were asked to select which of the four topics was the most appropriate for that response. Overall, validation raters selected the topic with the highest probability 65.2% of the time, substantially greater than random chance ( $t = 225.77, df = 1,203, p < 0.001$ ).

#### Data Analysis Strategy

*RQ1: Topic Modeling Within Equity Simulations and Associations With Equity Mindsets.* We estimated a separate

structural topic model for each of the four simulations in the course. The model produced a set of  $K$  topics that then needed to be labeled and interpreted in terms of the equity teaching literature. For each topic, the structural topic model estimates the posterior probability that the topic would appear given the text features of the document ( $\theta_T$ ; Roberts et al., 2014). We then used the *estimateEffect* function in the *stm* package to estimate the relationship between topic prevalence ( $\theta_T$ ) and participants’ survey subscale score for the specific mindset measured in the simulation. The estimate represents the difference in how likely a participant is to mention a topic in their simulation responses based on their survey responses for that mindset. A nonnull relationship indicates that there is a relationship between the prevalence of topics within a participants’ response and that mindset.

*RQ2: Changes in Simulation Response Behavior and Associations With Changes in Mindsets and Equity-Promoting Practices.* For this question, we explored whether participants changed how they responded across each successive simulation. Because each of the simulations presented new sets of scenarios and content, we could not measure changes in topic proportions directly. To address this issue, we developed a novel method to compare the topic distributions relative to a fixed group of participants over the course of several simulations. We assumed that participants who reported higher equity mindsets on the presurvey would from the beginning of the course respond more equitably in the simulations. We then chose to compare all participants’ responses to those who were in the top quartile (top 25%) for the

associated equity attitude on the precourse survey. We refer to these participants as HE participants. If the remaining participants, who we refer to as low-equity (LE) participants, became more similar in their responses to the HE participants, it would suggest that these LE participants were responding more equitably in the simulations. However, if their responses remained different, it would suggest that the LE participants' behavior within the simulations did not change to become more equitable during the course.

To measure how individual LE participants changed in their responses within the simulation, we aggregated the posterior probabilities across each participant's responses in the simulation and selected the maximum posterior probability for each topic ( $\max[\theta_T]$ ). This represented the maximum probability of a topic appearing in a participant's simulation responses; in other words, how likely the participant was to mention this topic in any of their responses.

We then analyzed the extent to which changes in attitudes were evident across simulations by measuring the average Euclidean distance from the HE participants. Euclidean distance can be used to measure the relative similarity between the topic distributions of two documents in the same corpus (Du et al., 2015). The smaller the average Euclidean distance from HE participants the more similar the LE participants simulation responses were to those participants. We normalized the Euclidean distance so that it was consistent across simulations. We then calculated the difference in the change distance between the first and last simulation. A larger decrease in distance means that participants were becoming more similar in their simulation responses to the HE participants and that their behavior in the simulation therefore became more equitable.

To validate this approach, we used two different methods. First, we compared changes in simulation responses to changes in equity attitudes and self-reported equity-promoting behaviors on surveys. For equity mindsets, we calculated a composite score by averaging participants' scores on each of the four mindset dimensions of equity and then averaging across these four dimensions so each component was equally weighted. Restricting our analysis to only LE participants, we then conducted a paired  $t$  test to assess whether the mean difference in equity mindsets and self-reported equity-promoting practices between the pre- and postsurvey and the pre- and follow-up survey was statistically significantly different from zero. We also calculated the Cohen's  $d$  effect size ( $ES$ ) for each mean difference. If both the distance metric and the surveys indicate changes in attitudes and practices it suggests that they are measuring similar underlying changes.

Second, for the *Roster Justice* simulation we compared the distance metric to ratings that were generated by human raters using a shared rubric. The rubric was designed to measure participants' use of Avoidant versus Aware mindsets reasoning in their responses (see the online Supplemental Appendix for the full rubric). Each response within *Roster Justice* was coded using a 5-point scale, where 1 represented

the most Avoidant response and 5 the most Aware (e.g., most equitable) response (Borneman, Smith, & Littenberg-Tobias, 2020). Raters were research assistants (supervised by the authors) and did not see the distance measure for any of the responses they rated. We randomly sampled 25% of responses to be double-rated. In all cases, raters reviewed their assignments independently and did not discuss ratings. We calculated a weighted Krippendorff's alpha statistic 0.76. If the ratings are correlated with the distance metric, it offers additional validity evidence that distance from the HE reference group is an indicator of the extent to which responses reflect more equitable mindsets.

## Results

We begin by describing the topics extracted from each simulation and the associations between topic prevalence and the attitude survey scales (RQ1). We then examine whether participants changed in their patterns of simulation responses over successive simulations in the course and assess to what extent these changes are consistent with changes in attitudes and practices on surveys and ratings by external human raters (RQ2).

In each of the four simulations, we identified particular topics produced by the STM that were correlated with more equity-oriented mindsets and other topics that were correlated with less equity-oriented mindsets. When we reviewed these topics and correlations, they often reflected tensions and decision points intentionally built into the scenarios. Although not every difference in topic prevalence reflected differences in equity mindsets, we found connections frequently enough to suggest that STM can identify meaningful equity-based differences in simulations responses. Additionally, we found that participants with less-equitable beliefs at the beginning of the course converged over time with participants with initially higher equitable beliefs in terms of the topics identified in their simulation responses. This suggests that these LE participants were applying similar approaches with HE participants in the simulations by the end of the course. This shift was corroborated by LE participants also expressing more equity-oriented mindsets and described engaging in more equity-promoting practices on surveys (both immediate postcourse and follow-up). Additionally, for one of the simulations, *Roster Justice*, we found that LE participants whose responses were most similar topically with HE participants were also more likely to be rated as equitable by human raters. These findings suggest that STM and NLP generally can be useful tools for analyzing text-based data from equity simulations and identifying changes in simulation behavior over time.

### *RQ1: Topic Modeling Within Equity Simulations and Associations With Equity Attitudes*

We examined the topics identified by the unsupervised STM model and analyzed their correlation with equity

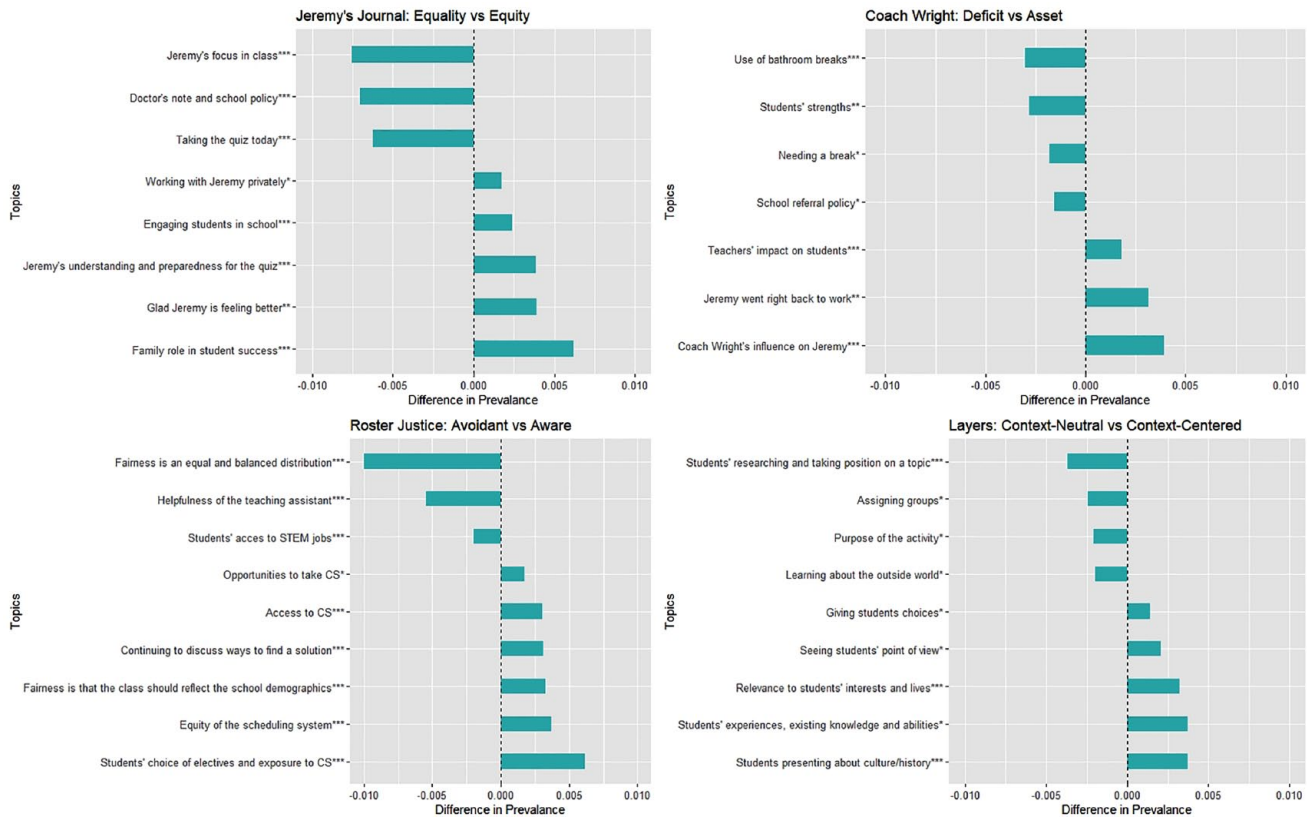


FIGURE 3. Correlations between equity mindsets and topic prevalence for selected topics.

attitudes. The topics often reflected meaningful choices the participants could make within the simulation. For example, Topic 11 (“Working with Jeremy in small groups”) in *Jeremy’s Journal* was related to responses about whether the teacher in the scenario should intervene, based on *Jeremy’s Journal* entry, by meeting with Jeremy and other students who seem confused in small groups (e.g., “By utilizing small groups, it gives the students more of a chance to share their thoughts and maybe bounce ideas off of each other). The full list of topics and labels can be found in the online Supplemental Appendix.

These findings suggest that STM detected patterns in simulation responses that were often associated with meaningful decision points in the simulation. These topics reflected meaningful differences in the mindsets that participants applied when responding to the fictional scenarios in the simulations. We measured this relationship by looking at the correlation between topic prevalence and participants’ responses to survey items about the specific dimension of equity measured within each simulation (e.g., Avoidant-Aware and *Roster Justice*). In Figure 3, we present the topics that were statistically significantly related to the survey items.

In the following sections, we briefly describe the key topics associated with equity mindsets in each of the

simulations and how they relate to mindsets about equity based on the existing literature.

*Jeremy’s Journal.* We found that certain topics were more likely to appear in the responses of participants based on the survey responses for the Equality-Equity mindset. Participants with more of an Equality mindset were more likely to mention the fact that school policy required students to submit a doctor’s note when they were sick (Topic 10, “Doctor’s note and school policy”). Participants with more of an Equality mindset are more likely to be concerned with giving all students equal treatment, regardless of students’ unique individual needs, for example, by equally enforcing school policies (Filback & Green, 2013). In contrast, participants with an Equity mindset are more likely to be concerned about meeting individual needs and taking into consideration factors outside of students’ control, such as thinking about Jeremy’s health (Filback & Green, 2013). Participants who demonstrated an Equity mindset were more likely to prioritize Jeremy’s health and well-being in response to receiving a note from his mother excusing him for being sick (Topic 7, “Glad Jeremy is feeling better”).

*Coach Wright.* Again, the differences in the prevalence of these topics often reflect differences in how participants

applied mindsets toward equity. In *Coach Wright*, participants with more of an Asset mindset were more likely to observe that Jeremy went right back to his work after taking his bathroom break (Topic 4, “Jeremy went right back to work”) and discuss the impact of Coach Wright on motivating Jeremy’s work (Topic 3, “Coach Wright’s influence on Jeremy”). In contrast, participants with more of a Deficit mindset were more likely to bring up the school’s referral policy for when students are wandering the hallways (Topic 10, “School referral policy”) and Jeremy’s long bathroom break (Topic 12, “Use of bathroom breaks”). Participants with more of a Deficit mindset were more likely to frame student behavior in terms of compliance with school rules (Battey & Franke, 2015; Horn, 2007). In contrast, participants with more of an Asset mindset were more likely to notice specific positive aspects of Jeremy’s interests, personality, and behavior and the supportive relationship with Coach Wright. Interestingly, mentioning the concept of leveraging student’s strengths was more likely to be associated with Deficit mindsets (Topic 2, “Student’s strengths”). However, many of these responses associated with this topic used the language of student strengths without actually referencing one of Jeremy’s strengths (e.g., “I advised Ms. Porter to recognize and point out Jeremy’s strengths”). This may be a case of “conceptual slippage” in the teaching where the meaning of the term changes from its original conception to meaning something entirely different in practice (Horn & Kane, 2019).

*Roster Justice.* The prevalence of certain topics was connected to certain aspects of the Avoidant-Aware mindsets. Participants with more of an Avoidant mindset recognized that the classes were imbalanced but they avoided solutions that explicitly named racial discrimination. They were more likely to describe difference in course composition in racially neutral terms without mentioning discrimination explicitly (Topic 10, “Fairness is an equal and balanced distribution”). Additionally, they were also more likely to agree with the principal’s offer of a teaching assistant alone to help balance the class loads (Topic 9, “Helpfulness of the teaching assistant”). Paradoxically, they were also more likely to mention how this imbalance would affect students’ access to jobs in STEM (science, technology, engineering, and mathematics) fields (Topic 5, “Students’ access to STEM jobs”). Although making connections to students’ future is more important, only focusing on career prospects within STEM fields without considering the racialized nature of how students understand their STEM identity is another example of Avoidant mindset (Collins, 2018). Many of the responses discussed the importance of STEM jobs without mentioning students’ racial identities (e.g., “What is the likelihood of those students just randomly ending up in tech industries if they don’t have any sort of a foundation?”).

Participants with more of an Aware attitude, by contrast, were more likely to see the racialized nature of the disparity. They were more likely to explicitly mention that the composition of the class did not match the school demographics (Topic 14, “Fairness is that the class should reflect the school demographics”). These teachers were more likely to note how stereotypes about who can be a computer scientist might influence whether students sign up for the course (Topic 2, “Access to CS [computer science]”; e.g., “there are a lot of stereotypes around who should study CS and who should not). Moreover, participants with more of an Aware mindset were also more likely to observe how relying on students to self-select into elective classes in high school can perpetuate inequities because students’ previous exposure to computer science would affect their choice of electives (Topic 15, “Students choice of electives and exposure to CS”). Participants with more of an Aware mindset were also more likely to mention that the current scheduling system perpetuates the racial imbalance (Topic 6, “Equity of the scheduling system”).

*Layers.* As with the other simulations, certain topics in *Layers* were related to the Context-Neutral and Context-Centered mindsets. Educators with more of a Context-Neutral mindset were less likely to view student needs in terms of their broader social identities or communities. Participants with more of a Context-Neutral mindset were more likely to describe ways of adapting content that did not incorporate aspects of students’ experiences, constraints, and responsibilities outside of school, such as reassigning groups (Topic 16 “Assigning groups”) or adjusting the amount of time given to the activity (Topic 9 “Time to complete the assignment”). Interestingly, participants with a more Context-Neutral mindset were more likely to discuss learning about culture and history in general terms (Topic 1, “Learning about the outside world”). Responses related to this topic, however, often did not mention incorporating the specific community where students were located (e.g., “I would have them focus on food, typical things found in houses around the world and clothing”).

In contrast, educators with more of a Context-Centered mindset believe students’ experiences in the world, and thus in school, are mediated through particular linguistic and cultural-historical practices, and it is important to include these in how they adapt their instruction (Gutiérrez & Rogoff, 2003; Nasir & Hand, 2006). Participants with more of a Context-Centered mindsets were more likely to mention drawing on students’ experiences in their communities and their home and family (Topic 8, “Students presenting about culture/history”); to mention making the lesson relevant to students’ interests and lives (Topic 5, “Relevance to students’ interests and lives”); drawing on students’ existing assets (Topic 4, “Draw on students’ existing experiences,

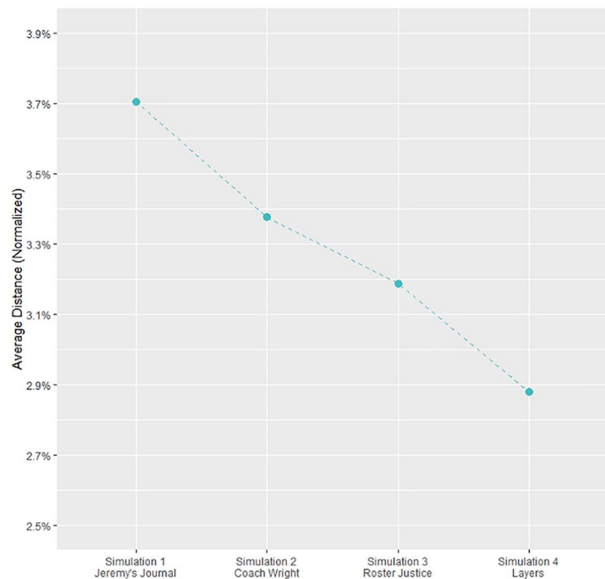


FIGURE 4. Changes in Euclidean distance from high-equity reference group.

knowledge, and abilities”); and including students perspectives in their lessons (Topic 14, “Seeing students’ of view”).

*Summary.* Across all four simulations, the STM model was able to successfully identify decision points within the simulations that were indicative of different mindsets toward equity in teaching. Many topics identified by the STM models within participants’ simulation responses were both theoretically linked to mindsets about equity and were correlated empirically with survey items about specific equity mindsets. This suggests that NLP tools such as STM, when applied to simulation responses, can provide insight into mindsets that participants are drawing from when responding within an equity simulation.

*RQ2: Changes in Simulations Response Behavior and Associations with Changes in Mindsets and Equity-Promoting Practices*

Next, we explored whether participants changed in their simulation responses over the four sequential simulations in the online DEI course. We found that as LE participants progressed through the course, they became more similar in their simulation responses to HE participants. We measured response similarity by looking at the average distance in topic prevalence using a Euclidean distance measure. In *Jeremy’s Journal*, LE participants were an average of 3.71 percentage points per topic from the reference group of HE participants. On *Coach Wright*, this distance decreased to 3.38 percentage points; in *Roster Justice*, to 3.19 percentage points, and finally to 2.88 percentage points in *Layers*, the final simulation (Figure 4). In terms of the topics identified

by the STM model, as LE participants progressed through the online course, their simulation responses became increasingly similar to those of HE participants.

We further extended this analysis to look at individual changes for LE participants compared with the HE reference group. LE participants significantly decreased in their average distance from the reference group for each topic by 0.64 percentage points ( $t=17.18, df=251, p<0.001, ES=1.08 SD$ ). Changes in simulation responses were consistent with shifts in equity attitudes and self-reported practices in surveys. LE participants significantly increased on the equity mindsets composite scale between the pre- and postsurvey by an average of 0.88  $SD$  ( $t=12.72, df=207, p<0.001$ ), and this increase largely persisted on the follow-up survey (0.60  $SD$ ;  $t=5.81, df=93, p<0.001$ ). These participants also significantly increased in their self-reported use of equity-promoting practices from the pre- to the post- and follow-up surveys (presurvey–postsurvey: 0.32  $SD, t=4.70, df=216, p<0.001$ ; presurvey–follow-up: 0.42  $SD, t=4.13, df=97, p<0.001$ ). Changes in equity mindsets and self-reported equity-promoting practices on surveys were thus consistent with changes in simulation responses toward more equitable behavior.

To further assess the validity of this method for measuring changes in simulation response behavior, we examined ratings given by human raters to participants’ responses in the *Roster Justice* scenarios. The average human rating for LE participants was correlated with their distance from the HE reference group ( $r=-0.18, df=198, t=2.52, p<0.05$ ). This suggests that simulations responses from LE participants that were closer topically to those from HE participants were also more likely to be rated as equitable by human raters. This indicates that as LE participants became more similar to the HE participants in their simulation responses, their responses were also likely to be rated as more equitable. This further suggests that the convergence in simulation responses between the HE and LE participants over the four simulations in the online course likely reflected the LE participants becoming more equitable in their simulation responses rather than the HE participants becoming less equitable.

**Discussion**

Although equity simulations are a potentially promising practice in DEI education (Robinson et al., 2018; Self, 2016), their utility as a teaching tool is somewhat constrained by the challenge of quickly interpreting and providing feedback on open-text responses. In this study, we explored the potential for using unsupervised NLP techniques to detect patterns within participants’ responses to prompts within equity teaching simulations. We hypothesized that the topics identified by the structural topic model would reflect different aspects of what teachers noticed and interpreted within the simulation and that these patterns would be associated

with survey measures of attitudes toward equity teaching. We also were interested in measuring changes in simulation responses of participants over the 8-week online course on DEI issues.

Overall, we found promising evidence for the use of NLP methods with simulations as a promising teaching tool for formatively assessing and evaluating equity-promoting teaching practices. Across all four simulations, the STM model identified discrete topics that reflected meaningful differences in what educators noticed and interpreted in the simulation. We also found preliminary evidence that topic modeling can be used to evaluate changes in simulation behavior over time. This suggests that NLP methods, such as STM, can be valuable tools when used in conjunction with equity teaching simulations.

The findings of this study point to a number of potential applications in the field of teacher education and in DEI training within educational settings more broadly. First, this study provides evidence supporting the use of simulation-based approaches to DEI training for educators. Although other studies have found benefits from using equity teaching simulations (e.g., J. A. Chen et al., 2020; Cohen et al., 2020; Self & Stengel, 2020), this is one of the first large-scale studies to detect changes in behavior within simulations. Digital simulation platforms in particular, such as *Teacher Moments*, have potential for wide impact because they are low-cost and relatively easy to customize and scale (Kaka et al., 2021). Cost and scalability are often important factors in selecting and implementing teacher professional learning (Kraft et al., 2018). As K–12 schools and educational organizations move toward incorporating more DEI-focused professional development, they should consider adopting digital equity simulations as part of a larger DEI professional learning strategy.

Second, this study suggests unsupervised models, such as STM, have the potential to automatically detect population-level patterns within simulation responses that can be used to create tools for teacher educators and professional development facilitators to interpret and evaluate data on teacher practice more easily. One of the barriers to scaling practice-based feedback in teacher education is the time and effort needed to manually analyze and provide feedback on individual practice (Peercy, 2014). Combining digital simulations with NLP tools can help instructors evaluate and interpret teacher behavior and provide feedback at scale without large increases in resources or cost. These patterns could also be disaggregated by demographic characteristics (e.g., gender, race), teaching experience, and exposure to coursework, which could help identify how different attributes are related to what participants noticed in the simulation. The results of these analyses could also be used to study, for example, if what participants noticed in the simulation changed in response to an intervention.

Third, the findings of this suggest the potential effectiveness of applying NLP tools to evaluate and provide feedback on individual educator performance within simulation. Although some researchers have proposed using NLP as an alternative to human observations of teaching (Jensen et al., 2020; Ramakrishnan et al., 2019); simulations allow for the collection of large-scale data without having to directly collect data from classrooms. The findings of this study suggest that NLP tools may be able to be integrated with digital simulations to provide automated feedback on participants' performance within the simulations. For example, in the *Jeremy's Journal* scenario, NLP could be used to detect when users are placing the majority of the blame on Jeremy for his performance in class. The system could then prompt the user to consider how external factors might play a role in Jeremy's performance and have them reattempt their response. Such an approach would work well within a large-scale online learning context, such as a MOOC, where individual feedback from instructors may not be logistically feasible. Automated feedback may allow simulation designers to tailor different feedback for participants with different levels of experience in discussing equity issues. For example, a less experienced user might receive recommendations on how inequitable educator mindsets affect student experiences in schools, while more experienced users would receive suggestions on how they can advocate for school-wide changes to inequitable school policies.

#### *Areas for Future Research*

This study is a preliminary investigation into how NLP models can be used to automatically detect patterns within participants' responses in equity teaching simulation responses among participants in an MOOC about equitable teaching. Although we targeted the course at a broad group of educators, because of the topic of the course, participants in this course were more likely to be interested in equity issues than a general educator audience. As a result, the findings may not transfer to other teacher populations. As a future direction, it would be informative to study the effect of these scenarios with the general K–12 teacher population. In subsequent work, we aim to study patterns among participants in other settings including teacher education courses, in-person professional learning, and nonequity-focused online courses. Additionally, because participants did not repeat any of the simulations, we cannot say conclusively whether the change in patterns between simulations was due to changes in attitudes or changes in the simulation. We will, in future analyses, study changes in participants' responses within the same or parallel simulations where changes in response can be more clearly linked to changes in behavior.

Finally, we aim to expand from the results of this study to explore the potential for automatic evaluation of individual participants' responses within equity teaching simulations.

Although the STM model was successful in identifying specific themes within participants' responses, in some cases it identified differences in topics that were not specifically related to equity. This may mean that the topics identified by the STM model may not be accurate enough to provide feedback for individual responses. In future work, we will explore using lightly supervised and hybrid approaches to improve the precision and accuracy of the NLP tools for evaluating individual responses without requiring extensive labeling.

### Conclusion

In education and every other sector of society, there is much work to be done to create a more just and equitable foundation for human flourishing. Education has an important role to play in helping people recognize how their biases can lead to prejudiced or discriminatory behavior while also helping them to change those behaviors. Given the scope of social change that needs to happen, scalable digital technology has the potential to play an important role in education around DEI.

The evidence from this study suggests promise for using digital simulations to prepare educators on DEI issues within large-scale asynchronous learning environments. We found initial evidence that NLP analyses of digital equity teaching simulation data can be used to formatively assess participants' use of equity teaching practices and capture changes over time. These findings expand the field's understanding of how technology can prepare educators to enact equitable teaching practices (J. A. Chen et al., 2020; Cohen et al., 2020; Okonofua et al., 2016) and offers tools to instructional designers and instructors develop and facilitate more effective DEI learning experiences.

### Authors' Note

Although coauthor Justin Reich is currently an associate editor of *AERA Open*, this submission was under consideration before he became the associate editor.

### ORCID iD

Joshua Littenberg-Tobias  <https://orcid.org/0000-0002-7204-3937>

### Note

1. An archived version of the course is available at <https://openlearninglibrary.mit.edu/courses/course-v1:MITx+0.503x+T2020/course/>

### References

Aldrich, C. (2009). *Learning online with games, simulations, and virtual worlds: Strategies for online instruction*. Wiley.

Anghel, E., & Littenberg-Tobias, J. (2020). *Measuring teachers' mindsets towards diversity*. EdArXiv. <https://doi.org/10.35542/osf.io/hdk4v>

Batthey, D., & Franke, M. (2015). Integrating professional development on mathematics and equity: Countering deficit views of students of color. *Education and Urban Society*, 47(4), 433–462. <https://doi.org/10.1177/0013124513497788>

Bautista, N. U., & Boone, W. J. (2015). Exploring the impact of TeachME™ lab virtual classroom teaching simulation on early childhood education majors' self-efficacy beliefs. *Journal of Science Teacher Education*, 26(3), 237–262. <https://doi.org/10.1007/s10972-014-9418-8>

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022. <https://doi.org/10.1016/b978-0-12-411519-4.00006-9>

Bonilla-Silva, E. (2015). The structure of racism in color-blind, “post-racial” America. *American Behavioral Scientist*, 59(11), 1358–1376. <https://doi.org/10.1177/0002764215586826>

Borneman, E., Littenberg-Tobias, J., & Reich, J. (2020). Developing digital clinical simulations for large-scale settings on diversity, equity, and inclusion: Design considerations for effective implementation at scale. In *L@S 2020: Proceedings of the 7th ACM Conference on Learning @ Scale*. Association for Computing Machinery. <https://doi.org/10.1145/3386527.3405947>

Borneman, E., Smith, B., & Littenberg-Tobias, J. (2020). *If you can't see color, you can't see patterns: Simulations raising awareness on structural inequity in K-12 computer science* [Unpublished manuscript]. Massachusetts Institute of Technology.

Bravo, M. A., Mosqueda, E., Solís, J. L., & Stoddart, T. (2014). Possibilities and limits of integrating science and diversity education in preservice elementary teacher preparation. *Journal of Science Teacher Education*, 25(5), 601–619. <https://doi.org/10.1007/s10972-013-9374-8>

Carter, E. R., Onyeador, I. N., & Lewis, N. A. (2020). Developing & delivering effective anti-bias training: Challenges & recommendations. *Behavioral Science & Policy*, 6(1), 57–70. <https://doi.org/10.1353/bsp.2020.0005>

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems* (pp. 288–296). <http://papers.nips.cc/paper/3700-reading-tea-leaves-how-humans-interpret-topic-models>

Chen, G. A. (2020). “That’s obviously really insensitive”: Attuning to marginalization in a parent-teacher encounter. *Cognition and Instruction*, 38(2), 153–178. <https://doi.org/10.1080/07370008.2020.1722128>

Chen, J. A., Tutwiler, M. S., & Jackson, J. F. L. (2020). Mixed-reality simulations to build capacity for advocating for diversity, equity, and inclusion in the geosciences. *Journal of Diversity in Higher Education*. Advance online publication. <https://doi.org/10.1037/dhe0000190>

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of *Computers & Education*. *Computers & Education*, 151, Article 103855. <https://doi.org/10.1016/j.compedu.2020.103855>

Civittillo, S., Juang, L. P., & Schachner, M. K. (2018). Challenging beliefs about cultural diversity in education: A synthesis and critical review of training with pre-service teachers. *Educational Research Review*, 24(June), 67–83. <https://doi.org/10.1016/j.edurev.2018.01.003>



- Cochran-Smith, M. (1995). Color blindness and basket making are not the answers: Confronting the dilemmas of race, culture, and language diversity in teacher education. *American Educational Research Journal*, 32(3), 493–522. <https://doi.org/10.3102/00028312032003493>
- Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher coaching in a simulated environment. *Educational Evaluation and Policy Analysis*, 42(2), 208–231. <https://doi.org/10.3102/0162373720906217>
- Collins, K. H. (2018). Confronting color-blind STEM talent development: Toward a contextual model for Black student STEM identity. *Journal of Advanced Academics*, 29(2), 143–168. <https://doi.org/10.1177/1932202X18757958>
- Dalinger, T., Thomas, K. B., Stansberry, S., & Xiu, Y. (2020). A mixed reality simulation offers strategic practice for pre-service teachers. *Computers & Education*, 144, Article 103696. <https://doi.org/10.1016/j.compedu.2019.103696>
- Datta, D., Phillips, M., Chiu, J., Watson, G. S., Bywater, J. P., Barnes, L., & Brown, D. (2021). *Improving classification through weak supervision in context-specific conversational agent development for teacher education*. arXiv.org. <https://arxiv.org/abs/2010.12710v1>
- Dotger, B. H., & Ashby, C. (2010). Exposing conditional inclusive ideologies through simulated interactions. *Teacher Education and Special Education*, 33(2), 114–130. <https://doi.org/10.1177/0888406409357541>
- Du, J., Song, D., & Liao, L. (2015, July). Topic modeling with document relative similarities. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)* (pp. 3469–3475). AAAI Press. <https://www.ijcai.org/Proceedings/15/Papers/488.pdf>
- Dutt, R., Hillaire, G., Fang, A., Larke, L., Rosé, C., & Reich, J. (2021). Investigating adoption and collaboration with digital clinical simulations by teacher educators. In E. Langran & L. Archambault (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 1209–1217). Association for the Advancement of Computing in Education. <https://www.learntechlib.org/primary/p/219277/>
- Ehrke, F., Ashoe, A., Steffens, M. C., & Louvet, E. (2020). A brief diversity training: Raising awareness of ingroup privilege to improve attitudes towards disadvantaged outgroups. *International Journal of Psychology*, 55(5), 732–742. <https://doi.org/10.1002/ijop.12665>
- Emdin, C. (2011). Moving beyond the boat without a paddle: Reality pedagogy, black youth, and urban science education. *Journal of Negro Education*, 80(3), 284–295.
- Ferry, B., Kervin, L., Cambourne, B., Turbill, J., Hedberg, J., & Jonassen, D. (2005). Incorporating real experience into the development of a classroom-based simulation. *Journal of Learning Design*, 1(1), 22–32. <https://doi.org/10.5204/jld.v1i1.5>
- Filback, R., & Green, A. G. (2013). *New directions for diversity at USC Rossier*. USC Rossier School of Education. <https://rossier.usc.edu/new-directions-for-diversity-at-usc-rossier/>
- Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education*, 53(2), 106–116. <https://doi.org/10.1177/0022487102053002003>
- Gibbons, L. K., & Cobb, P. (2016). Content-focused coaching: Five key practices. *Elementary School Journal*, 117(2), 237–260. <https://doi.org/10.1086/688906>
- Girod, M., & Girod, G. (2006). Exploring the efficacy of the cook school district simulation. *Journal of Teacher Education*, 57(5), 481–497. <https://doi.org/10.1177/0022487106293742>
- Grossman, P. L., Compton, C., Igra, D., & Williamson, P. W. (2009). Teaching practice: A Cross-professional perspective. *Teachers College Record*, 111(9), 2055–2100.
- Gutiérrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19–25. <https://doi.org/10.3102/0013189X032005019>
- Hillaire, G., Larke, L., & Reich, J. (2020, April). *Digital storytelling through authoring simulations with teacher moments*. In D. Schmidt-Crawford (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 1736–1745). Association for the Advancement of Computing in Education.
- Hirumi, A., Kleinsmith, A., Johnsen, K., Kubovec, S., Eakins, M., Bogert, K., Rivera-Gutierrez, D. J., Reyes, R. J., Lok, B., & Cendan, J. (2016). Advancing virtual patient simulations through design research and interplay: Part I: Design and development. *Educational Technology Research and Development*, 64(4), 763–785. <https://doi.org/10.1007/s11423-016-9429-6>
- Horn, I. S. (2007). Fast kids, slow kids, lazy kids: Framing the mismatch problem in mathematics teachers' conversations. *Journal of the Learning Sciences*, 16(1), 37–79. [https://doi.org/10.1207/s15327809jls1601\\_3](https://doi.org/10.1207/s15327809jls1601_3)
- Horn, I. S., & Kane, B. D. (2019). What we mean when we talk about teaching: The limits of professional language and possibilities for professionalizing discourse in teachers' conversations. *Teachers College Record*, 121(6).
- Jensen, E., Dale, M., Donnelly, P. J., Stone, C., Kelly, S., Godley, A., & D'Mello, S. K. (2020). Toward automated feedback on teacher discourse to enhance teacher learning. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI 2020)* (pp. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376418>
- Kaka, S. J., Littenberg-Tobias, J., Kessner, T., Francis, A. T., Kennett, K., Marvez, G., & Reich, J. (2021). Digital simulations as approximations of practice: Preparing preservice teachers to facilitate whole-class discussions of controversial issues. *Journal of Technology and Teacher Education*, 29(1), 67–90. <https://doi.org/10.35542/osf.io/95gyd>
- Kannan, P., Zapata-Rivera, D., Mikeska, J., Bryant, A., Long, R., & Howell, H. (2018). Providing formative feedback to pre-service teachers as they practice facilitation of high-quality discussions in simulated mathematics and science methods classrooms. In E. Langran & J. Borup (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 1570–1575). Association for the Advancement of Computing in Education. <https://www.learntechlib.org/primary/p/182737/>
- Kaufman, D., & Ireland, A. (2016). Enhancing teacher education with simulations. *TechTrends*, 60(3), 260–267. <https://doi.org/10.1007/s11528-016-0049-0>
- Kavanagh, S. S., & Danielson, K. A. (2020). Practicing justice, justifying practice: Toward critical practice teacher education.

- American Educational Research Journal*. <https://doi.org/10.3102/0002831219848691>
- Kavanagh, S. S., Metz, M., Hauser, M., Fogo, B., Taylor, M. W., & Carlson, J. (2020). Practicing responsiveness: Using approximations of teaching to develop teachers' responsiveness to students' ideas. *Journal of Teacher Education*, 71(1), 94–107. <https://doi.org/10.1177/0022487119841884>
- Kim, S. H., Lee, N., & King, P. E. (2020). Dimensions of religion and spirituality: A longitudinal topic modeling approach. *Journal for the Scientific Study of Religion*, 59(1), 62–83. <https://doi.org/10.1111/jssr.12639>
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3), 465–491. <https://doi.org/10.3102/00028312032003465>
- Love, B. L. (2019). *We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom*. Beacon Press.
- Matias, C. E., Montoya, R., & Nishi, N. W. M. (2016). Blocking CRT: How the emotionality of Whiteness blocks CRT in urban teacher education. *Educational Studies*, 52(1), 1–19. <https://doi.org/10.1080/00131946.2015.1120205>
- Milner, H. R. (2010). *Start where you are, but don't stay there: Understanding diversity, opportunity gaps, and teaching in today's classrooms*. Harvard Education Press.
- Milner, H. R. (2012). Beyond a test score. *Journal of Black Studies*, 43(6), 693–718. <https://doi.org/10.1177/0021934712442539>
- Mislevy, R. J. (2013). Evidence-centered design for simulation-based assessment. *Military Medicine*, 178(Suppl. 10), 107–114. <https://doi.org/10.7205/milmed-d-13-00213>
- Nasir, N. S., & Hand, V. M. (2006). Exploring sociocultural perspectives on race, culture, and learning. *Review of Educational Research*, 76(4), 449–475. <https://doi.org/10.3102/00346543076004449>
- Neville, H. A., Awad, G. H., Brooks, J. E., Flores, M. P., & Bluemel, J. (2013). Color-blind racial ideology theory, training, and measurement implications in psychology. *American Psychologist*, 68(6), 455–466. <https://doi.org/10.1037/a0033282>
- Neville, H. A., Lilly, R. L., Lee, R. M., Duran, G., & Browne, L. V. (2000). Construction and initial validation of the Color-Blind Racial Attitudes Scale (CoBRAS). *Journal of Counseling Psychology*, 47(1), 59–70. <https://doi.org/10.1037/0022-0167.47.1.59>
- Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences of the United States of America*, 113(19), 5221–5226. <https://doi.org/10.1073/pnas.1523698113>
- Paris, D. (2012). Culturally sustaining pedagogy. *Educational Researcher*, 41(3), 93–97. <https://doi.org/10.3102/0013189X12441244>
- Peercy, M. M. (2014). Challenges in enacting core practices in language teacher education: A self-study. *Studying Teacher Education*, 10(2), 146–162. <https://doi.org/10.1080/17425964.2014.884970>
- Ramakrishnan, A., Ottmar, E., LoCasale-Crouch, J., & Whitehill, J. (2019, May 1). Toward automated classroom observation: Predicting positive and negative climate. In *Proceedings of the 14th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2019* (pp. 1–8). IEEE. <https://doi.org/10.1109/FG.2019.8756529>
- Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, 16(9), 13–20. <https://doi.org/10.2307/1175725>
- Roberts, M. E., Stewart, B. M., & Airoidi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515), 988–1003. <https://doi.org/10.1080/01621459.2016.1141684>
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v091.i02>
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082. <https://doi.org/10.1111/ajps.12103>
- Robinson, K., Jahanian, K., & Reich, J. (2018). Using online practice spaces to investigate challenges in enacting principles of equitable computer science teaching. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education—SIGCSE '18* (pp. 882–887). Association for Computing Machinery. <https://doi.org/10.1145/3159450.3159503>
- Sayers, J., Petersson, J., Marschall, G., & Andrews, P. (2020). Teachers' perspectives on homework: Manifestations of culturally situated common sense. *Educational Review*. Advance online publication. <https://doi.org/10.1080/00131911.2020.1806786>
- Self, E. (2016). *Designing and using clinical simulations to prepare teachers for culturally responsive teaching* [Unpublished doctoral dissertation]. Vanderbilt University.
- Self, E., & Stengel, B. (2020). *Toward anti-oppressive teaching: Designing and using simulated encounters*. Harvard Education Press.
- Schiera, A. J. (2020). Seeking convergence and surfacing tensions between social justice and core practices: Re-presenting teacher education as a community of praxis. *Journal of Teacher Education*, 72(4), 462–476. <https://doi.org/10.1177/0022487120964959>
- Sherin, M. G., Russ, R. S., & Colestock, A. A. (2011). Accessing mathematics teachers' in-the-moment noticing (1st ed.). In M. G. Sherin, V. Jacobs, & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 79–94). Routledge. <https://doi.org/10.4324/9780203832714>
- Slater, M., Lotto, B., Arnold, M. M., & Sanchez-Vives, M. (2009). How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicología*, 40(2), 193–210. [www.redalyc.org/articulo.oa?id=97017660004](http://www.redalyc.org/articulo.oa?id=97017660004)
- Sleeter, C. E. (2012). Confronting the marginalization of culturally responsive pedagogy. *Urban Education*, 47(3), 562–584. <https://doi.org/10.1177/0042085911431472>
- Sterling, J., Jost, J. T., & Hardin, C. D. (2019). Liberal and conservative representations of the good society: A (social) structural

- topic modeling approach. *SAGE Open*, 9(2). <https://doi.org/10.1177/2158244019846211>
- Sullivan, F., Hillaire, G., Larke, L., & Reich, J. (2020). Using teacher moments during the COVID-19 pivot. *Journal of Technology and Teacher Education*, 28(2), 303–313. <https://www.learntechlib.org/primary/p/216171/>
- Theelen, H., van den Beemt, A., & den Brok, P. (2019). Classroom simulations in teacher education to support preservice teachers' interpersonal competence: A systematic literature review. *Computers & Education*, 129(February), 14–26. <https://doi.org/10.1016/j.compedu.2018.10.015>
- Thompson, M., Owho-Ovuakporie, K., Robinson, K., Kim, Y. J., Slama, R., & Reich, J. (2019). Teacher moments: A digital simulation for preservice teachers to approximate parent–teacher conversations. *Journal of Digital Learning in Teacher Education*, 35(3), 144–164. <https://doi.org/10.1080/21532974.2019.1587727>
- van de Oudeweetering, K., & Agirdag, O. (2018). Demographic data of MOOC learners: Can alternative survey deliveries improve current understandings? *Computers and Education*, 122(July), 169–178. <https://doi.org/10.1016/j.compedu.2018.03.017>
- Wang, X., Thompson, M., Yang, K., Roy, D., Koedinger, K., Rose, C., & Reich, J. (2020). *Practice-based teacher education with ELK: A role-playing simulation for eliciting learner knowledge*. EdArXiv. <https://doi.org/10.35542/osf.io/w8yzb>
- Yeomans, M., Stewart, B. M., Mavon, K., Kindel, A., Tingley, D., & Reich, J. (2018). The civic mission of MOOCs: Engagement across Political Differences in Online Forums. *International Journal of Artificial Intelligence in Education*, 28(4), 553–589. <https://doi.org/10.1007/s40593-017-0161-0>
- Zeichner, K. (2012). The turn once again toward practice-based teacher education. *Journal of Teacher Education*, 63(5), 376–382. <https://doi.org/10.1177/0022487112445789>

### Authors

JOSHUA LITTENBERG-TOBIAS is a research scientist at the Massachusetts Institute of Technology in the Teaching Systems Lab. His research focuses on learning interventions in a large-scale learning environment using computational measures to assess and evaluate learning.

ELIZABETH BORNEMAN is a designer, writer, and researcher interested in how art, computation, and communication can combine to strengthen community structures and enhance learning across learner background.

JUSTIN REICH is the Mitsui Career Development Professor of Comparative Media Studies and director of the Teaching Systems Lab at the Massachusetts Institute of Technology. He is interested in learning at scale, practice-based teacher education, and the future of learning in a networked world.