Using Simulation Experiences to Address Bias in Teaching Practices

Rhonda Christensen and Gerald Knezek
University of North Texas
Denton, TX USA

Abstract. To address the diversity of student differences, educators need to actively recognize and counter patterns of bias in their teaching practices as well as in their classroom environments. This paper presents a tool for educators that is scalable for developing equitable, culturally responsive teaching practices through implementation in a simulated teaching environment.

Introduction

Classrooms are more diverse as they continue to reflect the society in which we live. However, there is much less diversity in the gender and ethnicity of classroom teachers. Gender, ethnicity, socioeconomic status, and English language learning status have been linked to differences in teacher perceptions of students for whom educators may hold implicit negative attitudes and stereotypes. To address the diversity of differences, educators need to actively recognize and counter patterns of bias in their teaching practices as well as classroom environments. This paper presents a tool for educators that is scalable for developing equitable, culturally responsive teaching practices through implementation in a simulated teaching environment that is adaptive and interactive.

Improving teaching strategies through a simulated teaching environment has been shown to improve teacher self-efficacy, teaching skills, classroom management and multicultural awareness. The current study is using a simulation program to help educators recognize possible bias with the goal of reflecting on and remediating any biases that may exist. Both self-report survey data and simulation-captured data are used to understand the changes that occur as educators complete many trials of simulated teaching, using feedback for improvement each time.

First year project findings have shown that there were significant ($p < .05$) positive changes pre to post resulting from simulation experiences, based on self-report data for self-efficacy for instructional strategies as well as culturally responsive teaching self-efficacy. There were also large gains pre to post based on machine captured data during the simulation for equitable teaching strategies focused on the simStudents. Simulations allow users to practice multiple iterations with guided feedback to improve their teaching strategies for specific student needs. In addition, the simulation provides the opportunity to experience students with a variety of learning needs. This may not be possible in real life for teacher preparation candidates during, their teacher preparation programs.

ISTE Standards

This research paper involves several of the ISTE standards including: Equity and inclusion, Connected learner and the Learner categories of teacher standards. The focus of the research explores the impact of teacher simulation on increase in equitable teaching practices to support learning based on recognized student needs. The use of digital simulation experiences allows

users to reflect on practices that allow them to connect to all types of learners. Research findings using artificially intelligent simStudents to improve student learning will be described.

Theoretical Framework

The concept of perceived self-efficacy is rooted in social cognitive theory and is "concerned with judgments about how well one can organize and execute courses of action required to deal with prospective situations containing many ambiguous, unpredictable, and often stressful elements" (Bandura & Schunk, 1981, p 587). An individual's perception of ability to impact a situation is critical for whether or not they actually do affect change (Bandura, 2012). Teachers' beliefs about their ability to make a difference for students impacts their resilience and persistence in difficult situations (Gibson & Dembo, 1984; Milner, 2003).

Many research studies have found that a teacher's sense of self-efficacy was one of the variables highly related to student achievement (Medgley, Feldlaufer, & Eccles, 1989; Tucker et al., 2005) while others specifically noted the impact on Black students (Tucker et al., 2002). 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). Teachers who report a strong sense of self-efficacy are more likely to believe student learning outcomes are within their control and within their influence to impact. Connections between teachers' sense of efficacy, culturally responsive pedagogy (Callaway, 2016), and student achievement (Oyerinde, 2008; Tucker et al., 2005) have been shown to exist.

Research on teacher efficacy and its relationship with culturally responsive teaching illustrates a need to address teacher self-efficacy with respect to working with children from diverse backgrounds (APA, 2012; Oyerinde, 2008; Tucker et al., 2005). In order to increase low academic achievement among culturally diverse students, efforts should be made to increase teacher self-efficacy (Callaway, 2016; Tucker et al., 2005). Highly efficacious teachers are more likely to persist in helping struggling students, and are able to design and create more engaging lessons for their students (Bandura, 1997; Kitsantas, 2012; Protheroe, 2008). The project's researchers have demonstrated simSchool's impact on teachers' instructional self-efficacy (Knezek & Christensen, 2009).

About the Project

While refining their own best practices in this "flight simulator for teachers," participants interacted with this cognitive model over several sessions spanning several weeks, with micro-teaching interactions lasting from 15 to 30 minutes. The sessions were conducted within one of

the assigned modules. The modules were selected as professional development for classroom teachers and focused on culturally related topics such as race, ethnicity, bullying and bias (see Table 1). Prior to beginning the sessions, users reviewed the student profiles that contained information on student strengths, preferences and academic performance so they could attempt to match instruction with learner needs. During the sessions, participants attempted to negotiate the simulated classroom environment while adapting their teaching to the diversity of students they encountered.

Table 1. *Modules Completed by Classroom Teachers*

Elementary Teacher Modules
Module 1: Introduction to Teaching in simSchool
Module 2: Cultural Intelligence and Inclusion 2.0
Module 3: ELE 3-5 Bullying and Bias the First Coconut Tree
Module 4: ELE 3-5 Gender and Identity Supermom Saves the Day Why Can't Girls Be Superheros
Middle School Teacher Modules
Module 1: Introduction to Teaching in simSchool
Module 2: MS 6-8 Gender and Identity: The Misfits
Module 3: MS 6-8 History Empowering Learners to Change the World
Module 4: MS 6-8 (Race, Ethnicity, Class, Immigration) A Tale of Two Schools

At the end of each simulation session, participants received multiple types of 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 were documented for review were impacts on individual students of conversational stances, communication patterns, and attentional habits of the teacher. This feedback can reveal aspects of a teacher's implicit biases. Participants must view the feedback prior to completing additional sessions in each module.

Methods

Participants included classroom teachers from a western state in the US. Demographics available for 12 of the 13 respondents indicated data were from primarily middle school teachers (9/12 = 75%) with the remainder from elementary school teachers. The respondents were predominately female (10/12 = 83%). All reported they were born in the USA, although parents of two were born in another country.

For the study reported in this paper, participants rated how they thought students performed after they completed one simulation. First the participants rated how they perceived the simStudent performed by looking at their avatar image and then by simStudent names. The intention was to detect any possible bias due to images or names of students. Completing the surveys was required in order to progress to the next simulation and complete the module. The question was worded: *"Using the scale below, give a prediction for each student's ability to be successful with the lesson plan used."* The scale included five options: "Very Unlikely", "Somewhat Unlikely", "Neutral", "Somewhat Likely", "Very Likely" (Figure 1.) Teacher rating data recorded within the simSchool system used 1 = very unlikely to 5 = very likely as the rating

scale. The same rating scale and question to the teacher was used for a separate presentation based on Name, as shown in Figure 2.

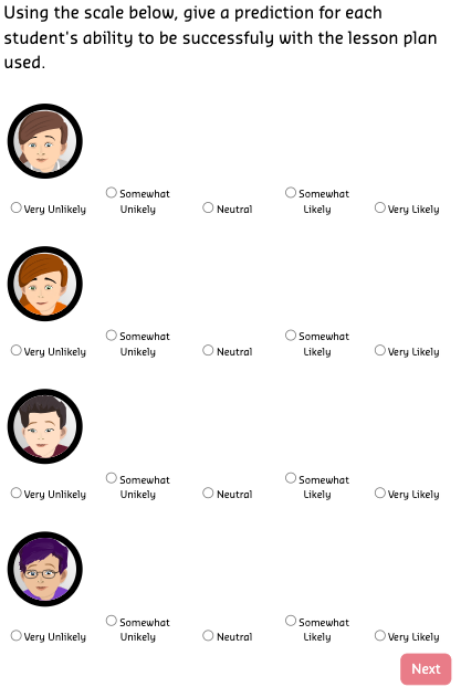

Figure 1. Predicting success using avatar images in simSchool

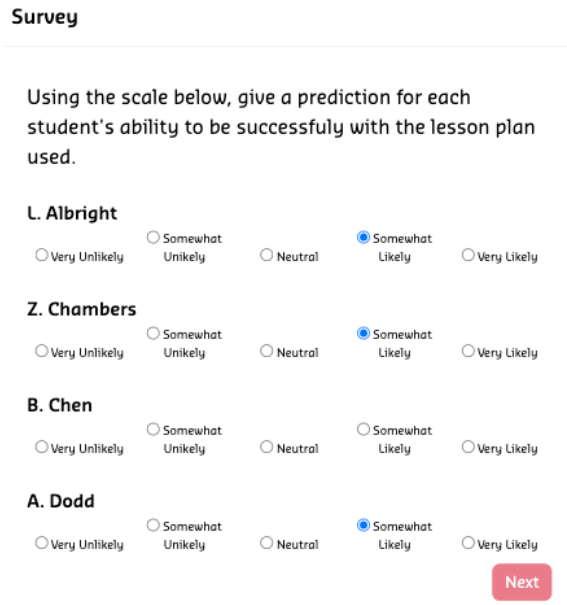

Figure 2. Predicting success using simStudent names.

Results

Preliminary analyses for the year two data available as of the submission of this paper have been completed for teachers who had finished all modules. A high-level analysis was completed by simply counting the number of times teacher ratings after completing one simulation were higher

(greater rating of likelihood of success in lesson; see Fig. 1) versus lower (rated less likely to be successful in lesson), for IEP or ELL labeled students. Teacher ratings where there were no variations across all six simStudents, post simulation, were not included in tally.

Two examples of qualifying units of analysis, based on six simStudents that had identical underlying student performance models but systematically altered IEP or ELL accommodations across varying avatar or name attributes, are shown in Figures 3 and 4. In Figure 3, the mean rating (3.0) by one teacher for Avatar Rating (based on Fig. 1) was the same across all three pairs of "twins" in terms of ELL N/Y labeling, so this instance was not counted in the "higher or lower" analysis for possible rating biases. No positive or negative bias was indicated. However, this teacher produced a mean rating of 2.5 for the ELL-Yes labeled students versus 3.5 and 4.0 respectively, for the two other pairs of ELL-N labeled students in the class, just based on seeing the student names (see rating scale in Fig. 2) for the simStudents who had been in the simulation just completed. For the example in Fig. 3, one tally was counted toward possible negative bias in the domain of Names, for ELL-Yes labeled simStudents. The tally was placed in the category of "lower" for Likert rating based on Name.

Who, Me? A Scientist

Base Profile	Gender	Ethnicity	IEP	ELL	Avatar Rating	Ave. Raing for Pair	Name Rating	Ave. Raing for Pair
Ashley Dodd	F	white	N	N	3		4	
Cameron Fields	M	white	N	N	3	3.0	3	3.5
Victoria Kramer	F	white	N	Y	3		2	
Trenton Knox	M	white	N	Y	3	3.0	3	2.5
Zoey Chambers	F	white	N	N	3		4	
Luke Albright	M	white	N	N	3	3.0	4	4.0

Figure 3. Example unit of analysis where Name is rated lower but Avatar not, across simStudents with same underlying student model (Variation: N/Y ELL).

As shown in Fig. 4, where the rating took place for a different teacher, with a different set of six simStudents, also having identical underlying student profiles, the mean rating for the ELL-Yes pair in Avatar Rating was 3.5, lower than both the mean rating of 5.0 for the pair above and the mean rating of 4.0 for the pair below. With regard to the Name Rating, the mean rating of the ELL-Yes pair was 4.5, lower than the mean rating of 5.0 for the pair immediately below and no higher than the pair immediately above, and therefore fitting the definition of possible bias with respect to at least one of the other pairs. Thus, the ratings by this teacher resulted in one tally of "lower" for Avatar Rating and "lower" for Name Rating. Using this same set of definitions, occasionally the ratings by teachers for the ELL or IEP labeled simStudents would result in a tally in the "higher" category of counts as well.

Clone Profile	Gender	Ethnicity	IEP	ELL	Avatar Rating	Ave. for Pair	Name Rating	Ave. for Pair
Keyona Jackson	F	Black	N	N	5		4	
Ladarrius Washington	M	Black	N	N	5	5.0	5	4.5
Sophia Ruiz	F	Hispanic	N	Y	3		5	
Luis Morales	M	Hispanic	N	Y	4	3.5	4	4.5
Khadeeja Fidali	F	Asian	N	N	5		5	
Brandon Chen	M	Asian	N	N	4	4.0	5	5.0

Figure 4. Example unit of analysis where Avatar and Name are rated lower across simStudents with same underlying student model (Variation: N/Y ELL).

Results of the analysis were that among ratings completed with variations in Avatar, where IEP and ELL labels were interwoven in a counterbalanced manner throughout, 18 of the

23 qualifying independent tests resulted in one or both of each IEP or ELL pair of 2 students among the groups of six being rated lower (less likely to be successful in the lesson), while in 3 cases one or both of the IEP or ELL pair of two were rated higher (more likely to be successful in the lesson), after one simulation run. The binomial probability of 18 or more of 21 tests being lower by chance is $p < .002$ (Graphpad, 2023).

Similarly, among the cases where teachers were presented with variations in simStudent names, 14 of 17 sets of 6-student ratings resulted in one or both of each IEP or ELL pair of 2 students among the groups of six being rated lower (less likely to be successful in the lesson), while only three were rated higher. This event would also be considered very rare by chance ($p < .013$) based on a binomial test (Graphpad, 2023).

These two analyses together provide credible evidence that simply labeling a student as needing an IEP or ELL accommodation resulted in implicit bias among teachers regarding student ability to successfully complete the lesson presented in the simulator. It is important to note that the actual abilities of the simStudents were the same across all six in a "unit of analysis" group.

Discussion

Data samples for this study were aggregated across numerous simulation modules, elementary versus middle school teacher levels, and systematic simStudent variations in Avatars and Names across gender, ethnicity, and assigned accommodations. Patterns were observed when exploring the data with respect to lower or higher ratings being associated with accommodations (either IEP and ELL). This led to the formal tallies of lower versus higher ratings within each of Avatar and Name categories, and the binomial tests to confirm that each indicator of probable bias based on simStudents being labeled as needing an accommodation was extensive enough to be unlikely by chance ($p <.05$). One byproduct of carrying out two parallel analyses was to confirm that there were no great differences in lower vs. higher rating ratios, based on Avatar (picture image) or Name categories, which implies that either can likely be the source of implicit bias. Additional analyses with larger data sets are needed confirm this observed trend and are being collected as this paper is being completed.

Educational Significance

Identification of effective methods for bias reduction in teaching practices that can be applied on a broad scale for teachers throughout the nation is a key foundation for enabling every future member of society to achieve their highest innate and nurtured potential. This project has great potential to accelerate reflections and refinements in the science of teaching and learning.

The recognized importance of socio-emotional stability for the long-term well-being of current teachers and future productive citizens of our society has spotlighted the urgency of research such as that proposed in the *simEquity* project with its focus on mitigation of implicit bias. This project offers the prospect of finding a timely contribution to a problem in many education systems.

Conclusion

This study has demonstrated that a teaching simulator such as simSchool can serve as a test environment where actions of teachers can be recorded and later analyzed without any concern about influences on actions by the researchers running the experiment. The study has further

shown that educators may often have an implicit bias in the form of rating students less likely to succeed in completing their lessons if they are labeled as IEP or ELL and needing accommodations. This rating bias existed even though the students actually performed at the same level as their non-IEP and non-ELL peers in the simulator, in the lesson where the teacher taught all just before completing the ratings of "how likely to succeed" for each. The study builds confidence that implicit teacher bias can exist within the simulator, and thus it is credible to pursue the next step of the research planned for the simEquity project, to determine if targeted feedback and 'debriefing' recommendations provided after simulation runs can help ameliorate implicit biases. Limitations of this study are that the findings, although significant ($p < .05$), were based on a small number of teachers. The analysis should be replicated with a larger group in order to build confidence in the results.

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

References

American Psychological Association (APA), Presidential Task Force on Educational Disparities. (2012). *Ethnic and racial disparities in education: Psychology's contributions to understanding and reducing disparities.* Retrieved from http://www.apa.org/ed/resources/racial-disparities.aspxAnderson, Graham, & Thomas, 2019

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9 – 44.

Bandura, A., & Schunk, D.H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586-598.

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.

Graphpad (2023). Sign and Binomial Test. https://www.graphpad.com/quickcalcs/binomial1/

Gibson, S., & Dembo, M.H. (1984). Teacher efficacy: A construct validation. *Journal of Education Psychology, 76*, 569-582.Gordon, M., & Cui, M. (2014). School-related parental involvement and adolescent academic achievement: the role of community poverty. *Family Relations,63*(5), 616-626. doi:10.1111/fare.12090.

Kitsantas, A. (2012). *Teacher efficacy scale for classroom diversity (TESCD): A validation study.* Vol. 16 no.1, Profesorado, Revista de curriculum y formacion del profesorado. Retrieved June 1, 2015, from http://www.ugr.es/local/recfpro/rev161ART3en.pdf

Knezek, G., & Christensen, R. (2009). Preservice educator learning in a simulated teaching environment. In *Research Highlights in Technology and Teacher Education* (Vol. 1, pp. 161–170).

Medgley, C., Feldlaufer, H., & Eccles, J.S. (1989). Student/teacher relations and attitudes toward mathematics before and after the transition to junior high school. *Child Development, 60*(2), 981-992.

Milner, H.R. (2003). Teacher reflection and race in cultural contexts: History, meanings, and methods in teaching. *Theory into Practice, 42*, 173-180.

Oyerinde, S.A. (2008). *A correlational study of teacher efficacy and culturally responsive teaching techniques in four public middle schools*. Dissertation University of Missouri-Kansas City.

Protheroe, N. (2008). Teacher efficacy: What is it and does it matter? *Principal*, *87*(5), 42-45.

Soodak, L.C., & Podell, D.M. (1994). Teachers' thinking about difficult-to-teach students. *Journal of Educational Research, 88*, 44-51.

Tucker, C.M., Porter, T., Reinke, W.M., Herman, K.C., Ivery, P.D., Mack, C.E., & Jackson, E.S. (2005). Promoting teacher efficacy for working with culturally diverse students. *Preventing School Failure, 50*(1), 29 -34.