Impact of Simulation-Based Teacher Professional Development on Student Perceptions of Teaching Practices

Gerald Knezek
University of North Texas
United States
gknezek@gmail.com

Rhonda Christensen
Institute for the Integration of Technology into Teaching & Learning
United States
rhonda.christensen@gmail.com

Abstract: Assessment of the impact of teacher professional development is seldom accomplished by asking their students. This study addresses whether self-reported changes in teacher practices align with their students' perceptions of changes in teaching practices. Participants were 39 teachers from two US states that completed at least 15 teaching sessions totaling more than 3.5 hours of teacher professional development (practice teaching) inside the simulated teaching environment of simSchool. The goals of the professional development were remediation of implicit bias in teaching practices and fostering equity in teaching. Pre-post surveys were completed by the teachers before and after the professional development sessions. Concurrent pre and post surveys were administered to 800 of the teachers' G3-12 students. This study presents the results of examining whether teacher-reported changes in their teaching practices can be shown to align with changes reported by their students.

Keywords: simulated teaching, professional development, student engagement, culturally responsive teaching practices; equity and bias

Introduction

How we approach other people is related to the environment and experiences that impacted our development of thoughts and actions. We are all enveloped in some cultural norms, and culture is central to learning. Culture informs how we communicate with each other, the way we receive information as well as shaping the thinking process of groups and individuals. Teachers need the opportunity to reflect on previously held beliefs and the way in which it may impact their teaching pedagogy. Students in classrooms where teachers effectively incorporate culturally responsive practices show evidence of improved learning experiences (Hamdan, 2012). Culturally responsive educators adopt the view that all students are capable of the success shown to be critical to student growth (Boser, Wilhelm, & Hanna, 2014). Culturally responsive teaching is an approach that challenges educators to recognize that, rather than deficits, students bring strengths into the classroom that should be leveraged to make learning experiences more relevant and effective for students (Muniz, 2019).

There are few studies that compare teachers' biases to the impact on students (Chin et al., 2020). While explicit bias may exist in educators, implicit biases are the most difficult to recognize and reduce. People can hold implicit bias even though they do not consciously recognize the underlying attitude or stereotype that may exist (Devine, 1989) and cannot intentionally control the impact these biases have in their perception and judgement during decision making. Many biases are reproduced through socialization during formative years of growing up (Yogeeswaran, Devos, & Nash, 2017) and are deeply rooted in actions, phrases, mindset, and perceptions of ability.

While most people have some type of implicit bias, the impact in education is worthy of exploration as it will likely impact educators in their interactions with students and parents. Bias is most likely to occur from teachers whose students do not share their racial, cultural, linguistic, socio-economic or gender traits (Pasternak et al., 2023). Because biases are likely to impact equitable teaching practices, the biases need to be recognized and addressed.

Teacher Self-Efficacy's Impact on Students

Many research studies have found that a teacher's sense of self-efficacy was one of the variables highly related to student learning (Medgley, Feldlaufer, & Eccles, 1989; Tucker et al., 2005). One way that teachers can develop their self-efficacy is by understanding the needs of learners in the classroom with strategies to teach them. "Teachers who believe that student learning can be influenced by effective teaching despite home and peer influence and who have confidence in their ability to teach persist longer in their teaching, efforts, provide greater academic focus in the classroom, give different types of feedback, and ultimately improve student performance" (Tucker et al., 2005, p. 29). Soodak and Podell (1994) also found a relationship between teacher self-efficacy and their beliefs about and

actions toward difficult to teach students. Teachers with high self-efficacy were more likely to believe their teaching could impact student learning while teachers with low self-efficacy were more likely to look for solutions outside the classrooms (Soodak & Podell, 1994). In a study of teacher efficacy, researchers found that teacher self-efficacy for working with students of diverse backgrounds can be significantly increased by targeted training (Tucker et al., 2005. Researchers have identified connections between teachers' sense of efficacy, culturally responsive pedagogy (Callaway, 2016), and student achievement (Oyerinde, 2008; Tucker et al., 2005).

Simulated Teaching Environment

Simulated teaching tools can provide opportunities to experiment with different teaching strategies for a variety of students without harming any real student learning. In addition, the data collected in a simulation can be valuable when provided in an objective manner, with no judgement. The system provides feedback based on actual actions rather than intentions. These data can be used to provide insight for recognizing and reducing bias. Data analytics can provide participants with a lens in which to view their interactions with simulated students in a computer generated, objective way. How can implicit bias that exists be reduced by becoming aware of the bias and being provided with evidence to inform decisions? The simEquity project aims to use a simulated teaching environment to explore the relationship between implicit bias and teaching practices with the objective to reduce educator bias by using data analytic feedback captured by the simulation. This paper includes the impact of bias and equity focused teacher simulated training on students.

SimSchool, a simulated teaching environment, uses learning analytics (LA) to capture user interactions in the simulated classroom environment and displays visual data to participants so they can self-reflect on their performance, adapt their practices (Kovanovic et al., 2021) and complete multiple iterations of interactions with simStudents while adapting their teaching strategies. The main goal of developing better teaching practices through simulation is to improve student learning outcomes in real students.

Research Design

This research analyses focused on targeted teacher professional development aimed at leading to more equitable teaching practices with the goal of changes impacted by implementation of the teachers impacting the perceptions of their students. In the current context, the logic flow being studied is:

Two research questions were addressed in this study: 1) To what extent do changes in teacher practices due to simulated teaching professional development impact student perceptions of their teachers' diverse teaching practices, cultural engagement and student voice in learning?; and 2) To what extent are there differential contributions among teacher measures to positive student gains?

Intervention

The simEquity project was created to enable participants to practice teaching strategies in a simulated teaching environment with objective feedback intended to improve equitable teaching practices for real students. All participants completed the tutorial teaching module with two simulations to allow for learning the system before advancing to modules containing meaningful content. The modules were selected for different grade bands to be appropriate for the level of student taught by the participants. Within the modules, the student avatars reflect the appropriate age level as well. Each participant completed three content modules with five class sessions in each one (15 sessions total). The participants were required to review their feedback from each session before they were allowed to move forward. The minimum amount of time in each session was 15 minutes before feedback would be generated. Some teachers spent longer in each session than 15 minutes. By the time participants had completed all three modules, they had interacted in simSchool for a minimum of 225 minutes (3 hours, 45 minutes).

Participants

There were 39 participating teachers from two different school systems in the US. The participants included 28 females (71.8%) and 11 males (28.2%) who taught grades 3-12, with 11 teaching in elementary school, 10 in middle

school and 18 at the high school level. The ethnicity reported by the participants included 4 Asian, 2 Black/African American, 5 Latinx/Hispanic, and 28 White. In addition, 822 students of these 39 teachers, in grades 3 through 12, were asked by their teachers to complete pretest and post test surveys during class time. K-12 students were almost equally balanced between males (49%) and females (51%). Ethnic affiliations were primarily Hispanic (30%), White (25%), Asian (13%), and Black/AA (12%). The mean rating score across all students of a given teacher became a single variable entry in the record for each teacher.

Data Sources

Data were collected within the simulator to measure pre-post changes that may have occurred during the intervention. In addition to system data, survey data were also collected. Teacher survey measures focused on self-efficacy, culturally responsive teaching, and self-awareness of bias. Student survey data were also collected before and after their teachers' participation in the project. Below are descriptions of the teacher and student self-report surveys. Each of the surveys is rated on a 6-point Likert scale from Strongly Disagree (1) to Strongly Agree (6). The teacher surveys included:

1. The *Teachers' Sense of Efficacy Scale* (TSES) (Tschannen-Moran & Hoy, 2001) was used to measure self-efficacy related to three subscales: instructional strategies, classroom management, and student engagement.
2. The *Culturally Responsive Self-Efficacy Survey* (Siwatu, 2007) was included to determine the level of competency in the skills and knowledge needed to engage in culturally responsive teaching that includes curriculum, assessment, classroom management and cultural enrichment.
3. Three scales from the *Educator Bias Inventory* (Collum et al., 2020) were included. These scales include: *Self-Awareness*, *Pedagogical environment*, and *Relationships with families and communit*y. This survey was recently used with simSchool research and is based on Chen et al. (2009).

Student measures focused on dispositions related to learning (student engagement, cultural identity, student voice).

- The *Student Engagement Inventory* (SEI) (Appleton, Christenson, Kim, & Reschly, 2006) measures the cognitive and psychological aspects of engagement. For this study, one of the six subscales (9 items), Teacher-Student Relationships, were administered.
- Two scales from the *Student Measure of Culturally Responsive Teaching* (Dickson, Chun & Fernandez, 2016) was administered to the students to measure their perceptions of their teachers' level of culturally responsive teaching. The first scale addresses the construct of Diverse Teaching Practices while the second scale assesses Cultural Engagement.
- Two scales representing independent constructs from the Student Participation Survey (SPS) (Anderson, Graham & Thomas, 2019) was administered to students. Voice about Schooling and Having influence will allow researchers to compare data regarding how students feel about having choices in their classrooms.

Results

Table 1 contains means and standard deviations for the pre-post gain scores for teacher survey scales, as well as scores for the average gains across all student survey ratings by students for each of the teachers. Gain scores represent the difference between pretest and post test ratings. Item ratings were averaged across all items contributing to a specified scale to produce a scale score for a single individual.

As shown in Table 1, the mean pre-post gains for teacher scales ranged from -.04 for AA Gain to +.40 for EE Gain. For students, the pre-post gains ranged from -.02 to +.12. Standard deviations varied widely for teacher gains, from .33 to 1.23, while they ranged from .42 to .59 for students. The sample size of N = 36 teachers represents the actual data set used for canonical correlation and regression analyses in latter sections of this paper, after three of the 39 teachers were removed as outliers (> 3 standard deviations).

Table 1. Descriptive Statistics for Pre-Post Gains in Teacher and Student Rating Scales

	N*	Mean	Std. Dev.
Teacher Measures			
Culturally Responsive Self-Efficacy (CR Gain Score)	36	.2444	.43992
Educator Bias Inventory Self-Awareness (EBSA Gain Score)	36	.0556	.42621
Educator Bias Inventory Pedagogical Environment (EBPE Gain Score)	36	.1706	.33077
Educator Bias Inventory Relationship with Family (EBRF Gain Score)	36	.3148	.64707
Teacher Efficacy Instructional Strategies (TEIS Gain Score)	36	.3090	.43351
Teacher Efficacy Classroom Management (TECM Gain Score)	36	.1806	.48499

Teacher Efficacy Student Engagement (TESE Gain Score)	36	.2951	.57333
Equality/Equity (EE Gain Score)	36	.4000	1.22824
Avoidant/Aware (AA Gain Score)	36	-.0389	.67919
Locus of Control (LC Gain Score)	36	.2056	.88993
Reflection Gain Score	36	.1667	1.00000
Student Measures			
Diverse Teaching Practices (Student Diverse Teaching Gain)	33	.0948	.42302
Cultural Engagement (Student Cultural Gain)	33	-.0161	.50472
Voice About School (Student Voice Gain)	33	.0274	.58868
Voice Having Influence (Student Voice Influence Gain)	32	.1166	.43386
Student Engagement (Student Engagement Gain)	33	.0939	.45338

*Note: Dataset reduced from 39 to 36 due to outliers.

Higher Order Factor Analysis of Teacher Scale Gains

A higher-order factor analysis was conducted on the pre-post gain scores for the teachers, in order to search for higher-order constructs that would simplify the number of concepts simultaneously under consideration. Based on the 10 teacher measures, analysis confirmed that the extraction of three higher order factors could account for 69% of the common variance among the scales, with the distribution among the higher order constructs being fairly evenly balanced at 29% for HF1, 21% for HF2, and 19% for HF3. The rotated pattern matrix shown in Table 2 indicates that TEIS gain score, TECM gain score, and TESE gain score belong to HF1, while LC gain score, AA gain score, and EE gain score belong to HF3. Not all items loading on HF2 achieved the desirable strength of association of >= .5 for pattern loadings, so the weakest item was excluded from HF2. The representatives retained for HF2 then became EBPE gain score, EBSA gain score, and CR gain score; EBRF was omitted.

Table 2. Factor Loadings for Three Higher Order Constructs Extracted from 10 Teacher Scales

	Component		
	1	2	3
Teacher Efficacy Instructional Strategies (TEIS Gain Score)	.964		
Teacher Efficacy Classroom Management (TECM Gain Score)	.948		
Teacher Efficacy Student Engagement (TESE Gain Score)	.940		
Educator Bias Inventory Pedagogical Environment (EBPE Gain Score)		.876	
Educator Bias Inventory Self-Awareness (EBSA Gain Score)		.756	
Culturally Responsive Self-Efficacy (CR Gain Score)		.673	
Educator Bias Inventory Relationship with Family (EBRF Gain Score)		.386	
Locus of Control (LC Gain Score)			.802
Avoidant/Aware (AA Gain Score)			.779
Equality/Equity (EE Gain Score)			.766

Note. Extraction method: Principal component analysis. Rotation method: Varimax with Kaiser normalization (Rotation converged in 5 iterations. Loadings <.3 suppressed.

Examination for Outliers

HFA scale scores (Z-scores) were then produced for each of the three principal components shown in Table 2. These HF1, HF2, and HF3 factor scores represent the values each teacher would have scored on each higher-order construct if all error was removed. Factor scores were produced primarily to construct a small number of variables (3) whose values represented most of the gains (69%) recorded across the original 10 teacher scales. Examination of HF1-3 scores through descriptive statistics revealed three outliers in the data, based on the criteria of > 3 standard deviations from the mean. Two cases with outliers in HF1 (-4.17 and -3.67) and one case with an outlier in HF3 (+3.13) were removed from further analyses. This reduced the data set from 39 to 36 teachers for subsequent analyses of the relationship of teacher reported gains to reported student gains.

Canonical Correlation Analysis: Teacher Gains with Student Gains

Canonical correlation was used to assess whether gains in student ratings were significantly ($p < .05$) associated with gains in teacher ratings. Exploratory canonical analyses (not shown) that included all teacher gain variables and all student gain variables indicated that approximately 19% of the gains in student measures could be accounted for by gains in teacher measures. Broadscale analyses also yielded some indications of which variables on the teacher side

and on the student side were the strongest contributors. However, the relatively small number of teachers in the study, combined with the large number of variables being studied, restricted the number of degrees of freedom in the model being tested and made it difficult for any of the extracted solutions to reach significance. Therefore, the analysis featured in this study examined teacher gain variables reduced to the set of HF1, HF3, and Reflection Gain which is a single item representing a construct found to be related to positive student perceptions of teaching practices in earlier studies such as Tondeur et. al (2016). Student gain variables were reduced in number from five to three, by selecting those with the strongest positive correlations in the broadscale analyses: Student Diverse Teaching Practices Gain, Student Cultural Engagement Gain, and Student Voice Gain. Since three variables were included on each side of the canonical correlation model, three independent canonical solutions could be derived. As shown in Table 3, the resulting canonical solution 1 indicated there is a significant (p = .033) association between the specified collection of teacher gain variables and the specified collection of student gain variables.

Table 3. Three Canonical Solutions for the Association of Gains Reported on Teacher Scales with Gains Reported on Student Scales

	Canonical Correlations						
	Correlation	Eigenvalue	Wilks Statistic	F	Num D.F	Denom D.F.	Sig.
1	**.620**	**.625**	**.527**	**2.202**	**9.000**	**65.862**	**.033**
2	.377	.166	.856	1.129	4.000	56.000	.352
3	.042	.002	.998	.052	1.000	29.000	.822

Since only Solution 1 from Table 3 yielded a significant ($p < .05$) association, attention was then focused on the canonical loadings in the first column of Tables 4 and 5. All loadings for the teacher set and the student set show non-trivial, positive associations with their underlying canonical construct.

Table 4. Canonical Loadings for Gains on Three Teacher Scales

Variable	**1**
HF1	**.425**
HF3	**.520**
Reflection Gain	**.592**

Table 5. Canonical Loadings for Gains on Three Student Scales

Variable	**1**	2	3
Student Voice Gain	**.296**	.697	.654
Student Cultural Engagement Gain	**.752**	-.106	.650
Student Diverse Teaching Practices Gain	**.919**	.394	-.028

The last step in canonical correlation is computation of the percentage of variance in the theorized outcome variable(s) that can be explained by the theorized input variable(s). As shown in the last column in Table 6, the proportion of variance in Set 2 (Student Gains) explained by the combination of the variables in Set 1 (Teacher Gains) is 19.2% for Solution 1. This indicates that approximately 1/5 (19%) of the gains in student survey ratings from pre to post can be explained by gains in reported teacher survey ratings, pre to post.

Table 6. Proportion of Variance Explained

Canonical Variable	Set 1 by Self	Set 1 by Set 2	Set 2 by Self	**Set 2 by Set 1**
1	.267	.103	.499	**.192**
2	.366	.052	.217	**.031**
3	.367	.001	.284	**.001**

Results of the canonical correlation analyses provide the answer to Research Question 1: To what extent do changes in teacher practices due to simulated teaching professional development impact student perceptions of their teachers' diverse teaching practices, cultural engagement and student voice in schooling? The answer is that approximately 1/5 (19%) of gains in student perceptions of the teaching practices can be attributed to self-reported gains by their teachers. Research Question 2 will be addressed through regression analysis, as explained in the following sections.

Regression Analyses Predicting Student Reported Gains from Teacher Self-Reported Gains

Regression analyses targeted at each measure of student gain included in the canonical correlation of the previous section were conducted to determine if specific types of teacher gains identified through higher-order factor analysis (HF1, HF2, HF3) were associated with student gains in the three key student areas identified through canonical correlation: Diverse Teaching Practices (CRTP subscale), Cultural Engagement (CRTP subscale), and Voice in Schooling. The independent analyses and findings are presented in the order of the greatest percentage of variance explained, in the sections that follow.

Regression Analysis for Student Perceptions of Diverse Teacher Practice Gains Predicted from Teachers' Gains

A linear regression predicting Student Diverse Teaching Practices Gain as function of HF1, HF2, HF3 was conducted to determine whether this category of student reported gain could be explained by one or more of the three types of teacher gains distilled into the three higher order factors shown in Table 2. A shown in Table 7, findings were that 29% (RSQ = .291) of Student Diverse Teaching Practices Gain could be explained through a linear combination of HF1, HF2, and HF3. The overall regression model was significant at the $p < .05$ level ($p = .018$). Both HF3 (beta = .552) and HF1 (beta = .349) were also significant ($p < .05$) individual contributors to the prediction equation.

Table 7. Regression Analysis for Student Diverse Teaching Practices Gain as a Function of Three Higher Order Factors of Teacher Reported Gains

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.539[a]	.291	.217	.37427

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.664	3	.555	3.960	.018[b]
	Residual	4.062	29	.140		
	Total	5.726	32			

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.040	.077		.522	.605
	REGR factor score 1 for analysis 1	.422	.208	.349	2.029	.052
	REGR factor score 2 for analysis 1	.018	.065	.045	.279	.782
	REGR factor score 3 for analysis 1	.272	.083	.552	3.279	.003

Regression Analysis for Student Perceptions of Cultural Engagement Gain Predicted from Teachers' Gains

A linear regression predicting student perceptions of teacher cultural engagement as a function of HF1, HF2, HF3 for teachers was also conducted to determine whether this category of student reported gain could be explained by one or more of the three types of teacher gains distilled through the higher order factor analysis presented

previously. Findings were that 16% of gains in student perceptions of cultural engagement could be explained through a linear combination of HF1, HF2, and HF3. However, the overall regression model was not significant (NS) at the $p < .05$ level ($p = .170$), so we must conclude it is possible these associations could have occurred by chance. Both HF1 (beta = .358) and HF3 (beta = .338) were borderline for being labeled as individually significant ($p < .05$) contributors, with p levels of .066 and .076, respectively.

Regression Analysis for *Student Voice in Schooling Gains Predicted from Teachers' Gains*

A linear regression predicting student voice in schooling gain as function of HF1, HF2, HF3 was also conducted to determine whether this category of student reported gain could be explained by one or more of the three types of teacher gains distilled through the higher order higher order factor analysis presented in Section 1. Findings were that 13% of student voice gain could be explained through a linear combination of HF1, HF2, and HF3. However, the overall regression model was not significant at the $p < .05$ level ($p = .262$), so we must conclude it is possible these associations could have occurred by chance. HF3 (beta = .335) approached being labeled as an individually significant ($p < .05$) contributor ($p = .083$).

Summary of Findings Across Regression Analyses

The trends across the three reported regression analyses were that HF1 and HF3 were common contributors to gains in individual student measures. On the other hand, HF2 never stood out as a significant ($p < .05$) contributor to gains in student pre-post measures. These findings imply that the areas of teacher gains that influenced changes perceived by their students lie primarily in the realm of increased teaching self-efficacy (confidence in competence), which is dominant in HF1, and in beliefs that all students can succeed (Locus of Control scale) plus intentionality toward equitable rather than simply equal teaching practices (Equality-Equity scale), both of which were strong contributors to HF3. (Avoidance-Awareness was also aligned with HF3 but showed very small gain, ES = .04, pre to post.) These findings provide the answer to Research Question 2: To what extent are there differential contributions among teacher measures to positive student gains? The answer is that there appear to be extensive differences in the areas of teacher professional development that contribute to different types of gains as perceived by their students. This topic will be further addressed in Discussion.

Discussion

As shown in Table 2, HF1 consists of the teacher scales TEIS gain score, TECM Gain Score, and TESE Gain Score, all from The Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran & Hoy, 2001). TSES measures teacher self-efficacy on three subscales: Instructional Strategies (TIES), Classroom Management (TECM), and Student Engagement (TESE). Table 2 indicated that the average gain scores on these three measures for the teachers in this study were all positive, and specifically TIES = .3090 ($p < .001$), TECM = .1806 ($p = .032$) and TESE = .2951 ($p = .004$). As shown in Table 8, student pre-post rating gains for Diverse Teaching Practices increased from a pretest mean of 4.55 to 4.67 ($p = .021$), pre to post.

Table 8. simEquity Y2 Pre and Post Mean Ratings by Students on Diverse Teaching Practices Scale

Measure		N	Mean	Std. Dev.	Sig	ES
Diverse Teaching Practices (CRTP subscale)	Pre	745	4.55	.83		
	Post	530	4.67	.85		
	Total	1275	4.60	.84	.021	.14

These findings together with the regression analysis shown in Table 7 provide empirical evidence that increasing teacher self-efficacy (confidence in their competence in teaching practices) contributes to measurable self-reported gains in Diverse Teaching Practices (CRTP subscale) as assessed by their students. Students provided more positive ratings at post test time, after the simSchool teacher professional development interventions, than at pretest time before the professional development began, on items such as *My teacher treats all students like they are important members of the classroom*; *My teacher uses real-life examples to help explain things*; *My teacher uses what I already know to help me understand new ideas;* and *My teacher wants parents to be involved in student learning.* A second significant ($p = .003$) contributor to Diverse Teaching Practices shown in the regression analysis of Table 7 was HF3. This higher-order construct (factor score) was a composite of the scales LC Gain Score, AA Gain Score and EE Gain Score (see Table 2). Locus of control (LC) was adapted from prior studies (Christensen et al., 2011) related to simSchool and used to measure the sense of control educators feel in changing their classroom environment or reaching difficult students. The higher the locus of control, the more strongly responders feel they

can make changes in situations in their classrooms. The other two scales, Avoidant vs. Aware (AA) and Equality vs. Equity (EE) were developed by Littenberg-Tobias, Borneman, & Reich (2021) to measure equity-promoting behaviors in digital teaching simulations. An equality perspective indicates that all students should be treated the same, whereas an equity perspective indicates that students should be provided resources based on their needs. Regarding the avoidant-aware scale, an avoidant perspective avoids mentioning or considering race in order to be racially unbiased whereas an aware perspective acknowledges the role race plays in students' experiences in schools and seeks to explicitly name and actively remove systemic practices that cause racial inequity. This HF3 predictor was actually stronger (beta = .552, Table 7) when viewed individually while holding the other two predictors constant, than was the case for HF1 (beta = .349), whose substantial contributions attributed to gains in teacher self-efficacy were described in the previous paragraph. The strength of the contribution of HF3 (beta = .552, p = .003) to pre-post gains in student ratings of Diverse Teaching Practices provides empirical evidence that the simSchool-based professional development completed by the teachers in the simEquity project for the 2022-2023 school year were effective in contributing to the primary simEquity goal of reduction in implicit biases in teaching practices.

There are several limitations to this study. First, the sample size of teachers is relatively small (N = 39, reduced to N = 36 in this study due to removal of extreme outliers). Second, there is no true (randomly-assigned) control group that completed all of the instruments pre and post, without participating in the simSchool-based professional development. Third, the significant ($p < .05$) positive effects reported here are for relationships that stood out as strong among the many in the teacher and K-12 student data sets that were possible. Key findings need to be reconfirmed with new and larger data sets. In part in recognition of these limitations, the simEquity research team is planning a 50% larger teacher pool for the 2023-2024 cycle of simEquity activities. The larger sample size may enable reconfirmation of major findings from 2022-2023 and also enable confirmation of additional noteworthy associations, such as those implied by the (NS) positive associations indicated in the regression analyses of Tables 7-9 in this paper.

Conclusions and Implications for Teacher Education

Simulations hold many possibilities as a pedagogical approach for teacher professional development related to equity-based teaching practices and are increasingly being used to approximate various teaching scenarios and support the transfer of learning into classroom situations (Dalinger, Thomas, Stansberry, & Xiu, 2020). Most research on simulations for teacher education that focus on equitable teaching practices include human actors (Cohen, Wong, Krishnamachari, & Berlin, 2020), a type of simulation that is not affordable or sustainable for large groups of educators. SimSchool provides a fully digital environment for supporting the improvement of teacher practices related to equity.

Findings from this study provide empirical evidence that the simSchool-based professional development designed and delivered through the NSF-supported simEquity project were effective in contributing to the reduction in implicit biases in teaching practices. Specifically, canonical correlation and regression analyses indicated that teacher gains in self-efficacy, locus of control, and equity (vs. simple equality) resulted in gains in diverse teaching practices, as assessed by their students.

Acknowledgement: This research was funded in part by NSF Grant # 2118849.

References

Anderson, D.L., Graham, A.P., & Thomas, N.P. (2019). Assessing student participation at school: Developing a multidimensional scale. *International Journal of Student Voice, 5*(1), 1-24.

Appleton, J.J., Christenson, S.L., Kim, D., & Reschly, A.L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of Psychology, 44*, 427-445.

Bialo, E. & Sivin-Kachala, J. (1996). The effectiveness of technology in schools: A summary of recent research. *School Library Media Quarterly, 25*(1), 51-57.

Boser, U., Wilhelm, M., & Hanna, R. (2014). *The power of the pygmalion effect: Teachers' expectations strongly predict college completion*. Washington DC: Center for American Progress. https://files.eric.ed.gov/fulltext/ED564606.

Chen, D.W., Nimmo, J., & Fraser, H. (2009). Becoming a culturally responsive early childhood educator: A tool to support reflection by teachers embarking on the anti-bias journey. *Multicultural Perspectives, 11*(2), 101-106. doi: 10.1080/15210960903028784

Chin, M.J., Quinn, D.M., Dhaliwal, T.K., & Lovison, V.S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher, 49*(8), 566-578.

Christensen, R., & Knezek, G. (2022). Using digital simulation to address implicit bias in teaching. In E. Langran (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 477-483). San Diego, CA, United States: Association for the Advancement of Computing in Education (AACE).

Christensen, R., Knezek, G., Tyler-Wood, T., & Gibson, D. (2011). SimSchool: An online dynamic simulator for enhancing teacher preparation. *International Journal of Learning Technologies.* 6(2), 201-220.

Cohen, J., Wong, V., Krishnamachari, A., & Berlin, R. (2020). Teacher coaching in a simulated environment. *Educational Evaluation and Policy Analysis, 42*(2), 208-231. Doi.org/10.3102/0162373720906217

Collum, D., Christensen, R., Delicath, T., & Knezek, G. (2020). Measuring changes in educator bias in a simulated learning environment. In G. H. Marks & D. Schmidt-Crawford (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 410-416). Online: Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/215788/

Dalinger, T., Thomas, K.B., Stansberry, S., & Xiu, Y. (2020). A mixed reality simulation offers strategic practice for pre-service teachers. *Computers & Education, 144*. https://doi.org/10.1016/j.compedu.2019.103696

Devine, P.G. (1989). Stereotypes and prejudice: Their automatic controlled components. *Journal of Personality and Social Psychology, 82*, 62-68.

Dickson, G.L., Chun, H., & Fernandez, I.T. (2016). The development and initial validation of the student measure of culturally responsive teaching. *Assessment for Effective Intervention, 41*(3), 141-154.

Littenberg-Tobias, J., Borneman, E., & Reich, J. (2021). Measuring equity-promoting behaviors in digital teaching simulations: A topic modeling approach. *AERA Open, 7*(1), 1-19. doi.org/10.1177/23328584211045685

Muniz, J. (2019). *Culturally responsive teaching: A 50-state survey of teaching standards.* http://newamerica.org/education-policy/reports/culturally-responsive-teaching

Pasternak, D.L., Harris, S.D., Lewis, C., Wolk, M.A., Wu, X., & Evans, L.M. (2023). Engaging culturally responsive practice: Implications for continued learning and teacher empowerment. *Teaching and Teacher Education, 122*. Doi:10.1016/j.tate.2022.103976

Sianjina, R.R. (2000). Educational technology and the diverse classroom. *Kappa Delta Pi Record*.

Siwatu, K.O. (2007). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education, 23*, 1086-1101.

Tondeur, J., van Braak, J., Siddiq, F., & Scherer, R. (2016). Time for a new approach to prepare future teachers for educational technology use: Its meaning and measurement. *Computers & Education, 94*, 134-150.

Tschannen-Moran, M., & Hoy, A.W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.

Yogeeswaran, K., Devos, T., & Nash, K. (2017). Understanding the nature, measurement, and utility of implicit intergroup biases. In C. G. Sibley & F. K. Barlow (Eds.), *The Cambridge handbook of the psychology of prejudice* (pp. 241–266). Cambridge University Press. https://doi.org/10.1017/9781316161579.011