

Year Two Report for the SimEquity Project: Changes in Student and Teacher Dispositions Regarding Culturally Responsive Teaching Strategies

Rhonda Christensen, Ph.D.
August 2023

Project Goals

The overarching goal of the National Science Foundation funded *simEquity: Improving Student Learning While Decreasing Bias in Teaching Through Simulation* project is to implement a scalable model for developing equitable, culturally responsive teaching practices through simSchool, a simulated teaching environment. The project focused on identifying best practices to help teachers recognize and mitigate implicit biases that often impact student success. The project team is in the initial stages of iteratively developing and testing an AI-driven set of bias reduction tools that can be added to existing simulation instruction modules.

For year two, an invitation was sent to all teachers in the Aspire Public Schools from three regions in California: Central Valley, Bay Area and Los Angeles. In addition, a high school from Northwest Independent School District in Texas provided participants for year two of the project. Each selected participant was sent a link to select from the optional days/times for one hour training times to introduce the research project and simSchool.

Because it is a research project, data from survey as well as simulation gathered data is important to be able to establish valid modules that lead to improvement in equitable teaching practices. Thirty-nine teachers were able to complete the training, pre and post test self-surveys, required modules and student pre and post surveys by the end of May 2023. The data from the 39 teachers and their students are included in this report. Teachers received a stipend for their participation.

In subsequent years, teachers who were part of year one and/or year two may continue in year three. In addition, a much larger sample (90) of teachers will be selected to participate and receive the stipend for year three. In addition to the stipend, the participating teachers' schools have access to simSchool modules for professional development on a large range of topics.

Background

SimSchool is a dynamic, online classroom simulation program that allows preservice and inservice teachers the opportunity to practice teaching. SimSchool was designed to provide future and current teachers with a safe environment for experimenting and practicing techniques, especially methods of addressing different learning needs, and wide variations in academic and behavioral performance of students. Using student profiles, teachers need to be able to plan and deliver culturally relevant instructional challenges and supports that build on the strengths of students to address their learning needs (Sianjina, 2000).

Research on the use of simSchool has shown improved educator understanding in *teaching skills* (Christensen, Knezek, Tyler-Wood, & Gibson, 2011; Knezek, Hopper, Christensen, Tyler-Wood, & Gibson, 2015), *classroom management* (Christensen et al., 2007), *motivation* (Tyler-Wood, Estes, Christensen, Knezek, & Gibson, 2017), *multicultural awareness*, *literacy* (Collum, Christensen, Delicath, & Johnston, 2019), self-reported *educator bias* (Collum, Christensen, Delicath, & Knezek, 2020) and *instructional self-efficacy* (Knezek & Christensen, 2009). The key innovation of the program is that it provides teachers and teacher trainees many learning trials with simulated students, thereby increasing teacher confidence and competence,

which in turn improves student learning. Repetition of many trials is important in changing habit complexes such as implicit bias (Malone, 2016).

As shown in Figure 1, simSchool users have options to review detailed student profiles before or during any session. Through the “Teach” button, users select which students to assign activities, either as individuals, groups, or as a whole class. The simSchool user can also make different types of comments through the “Talk” button in which they can encourage, redirect or discipline students. The “Progress” button allows the user to see how each student is performing in many dimensions including academics and emotions. The “Logs” button allows the user to see how the class and each individual student performed for each action the user selected including tasks and comments. These features provide a robust system that allows each user to “pause” the simulator to assess whether to make changes in the activities or comments. Users can also see the overall performance in the simulator after it is complete. Users are able to reflect and make corrections prior to beginning another module with the intention of improving their teaching performances.

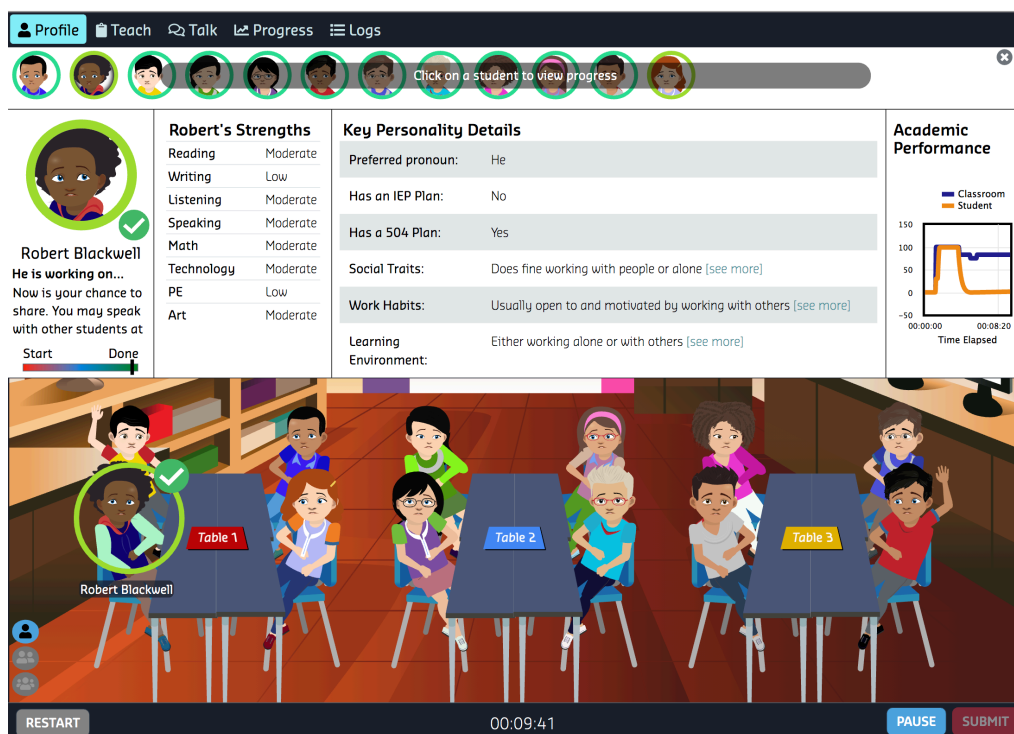


Figure 1. SimSchool classroom highlighting student profiles.

The simEquity project is using the simulation program to help educators recognize possible bias with the goal of recognizing, reflecting and remediating any biases that may exist.

Activities

Classroom teachers in several schools in a school system in California and a high school in Texas participated in one of three sets of modules depending on the grade level they teach. Each participant first completed an introductory module to understand how to navigate simSchool. The listing of the modules by grade level band is shown in Table 1.

Table 1. *Modules Completed by Classroom Teachers*

Elementary School Modules

- Tutorial: Teaching in simSchool
- Module 1: Cultural Intelligence and Inclusion 2.0
- Module 2: Bullying and Bias the First Coconut Tree
- Module 3: Gender and Identity: Supermom Saves the Day

Middle School Modules

- Tutorial: Teaching in simSchool
- Module 1: Gender and Identity: The Misfits
- Module 2: History Empowering Learners to Change the world
- Module 3: Race, Ethnicity, Class, Immigration: A Tale of Two Schools

High School Modules

- Tutorial: Teaching in simSchool
- Module 1: Showing Empathy
- Module 2: Sounds of Change
- Module 3: Why Local Elections Matter

At the end of each simulation session, participants receive graphical feedback displaying degree of success at promoting academic (learning) increase in the class overall, as well as feedback regarding the degree of suitability of the instructional activities selected for each individual simulated student in the class. Among the aspects of instructional activities that are documented for review are impacts on individual students of conversational stances, communication patterns, and attentional habits of the teacher. Figures 2 and 3 illustrate examples of graphical feedback as well as an observation report that participants receive. Participants must view the feedback prior to completing another session in the module. Figure 4 illustrates the equity index that has been developed for this project. This feedback can reveal aspects of a teacher’s implicit biases for gender and/or ethnicity.

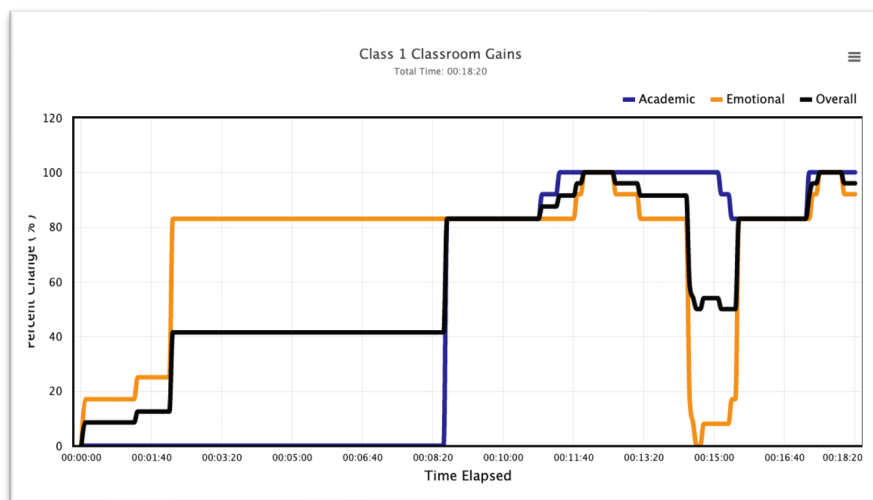


Figure 2. SimSchool graphical feedback based on interactions with students.

	Description	Recommended Practice	N/A	Needs Attention	Satisfactory	Good
1	Regularly assesses individual and group performance.	Checks the class progress tab at a regular interval.			✓	
2	Uses information of differing student strengths and needs to further each learner's development.	Clicks on each student's profile and views their traits or progress at least once throughout the simulation.				✓
3	Designs, adapts, and delivers instructions to address each student's diverse learning strengths and needs and creates opportunities for students to demonstrate their learning in different ways.	User uses a variety of tasks during the simulation.				✓
4	Makes appropriate and timely provisions for individual students with particular learning differences and needs.	Assigns tasks in a way that maximizes students' performance (academic and behavioral).		✓		
5	Understands students with exceptional needs, including those associated with disabilities and giftedness, and knows how to use strategies and resources to address these needs.	User chooses appropriate strategies when students with accommodations are present.	✓			
6	Gives students adequate time to transition between instructional activities.	Selects the appropriate strategy 'give break' before assigning a new task.		✓		
7	Pacing maintains student involvement and engagement.	Changes task or makes a comment when students become distressed.				✓
8	Uses proximity control.	Walks by students who have been disruptive.				✓
9	Uses praise and encourages positive behavior	Makes friendly conversation or the 'give praise' strategy for students who have moved up a behavior band.				✓

Figure 3. SimSchool observation report based on participant actions in the simulated classroom.

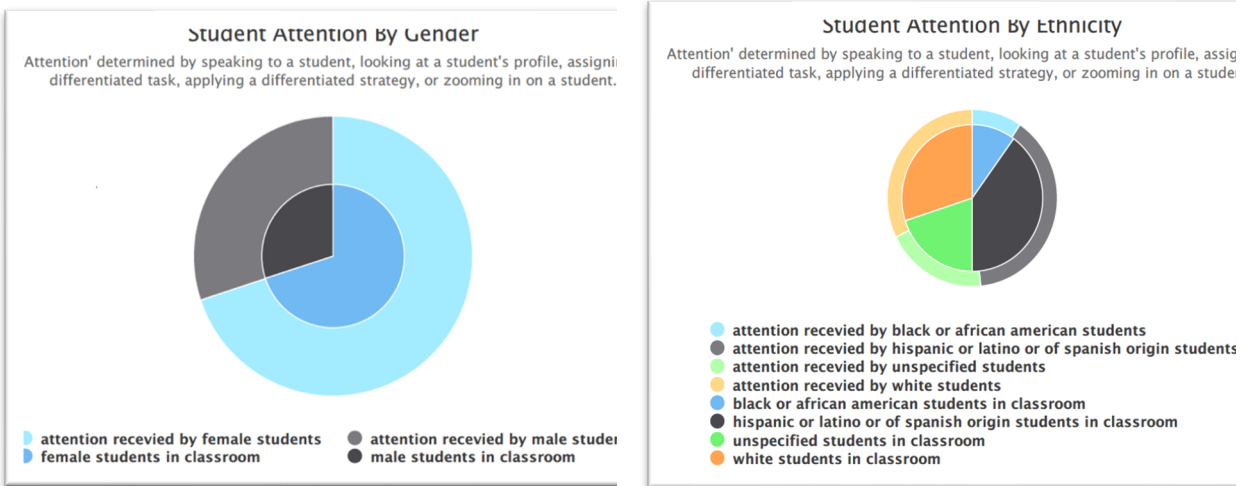


Figure 4. SimSchool equality index feedback by gender and ethnicity.

Participants

Complete data that included pre-post self-report surveys for the teachers and the students as well as completion of the required modules was used for the analysis in this report. The data set included 23 teachers from the Aspire Public Schools and 16 teachers from Byron Nelson High School. The Aspire schools data included elementary (11), middle school (10) and high school (2) teachers and students while Byron Nelson only included high school participants. Twenty-eight (71.8%) of the teachers reported being females and eleven (28.2%) of the teachers reported being male.

Table 2. *Reported Ethnicity of Participating Teachers*

Teacher Ethnicity	Frequency	Percent
Asian	4	10.3
Black or African American	2	5.1
Hispanic	4	10.3
Latinx	1	2.6
White	28	71.8
Total	39	100.0

Table 3. *Frequency of Grade Level Bands*

School Level	Frequency	Percent
Elementary	11	28.2
Middle	10	25.6
High	18	46.2
Total	39	100.0

Data Collection

Data were collected within the simulator to measure changes during the modules for academic, emotional and equality performances by participants. In addition, teacher survey measures focused on self-efficacy, culturally responsive teaching, and self-awareness of bias were embedded within the simSchool platform and collected pre-post participation. Student survey data were collected prior to and following their teachers' participation in the project. Below are descriptions of both teacher and student surveys. Each of the surveys is rated on a 6-point Likert scale from Strongly Disagree (1) to Strongly Agree (6).

The teacher surveys include:

1. The *Teachers' Sense of Efficacy Scale* (TSES) short form (Tschannen-Moran & Hoy, 2001) was used to measure self-efficacy related to three subscales: instructional strategies, classroom management, and student engagement. The survey has been used in multiple studies and shown to have validity and reliability as described in Tschannen-Moran et al. (2001).
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 (*Self-Awareness*, *Pedagogical environment*, and *Relationships with families and community*) will be used to cross validate with the simSchool bias index derived by the simSchool system. This survey was recently used with simSchool research and is based on Chen, Nimmo, & Fraser (2009).
4. Two scales that were used with a simulated environment were also included and are designed to measure individual pair of educator mindsets (Equality vs. Equity; and Avoidant vs. Aware). These surveys 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.

5. Locus of control 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.

Student surveys focused on dispositions related to learning (student engagement, cultural identity, student voice and their perceived culturally relevant teaching of their environment). Selected scales from three surveys included 44 items administered pre and post for year two. The administered survey scales include:

1. 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 of student engagement (9 items), Teacher-Student Relationships, was administered.
2. Two scales from the *Student Measure of Culturally Responsive Teaching* (Dickson, Chun & Fernandez, 2016) were administered to the students to measure their perceptions of their teachers' level of culturally responsive teaching. The first scale addresses the construct of Cultural Engagement while the second scale assesses Diverse Teaching Practices from the student point of view.
3. Two scales representing independent constructs from the Student Participation Survey (SPS) (Anderson, Graham & Thomas, 2019) were administered to students. Voice about Schooling and Voice about Having influence was gathered to assess how students feel about the level of choices they have in their classrooms.

Results

For the second year of the project, there were 39 educators who completed each component of the project including a pretest survey, completion of an introductory module as well as three additional district-selected modules in simSchool, pre-post surveys and pre-post surveys from their students. Survey data from students of the educators were also collected pre and post. However, there were no identifiers attached to the students and therefore could not be matched pre and post. The student results will be based on the 540 pre and 435 post surveys. Results will be reported for educators first and then students.

Educator Results

Data were collected both during the simulations and through pre-post surveys related to teacher cultural self-efficacy and educator bias. Of the 39 teachers who completed all modules and data collection, there were 23 from Aspire Public Schools and 16 from a high school in Northwest ISD.

As shown in Table 4, there were significant ($p < .05$) positive changes from pre to post for Culturally Responsive Teaching Self-efficacy and two scales on the Educator Bias Inventory:

Pedagogical environment and Relationship with families and communities. Effect size is a meaningful indicator for looking at educational significance. As shown in Table 4, the effect sizes were educationally meaningful at .30 and above for three of the seven measures. Fourteen of the 25 individual items on the Culturally Responsive Self-Efficacy Survey and five of the Educator Bias Inventory were significant and are shown in Table 5. Effect sizes for each of these are also included in Table 5. An additional single item related to reflection of participants' own identity was also included in the survey items. As shown in Table 6, while not significant at the $p < .05$ level, the effect size was .21 (Cohen's d) regarding the magnitude of the gain for this item, meaning the teachers became more reflective on their identity following the simEquity module completions.

Table 4. Comparison of Paired Pre and Post T-tests for Equity-Related Subscales for Teachers

Scale	PrePost	Mean	Std. Dev	Sig	Cohen's d
Teacher Efficacy (TE) for Instructional Strategies	Pre	4.98	.51	.640	.09
	Post	5.06	1.1		
Teacher Efficacy (TE) for Classroom Management	Pre	5.08	.58	.681	-.07
	Post	5.00	1.13		
Teacher Efficacy (TE) for Student Engagement	Pre	4.93	.56	.702	.08
	Post	5.00	1.09		
Culturally Responsive Teaching Self-Efficacy Survey	Pre	4.94	.58	.001	.44
	Post	5.19	.55		
Educator Bias Inventory: Self Awareness	Pre	5.32	.46	.394	.14
	Post	5.38	.41		
Educator Bias Inventory: Pedagogical environment	Pre	5.11	.55	.002	.32
	Post	5.28	.50		
Educator Bias Inventory: Relationship with families and community	Pre	4.53	.96	.006	.33
	Post	4.82	.82		

Note: $n = 39$; * Significant at the $p = .05$ level. Cohen's (1988) effect size guidelines .2 = small, .5 = moderate, .8 = large.

Table 5. Individual Educator Items that were Significant from the Subscales

		Mean	N	Std. Dev	p	Cohen's d
Culturally Responsive Self-Efficacy Items						
<i>I feel confident I could...</i>						
obtain information about my students' academic strengths.	CR2pre	5.08	39	.81	.014	.411
	CR2post	5.36	39	.63		
identify ways that the school culture (e.g., values, norms, and practices) is different from my students' home culture.	CR4pre	4.77	39	.87	.002	.531
	CR4post	5.21	39	.80		
implement strategies to minimize the effects of the mismatch between my students' home culture and the school culture.	CR5pre	4.46	39	1.05	.004	.489
	CR5post	4.95	39	.94		
	CR6pre	5.03	39	.67		

assess student learning using various types of assessments.	CR6post	5.36	39	.58	.003	.503
obtain information about my students' home life.	CR7pre	4.69	39	.95		
	CR7post	5.13	39	.86	.006	.464
develop a community of learners when my class consists of students from diverse backgrounds.	CR10pre	5.13	39	.77		
	CR10post	5.38	39	.63	.016	.402
obtain information about my students' cultural background.	CR13pre	4.82	39	.72		
	CR13post	5.13	39	.70	.038	.345
help students to develop positive relationships with their classmates.	CR16pre	5.00	39	.83		
	CR16post	5.23	39	.71	.048	.327
revise instructional material to include a representation of different cultural groups.	CR17pre	4.64	39	.99		
	CR17post	4.92	39	.96	.032	.356
help students feel like important members of the classroom.	CR19pre	5.26	39	.72		
	CR19post	5.46	39	.56	.058	.313
identify ways that standardized tests may be biased towards culturally diverse students.	CR20pre	4.54	39	1.00		
	CR20post	4.95	39	.97	.006	.466
use examples that are familiar to students from diverse cultural backgrounds.	CR21pre	4.67	39	1.03		
	CR21post	5.05	39	.92	.014	.411
obtain information regarding my students' academic interests.	CR23pre	5.13	39	.66		
	CR23post	5.41	39	.50	.026	.372
use the interests of my students to make learning meaningful for them.	CR24pre	5.05	39	.69		
	CR24post	5.31	39	.57	.031	.359
<i>Educator Bias Inventory Items</i>						
I actively encourage critical thinking about differences, stereotypes and biases.	EBPE20pre	5.08	39	.74		
	EBPE20post	5.33	39	.66	.048	.327
I teach about minority and non-minority groups who have devoted their lives to ending injustice.	EBPE21pre	5.00	39	.73		
	EBPE21post	5.23	39	.67	.018	.396
I have high expectations for learning for all students.	EBPE23pre	4.87	39	.83		
	EBPE23post	5.18	39	.85	.017	.402
I provide the option of translations for families who do not speak English.	EBRF25pre	4.44	39	1.55		
	EBRF25 post	4.79	39	1.08	.046	.330
I include families in creating the learning environment for children.	EBRF27 pre	4.36	39	1.35		
	EBRF27 post	4.69	39	1.20	.036	.349

Table 6. *Pre and Post Paired Means for Reflection on Own Identity*

Reflection Item	PrePost	Mean	N	Std. Dev	Sig	ES
I reflect on how my own identity influences my interactions with students.	Pre	4.85	39	1.01		
	Post	5.05	39	.86	.210	.21

Changes in Disposition by Teacher Gender

As shown in Table 7, from pre to post, male teachers increased significantly ($p < .05$) on two of the seven measures. However, female teachers increased significantly on six of the seven measures. The intervention appears to have had a greater impact on female teachers than male teachers. It is important to note that females started lower on five of the seven scales.

Table 7. *Differences in Teacher Dispositions by Gender*

Pre-Post Scales Paired Scales		Male Teachers				Female Teachers			
		Mean	N	SD	Sig	Mean	N	SD	Sig
Teacher Efficacy (TE) for Instructional Strategies	Pre	5.05	11	.52		4.91	25	.50	
	Post	5.45	11	.50	.009	5.17	25	.54	.006
Teacher Efficacy (TE) for Classroom Management	Pre	5.25	11	.47		4.91	25	.57	
	Post	5.32	11	.65	.506	5.14	25	.64	.043
Teacher Efficacy (TE) for Student Engagement	Pre	4.98	11	.64		4.86	25	.49	
	Post	5.20	11	.66	.082	5.19	25	.51	.018
Culturally Responsive Teaching Self-Efficacy Survey	Pre	5.01	11	.74		4.86	25	.47	
	Post	5.17	11	.80	.109	5.14	25	.40	.008
Educator Bias Inventory: Self Awareness	Pre	5.30	11	.49		5.29	25	.43	
	Post	5.35	11	.47	.752	5.35	25	.37	.469
Educator Bias Inventory: Pedagogical environment	Pre	5.01	11	.68		5.11	25	.45	
	Post	5.26	11	.66	.027	5.25	25	.39	.054
Educator Bias Inventory: Relationship with families and community	Pre	4.44	11	1.15		4.47	25	.85	
	Post	4.74	11	1.10	.114	4.79	25	.64	.028

Results from the Simulation Data

Teaching behaviors are captured within the simSchool system that allow computation of academic gains, emotional gains, and equality gains while teaching within a module. In addition, ratings for “How likely to succeed in future lessons” were also recorded in the system by each teacher for each of the 12 simStudents in a teacher’s class. These were presented to the teachers following the first 15-minute simulation, and the last (fifth) 15-minute simulation, for each of the three modules. Separate ratings were completed by each teacher, first based on just reflecting on the image of each simStudent (Avatar featuring different skin tones) and then based on name of the student with no image.

Trends for the major findings within and across modules are presented in Table 7 and graphically highlighted in Figures 5-7. Across the three modules completed by each teacher, from the first time of capturing data at the end of completing the first 15-minute simulation (of 5) for Module 1, to the end of the last 15-minute simulation (of 5) for Module 3, the overall gains shown in Table 8 and graphically displayed in Figures 5, 6, and 7 were highly significant ($p < .005$) for Academic Index (ES = .42), Avatar Rating (ES = .37), and Name Rating (ES = .29). The magnitudes of these gains are at or beyond levels that would be widely accepted as

educationally meaningful in the research literature (Bialo & Sivin-Kachala, 1996). The simulator-produced Emotional Index and Equality Index showed non-significant (at the $p < .05$ level) positive gains overall with effect sizes of $ES = .08$ and $ES = .07$ respectively. These would be considered small positive effects by Cohen (1988) and would be considered “Developmental Effects” (0 - .1) by Hattie (2009).

Within each module there were differences that are beginning to emerge and will be studied in more detail. Specifically, for Module 1, First to Last Academic, Emotional, Equality and Avatar plus Name ratings all become more positive from the first of five simulations to the last of five simulations, and all but the Equality Index gains were significant ($p < .05$). Within Module 2, Teachers appear to have concentrated on Academic Gains ($p < .0005$, $ES = .41$). They began on the first simulation in M2 near the same level of proficiency as where they began (with a different lesson to teach) in M1, but advanced more. None of the other indices changed significantly ($p < .05$) and magnitude/ES was typically very small in M2. In Module 3, with a third new lesson to teach, educators held steady at what we can now call a “high” level of First and Last in Academic Index (.85 vs. .72 in the beginning), Avatar Ratings and Name ratings (3.7+ vs. ≤ 3.4 in the beginning). Mean ratings in these three areas changed little first to last simulation. Teachers advanced first to last significantly ($p < .05$) in the areas of Emotional Index and Equality Index but the magnitude/ES was $ES = .12$ and $ES = .09$, small per Cohen (1988). Note that these latter two areas were the only ones without significant gains in Module 1.

Table 8. *Simulation Data Gains for Academic, Emotional and Equality Gains by Module*

Module Number	First Measure	Last Measure
	<i>First Academic Index</i>	<i>Last Academic Index</i>
M1	.7170	.8108
M2	.7262	.8505
M3	.8551	.8549
	<i>First Emotional Index</i>	<i>Last Emotional Index</i>
M1	.1897	.2465
M2	.2954	.2615
M3	.1813	.2195
	<i>First Equality Index</i>	<i>Last Equality Index</i>
M1	.9880	.9885
M2	.9892	.9879
M3	.9869	.9887
	<i>First Avatar Rating</i>	<i>Last Avatar Rating</i>
M1	3.32	3.60
M2	3.60	3.56
M3	3.78	3.77
	<i>First Name Rating</i>	<i>Last Name Rating</i>
M1	3.40	3.63
M2	3.61	3.66
M3	3.70	3.74

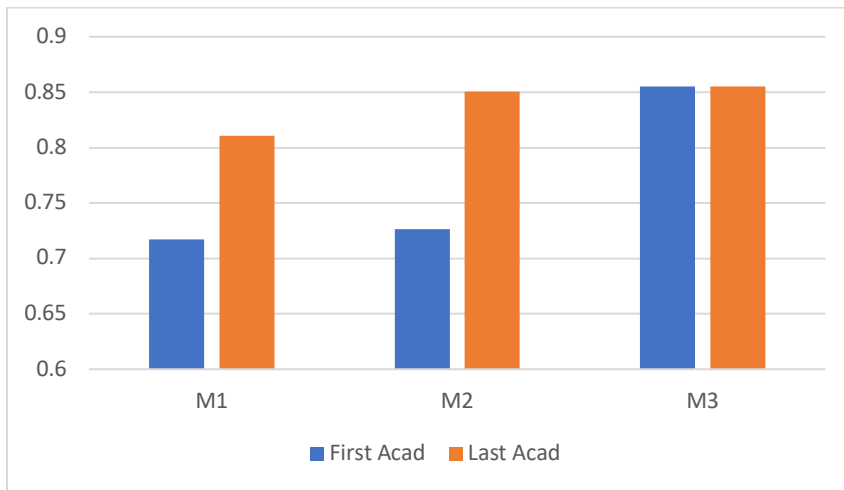


Figure 5. SimSchool Academic Index group mean values for M1, M2 and M3.

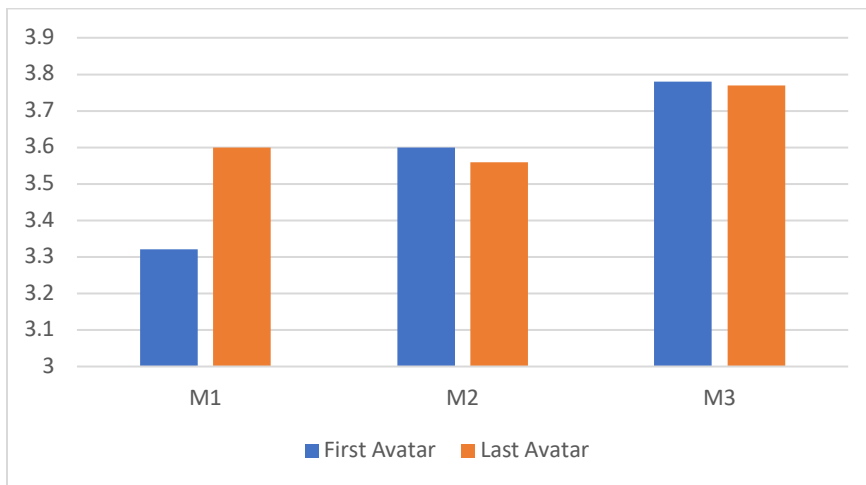


Figure 6. SimSchool Avatar Rating group mean values for M1, M2 and M3.

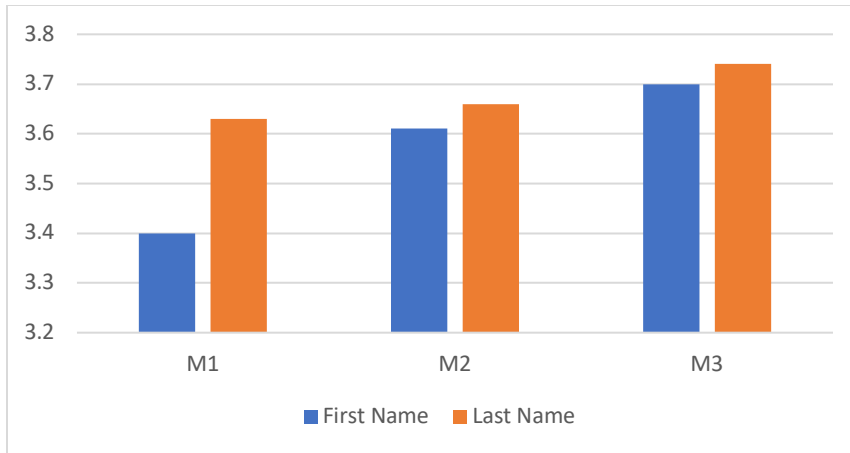


Figure 7. SimSchool SimSchool Name Rating group mean values for M1, M2 and M3.

Student Results

While only post data were collected for students in year one, both pre and post test survey data were collected from students of those teachers who participated in year two. Because the data were anonymous, there was no way to pair the pre and post test data. The reported results are from pre and post test comparisons, but not paired. For pretest data, 822 students completed the surveys prior to their teachers' participation in simSchool activities while 574 students completed surveys following their teachers' participation. Complete data varied by scale and is reported with the usable, complete data for each measure. The breakdown by school system is shown in Table 9. The demographics of students by gender, grade level and ethnicity by pre and post are shown in Tables 10 - 12 below.

Table 9. *Student Survey Respondents by Site*

School System	Pre	Post	Total
Aspire	540	435	975
BNHS	282	139	421
Total	822	574	1396

Table 10. *Self-Reported Gender of Students*

Gender	Pre	Post	Total
Male	389	256	645
Female	401	296	697
Total	790	552	1342

Table 11. *Self-Reported Grade Level of Students*

Grade Level	Pre	Post	Total
-------------	-----	------	-------

3	51	95	146
4	59	51	110
5	139	63	202
6	79	60	139
7	72	74	146
8	138	82	220
9	97	60	157
10	47	11	58
11	68	30	98
12	72	48	120
Total	822	574	1396

Table 12. *Self-Reported Ethnicity of Students*

Ethnicity	Pre	Post	Total
American Indian	10	9	19
Asian	105	83	188
Black/AA	95	72	167
Hispanic	250	150	400
No response	105	98	203
Latinx	43	20	63
Native Hawaiian/PI	9	4	13
White	205	138	343
Total	822	574	1396

Results from self-reported survey data included five indicators: students’ perceptions of their voice in schooling, voice in having influence about schooling, student engagement, how they feel their teachers practice culturally responsive teaching practices (which is measured in two separate subscales of cultural engagement and the perception of diverse teaching practices by their teachers). As shown in Table 13, analysis of variance was used to determine the pre-post differences on each of the subscales. Three of the five subscales showed a significant ($p < .05$) increase: voice having influence, diverse teaching practices of their teachers and the engagement they felt as students.

Table 13. *Overall Means and Standard Deviations for Each of the Student Scales*

Measure		N	Mean	Std. Dev.	Sig	ES
Voice in Schooling	Pre	822	3.75	1.04	.100	
	Post	574	3.84	1.02		
	Total	1396	3.79	1.03		
Voice Having Influence	Pre	803	3.94	1.08		
	Post	556	4.15	1.07		

Cultural Engagement (CRTP subscale)	Total	1359	4.03	1.08	.000	.20
	Pre	745	3.60	1.05		
	Post	530	3.68	1.09		
Diverse Teaching Practices (CRTP subscale)	Total	1275	3.64	1.07	.194	
	Pre	745	4.55	.83		
	Post	530	4.67	.85		
Student Engagement	Total	1275	4.60	.84	.021	.14
	Pre	709	4.31	1.05		
	Post	524	4.43	.99		
	Total	1233	4.36	1.03	.041	.12

Analysis of variance by gender at pretest and post test times showed that there were initially significant differences between males and females on two of the subscales. Specifically, at pretest time males were significantly ($p = .048$) higher than females on their voice having influence and significantly ($p = .007$) higher on student engagement. By post test time, there were no significant ($p < .05$) differences between males and females on any of the measures. When comparing pre-post for males and females separately, there were no significant pre-post changes for males (Table 14) but for females, three of the measures increased significantly ($p < .05$) pre-post (Table 15). Females gained in voice having influence, perception of their teachers' diverse teaching practices and student engagement.

Table 14. *Comparison of Males by Each of the Student Scales*

Measure		N	Mean	Std. Deviation	Sig
Voice in Schooling	Pre	389	3.82	1.03	
	Post	256	3.89	1.04	
	Total	645	3.85	1.03	.377
Voice Having Influence	Pre	376	4.02	1.09	
	Post	248	4.13	1.16	
	Total	624	4.07	1.12	.225
Cultural Engagement (CRTP subscale)	Pre	352	3.63	1.08	
	Post	238	3.66	1.16	
	Total	590	3.65	1.11	.725
Diverse Teaching Practices (CRTP subscale)	Pre	352	4.58	.84	
	Post	238	4.61	.92	
	Total	590	4.60	.88	.714
Student Engagement	Pre	333	4.42	1.03	
	Post	235	4.45	1.04	
	Total	568	4.43	1.04	.749

Table 15. *Comparison of Females by Each of the Student Scales*

Measure		N	Mean	Std. Dev.	Sig.
Voice in Schooling	Pre	401	3.68	1.03	.180
	Post	296	3.78	1.00	
	Total	697	3.72	1.02	
Voice Having Influence	Pre	395	3.87	1.05	.000
	Post	287	4.16	.98	
	Total	682	3.99	1.03	
Cultural Engagement (CRTP subscale)	Pre	363	3.56	1.01	.219
	Post	270	3.66	1.03	
	Total	633	3.60	1.02	
Diverse Teaching Practices (CRTP subscale)	Pre	363	4.52	.83	.002
	Post	270	4.72	.77	
	Total	633	4.60	.81	
Student Engagement	Pre	347	4.21	1.02	.010
	Post	268	4.42	.95	
	Total	615	4.30	.99	

Findings by Student Grade Level

Findings by grade level included all students combined and revealed significant findings pre to post for measures for the grade levels. The significant differences are shown in Table 16 by grade level and measure. As shown in Table 16, voice about schooling significantly ($p < .05$) increased from pre to post for students in grades 3,4,5 and 8.

Table 16. *Significant Findings by Grade Level*

Grade Level	Measure	Pre/Post	N	Mean	Std. Dev	Sig.
Grade 3	Voice in Schooling	Pre	51	4.35	.86	.017
		Post	95	3.95	.99	
		Total	146	4.09	.96	
Grade 4	Voice in Schooling	Pre	59	3.83	.91	.026
		Post	51	4.18	.73	
		Total	110	3.99	.85	
Grade 5	Voice in Schooling	Pre	138	3.98	.88	.002
		Post	63	3.55	.98	
		Total	201	3.85	.93	
	Cultural Engagement	Pre	127	3.87	.97	.003
		Post	59	3.40	1.02	
		Total	186	3.72	1.01	
Grade 8	Voice in Schooling	Pre	138	3.84	.98	
		Post	82	4.12	1.01	

		Total	220	3.94	1.00	.045	
	Voice Having Influence	Pre	134	3.92	1.04		
		Post	79	4.40	1.11		
		Total	213	4.10	1.09	.002	
	Diverse Teaching Practices	Pre	129	4.39	.88		
		Post	74	4.75	.89		
		Total	203	4.52	.90	.007	
Grade 10	Voice Having Influence	Pre	47	3.80	1.05		
		Post	10	4.55	1.17		
		Total	57	3.93	1.10	.050	
	Diverse Teaching Practices	Pre	44	4.36	.78		
		Post	10	5.04	.96		
		Total	54	4.49	.85	.022	
Grade 12	Student Engagement	Pre	62	4.28	1.08		
		Post	46	4.78	.75		
		Total	108	4.49	.98	.008	

Results Related to Reported Student Ethnicity

Analyses were completed to examine pre to post changes for the ethnic groups to which each of the students reported their affiliation. Asian and White students each had significant ($p < .05$) pre-post gains for three measures while American Indian and Hispanic students had significant ($p < .05$) gains for one of the measures. Details of the analyses are shown in Table 17. There were no significant changes from pre to post on any of the measures for Black/African American or Latinx students.

Table 17. *Student Measures Significant by Ethnicity*

Student Ethnicity	Measure	Pre/post	N	Mean	SD	Sig.	ES
American Indian/Alaska Native	Voice About Schooling	Pre	10	3.58	.51		
		Post	9	4.29	.52	.007	1.38
Asian	Voice Having Influence	Pre	103	4.00	1.02		
		Post	80	4.39	.92	.008	.40
	Diverse Teacher Practices	Pre	100	4.60	.79		
		Post	76	4.83	.77	.053	.29
	Student Engagement	Pre	97	4.26	.85		
		Post	75	4.60	.80	.009	.41
Hispanic	Voice Having Influence	Pre	243	3.92	.95		
		Post	144	4.18	.92	.009	.28
White	Voice Having Influence	Pre	204	3.77	1.23		
		Post	137	4.18	1.20	.003	.34
	Diverse Teacher Practices	Pre	191	4.46	.87		
		Post	133	4.72	.88	.009	.30

Student Engagement	Pre	182	4.15	1.22		
	Post	131	4.58	1.02	.001	.38

Differences by Location for Students

Included in this year’s project participants were teachers and students from two different school systems, one in California and one in Texas. As previously stated, the schools in California included elementary, middle and high school participants while the Texas school only included high school participants. While there were no significant pre-post changes for the Aspire schools students (Table 18), all five of the measures indicated pre-post gains that were significant ($p < .05$) for the Byron Nelson students (Table 19). The magnitudes of the gains in each area were at or near the $ES = .3$ level widely accepted as the criterion for considering gains to be educationally meaningful (Bialo & Sivin-Kachala, 1996).

Table 18. *Pre-Post Comparisons for Aspire Public School Students*

Measure		N	Mean	Std. Dev	Sig	ES
Voice in Schooling	Pre	540	3.92	.94		
	Post	435	3.85	.98		
	Total	975	3.89	.96	.236	-.07
Voice Having Influence	Pre	528	4.10	.99		
	Post	421	4.21	1.04		
	Total	949	4.15	1.01	.083	.11
Cultural Engagement (CRTP subscale)	Pre	484	3.74	1.04		
	Post	401	3.69	1.07		
	Total	885	3.72	1.05	.549	-.05
Diverse Teaching Practices (CRTP subscale)	Pre	484	4.69	.79		
	Post	401	4.70	.84		
	Total	885	4.69	.82	.752	.01
Student Engagement	Pre	463	4.46	1.00		
	Post	397	4.46	1.00		
	Total	860	4.46	1.00	.919	.00

Table 19. *Pre-Post Comparisons for Byron Nelson High School Students*

Measure		N	Mean	Std. Dev	Sig	ES
Voice in Schooling	Pre	282	3.42	1.13		
	Post	139	3.82	1.13		
	Total	421	3.55	1.14	.001	.35
Voice Having Influence	Pre	275	3.64	1.18		
	Post	135	3.97	1.12		
	Total	410	3.75	1.17	.007	.28
	Pre	261	3.36	1.04		

Cultural Engagement (CRTP subscale)	Post	129	3.65	1.16		
	Total	390	3.45	1.09	.013	.27
Diverse Teaching Practices (CRTP subscale)	Pre	261	4.31	.86		
	Post	129	4.55	.86		
	Total	390	4.39	.86	.011	.28
Student Engagement	Pre	246	4.01	1.08		
	Post	127	4.34	.98		
	Total	373	4.12	1.06	.005	.32

Combined Teacher/Student Results

Teacher demographics were coded into the student data to find whether there were any differences based on teacher ethnicity or gender. Analysis by teacher ethnicity showed there to be one significant difference by ethnicity for the cultural engagement factor as reported by students. At pretest time, students of White teachers were significantly lower ($p < .005$) on the cultural engagement factor than students of Black teachers. However, by posttest time, there were no significant differences on any of the measures.

Comparing students by their teachers' gender, there were no significant differences on any of the five measures at pretest time. However, by post test time, four of the five measures showed students of female teachers were significantly higher on four of the five measures (Table 20). Separating the analysis by gender of the teacher and analyzing the data from pre to post, the students of female teachers went up significantly from pre to post on each of the five measures (Table 21). However, while not significant, the students of male teachers tended to decrease on all five measures from pretest to posttest (Table 22).

Table 20. *Pre-post Student Measures by Gender of Their Teacher*

Student Measures	Teacher Gender	Student Pretest				Student Posttest			
		N	Mean	Std. Dev	Sig	N	Mean	Std. Dev	Sig
Voice About Schooling	M	214	3.77	1.04		141	3.70	1.03	
	F	608	3.75	1.04		433	3.89	1.01	
	Total	822	3.75	1.04	.793	574	3.84	1.02	.049
Voice Having Influence	M	207	3.97	1.07		136	3.93	1.07	
	F	596	3.93	1.08		420	4.23	1.05	
	Total	803	3.94	1.08	.644	556	4.15	1.06	.005
Cultural Engagement	M	198	3.55	1.06		128	3.39	1.09	
	F	547	3.62	1.05		402	3.77	1.08	
	Total	745	3.60	1.05	.440	530	3.68	1.09	.001
Diverse Teaching Practices	M	198	4.59	.78		128	4.51	.80	
	F	547	4.54	.85		402	4.72	.86	
	Total	745	4.55	.83	.526	530	4.67	.85	.015
Student Engagement	M	188	4.30	1.03		128	4.32	.93	
	F	521	4.31	1.06		396	4.46	1.02	
	Total	709	4.31	1.05	.894	524	4.43	1.00	.155

Table 21. *Pre-post Measures of Students of Female Teachers*

Student Measures	PrePost	N	Mean	Std. Dev.	Sig
Voice About Schooling	Pre	608	3.75	1.04	.024
	Post	433	3.89	1.01	
	Total	1041	3.81	1.03	
Voice Having Influence	Pre	596	3.93	1.08	.000
	Post	420	4.23	1.05	
	Total	1016	4.05	1.08	
Cultural Engagement	Pre	547	3.62	1.05	.029
	Post	402	3.77	1.08	
	Total	949	3.69	1.06	
Diverse Teaching Practices	Pre	547	4.54	.85	.002
	Post	402	4.72	.86	
	Total	949	4.62	.86	
Student Engagement	Pre	521	4.31	1.06	.028
	Post	396	4.46	1.02	
	Total	917	4.38	1.04	

Table 22. *Pre-post Measures of Students of Male Teachers*

Student Measures	PrePost	N	Mean	Std. Dev.	Sig
Voice About Schooling	Pre	214	3.77	1.04	.533
	Post	141	3.70	1.03	
	Total	355	3.74	1.03	
Voice Having Influence	Pre	207	3.97	1.07	.730
	Post	136	3.93	1.07	
	Total	343	3.95	1.07	
Cultural Engagement	Pre	198	3.55	1.06	.191
	Post	128	3.39	1.09	
	Total	326	3.49	1.08	
Diverse Teaching Practices	Pre	198	4.59	.78	.370
	Post	128	4.51	.80	
	Total	326	4.56	.79	
Student Engagement	Pre	188	4.30	1.03	.858
	Post	128	4.32	.93	
	Total	316	4.31	.99	

Summary and Discussion of Findings

Many interesting findings promote continuing and expanding the approach to using simulations to help educators with equity-based teaching practices. There were many significant as well as meaningful findings for both teachers and their students. From self-reported pre-post data, there were significant ($p < .05$) positive changes from pre to post for Culturally Responsive Teaching Self-efficacy and two scales on the Educator Bias Inventory: Pedagogical environment and Relationship with families and communities. The effect sizes were educationally meaningful at

.30 and above for 3 of the 7 measures. (Bialo & Sivin-Kachala, 1996). There were many individual items that also increased significantly ($p < .05$) from pre to post test.

Significant findings emerged from the data gathered and/or produced by the simulator. First, across the three modules completed by each teacher, from the first time of capturing data at the end of completing the first 15-minute simulation (of 5) for Module 1, to the end of the last 15-minute simulation (of 5) for Module 3, overall gains for Academic Index, Avatar Rating, and Name Rating were highly significant ($p < .005$) with effect sizes (Cohen's d) for Academic Index = .42, Avatar Rating = .37, and Name Rating = .29). The magnitudes of these gains are at or beyond levels that would be widely accepted as educationally meaningful in the research literature (Bialo & Sivin-Kachala, 1996). Second, analyses focused on simStudents labeled by the simulator as English Language Learners (ELL) confirmed that after the first 15-minute simulation with a simClass, teachers rated ELL-labeled students as "less likely to succeed in future lessons" even though the learning characteristics of the ELL-labeled students did not actually differ from the profiles of the non-ELL simStudents. This could be considered a form of bias by the teachers. However, in a third major finding based on simulator data, teachers produced a higher ratio of targeted actions to tips ($p < .05$) for the same students labeled ELL versus their non-ELL counterparts, when average ratios for the two groups were compared. In a related finding, the ratios of Targeted Actions to Tips increased significantly ($p < .05$) across all simStudents as an aggregated class, from first 15-minute simulation to last 15-minute simulation for Module 2 and Module 3. This ratio also rose slightly from first 15-minute simulation to last 15-minute simulation for Module 1, but the difference / gain was not significant for Module 1.

Results from self-reported pre-post student survey data included six indicators including students' perceptions of their voice in schooling, voice in having influence about schooling, student engagement, how they feel their teachers practice culturally responsive teaching practices (which also is measured in two separate subscales of cultural engagement and the perception of diverse teaching practices by their teachers). Three of the five subscales showed a significant ($p < .05$) increase for their voice having influence, the diverse teaching practices of their teachers and the engagement they felt as students.

In addition, pretest and post test showed that there were significant differences between males and females on two of the subscales. Males were significantly ($p = .048$) higher than females on their voice having influence and significantly ($p = .007$) higher on student engagement. By post test time, there were no significant ($p < .05$) differences between males and females on any of the measures. When comparing pre-post for males and females separately, there were no significant pre-post changes for males but for females, three of the measures increased significantly ($p < .05$) pre-post. Females gained in voice having influence, perception of their teachers' diverse teaching practices and student engagement.

The student participants were from the grade range of 3 through 12. Analysis was completed for each grade level. The pre-post increases discussed in this report showed that most of the significant ($p < .05$) findings were related to voice – either voice in schooling or their voice having influence.

Analyses were completed to examine pre to post changes for the ethnic groups to which each of the students reported themselves. Asian and White students each had significant ($p < .05$) pre-post gains for three measures while American Indian and Hispanic students had significant ($p < .05$) gains for one of the measures. There were no significant changes from pre to post on any of the measures for Black/African American or Latinx students.

For year two, there were participants from different school systems in two different states. There were differences between the groups, but the reasons are unclear. One system included elementary, middle and high school participants while the other system only included high school participants. While there were no significant pre-post changes for the Aspire/California schools students, all five of the measures indicated pre-post gains that were significant ($p < .05$) for the Byron Nelson (Texas) students. The magnitudes of the gains in each area were at or near the $ES = .3$ level widely accepted as the criterion for considering gains to be educationally meaningful (Bialo & Sivin-Kachala, 1996).

Analysis of student data by teacher ethnicity showed there to be one significant difference by ethnicity for the cultural engagement subscale. At pretest time, students of White teachers were significantly lower ($p < .005$) on the cultural engagement scale than students of Black teachers. However, by posttest time, there were no significant differences on any of the measures related to ethnicity of their teachers.

Analyzing student data by their teachers' gender revealed no significant differences on any of the five measures at pretest time. However, by post test time, four of the five measures showed students of female teachers were significantly higher on four of the five measures. Separating the analysis by gender of the teacher and analyzing the data from pre to post, the students of female teachers went up significantly from pre to post on each of the five measures (Table 20). However, while not significant, the students of male teachers tended to decrease on all five measures from pretest to posttest.

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.

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

Thanks to Gerald Knezek for the analysis on the simulated data findings and to Jenna Kelley for grant support.

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.

- 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
- 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. & Johnston, V. (2019). SimSchool: SPARCing New Grounds in Research on Simulated Classrooms. In K. Graziano (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 733-739). Las Vegas, NV, United States: Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/primary/p/207723/>.
- 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>
- 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.
- Hattie, J. (2009). *Visible Learning*. London: Routledge.
- Knezek, G., Hopper, S., Christensen, R., Tyler-Wood, T., & Gibson, D. (2015). Assessing pedagogical balance in a simulated classroom environment. *Journal Digital Learning in Teacher Education*, 31(4), 148-159.
- 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
- Malone, J.C. (2016). *E.R. Guthrie: A behaviorism for everyone*. In D. Zilio & K. Carrara (eds.). *Behaviorisms: Historical and conceptual issues* (Vol. 1). Sao Paulo: Nucleo Paradigma Press.
- 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.
- Tschannen-Moran, M., & Hoy, A.W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Tyler-Wood, T., Estes, M., Christensen, R., Knezek, G., & Gibson, D. (2017). SimSchool: An opportunity for using serious gaming for training teachers in rural areas. *Rural Special Education Quarterly*, 34(3), 17-20. <https://doi.org/10.1177/875687051503400304>