Effect of Labeling Simulated Students as Having Special Learning Needs on the Simulated Teaching Behaviors of K-12 Educators

Gerald Knezek
Rhonda Christensen
University of North Texas
United States
gknezek@gmail.com
rhonda.christensen@gmail.com

Abstract: Fifty-five (55) K-12 teachers in two US states completed five repetitions of a 15-minute simulated teaching lesson in which six of their 12 simulated students in a class were labeled as having special learning needs while the other six were not labeled. Analysis of three measures of differential focus in instructional attention revealed that educators teaching within the simulator tended to increase attention ($p < .05$) in the areas where they perceived the greatest needs. That is, they tended to focus increased guidance toward the simulated students that were labeled as having special learning needs. Findings provide credible evidence of the fidelity of the simulated teaching environment as perceived by actual teachers, in that teachers focused actions in the simulator comparable to the manner in which teachers commonly focus guidance in real classrooms, targeting extra assistance where they perceive it is needed. This finding has reconfirmed preliminary indications from a smaller participant pool ($n = 40$) studied in a previous year and has expanded confirmation of the effect to additional types of special learning needs labels in addition to English Language Learner (ELL).
Keywords: simulated teaching, instructional decisions, learners with special needs

Introduction

Prior research has found evidence that teachers engaged in simulated teaching professional development focused more of their attention on simulated students labeled as English Language Learners (ELL), even though the labeled simStudents were actually identical in personality and learning characteristics to their non-labeled clones (Christensen et al., 2023). This finding showed evidence of the fidelity of the simulation environment regarding its alignment with real life teaching situations, in that educators tended to focus their attention in areas where they were informed through reading student profiles prior to beginning a teaching module, that extra attention to learner needs was implied as possibly needed.

Previous studies have shown that a typical classroom teacher makes up to 3,000 important decisions during a day of instruction (Danielson, 1996, 2007), often involving responses to hand raising and body posture cues, but this may also be due to diverse learner attributes made known the teacher in advance through student profiles. Teaching in a K-12 classroom is commonly a dynamic, real-time orchestration of guided learning for an entire group of students with special attention paid by the teacher to nuances for individual learner needs. One goal of studying simulated teaching environments is to test the efficacy of using these environments to confirm that classroom teachers come to "feel" that their role within the simulator is like real classroom teaching. Since quality teachers are known to focus attention toward groups or individual learners with greatest perceived needs (OECD, 2009), supporting evidence for the fidelity (perceived authenticity) of the simulated teaching environment can be provided in the form of measuring increased (meaningful) pedagogical efforts by teachers toward the simulated learners labeled as having specified learning needs. The current paper reports on an attempt to replicate and extend previous findings based using data gathered from 55 teachers participating in a simulated teaching experience.

Background

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. The simEquity project (Christensen & Knezek, 2022) was funded by the US National Science Foundation 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 in order to aid teachers in providing more equitable teaching practices. The project uses a well-designed and tested simulated online teaching environment, simSchool, that is currently being used by many educator preparation programs as well as school districts for professional development.

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 Question

Analysis comparing teaching actions by project participants within the simulator, for simStudents labeled as having a special learning need (IEP/504 or ELL), vs. simStudents with no labeled special needs, were conducted. Three measures including Tip Ratio, Hand Ratio, and Accommodation Effectiveness were the dependent (outcome) variables used as primary measurement indices, while being labeled for a special learning need or not (yes/no) formed the independent variable. The specific research question addressed in this study was:

RQ1: To what extent does labeling simulated students as having specified learning needs result in classroom teachers paying increased, targeted attention to these students during a simulated teaching / learning exercise, as measured by:

a. Tip Ratio (number of targeted actions divided by number of tips provided, calculated per simStudent)
b. Hand Ratio (number of hand raise responses divided by number of hand raises, per simStudent)
c. Accommodation Effectiveness (appropriate actions divided by number of declining states)

Methods

Research Design

Within the simulated teaching environment constructed in simSchool for Year 3 of the simEquity Project, each simulated learner labeled with a special need was actually a clone (unknown to participants) of another simulated learner in the same class, without the special needs label but possessing the same learning characteristics as the unlabeled clone. Each simulated class of 12 simStudents was counterbalanced so that six had special needs labels and six were without labels, which created a treatment/comparison unit for the module completed by each teacher. Operationally, simClasses with varying student personalities and learning attributes were randomly created and assigned to each of 55 teachers completing five repetitions of 15 minutes of teaching in the first module of their simEquity professional development. This created six special learning needs labeled simStudents x 55 teachers = 330 treatment cases and an equal number of comparison / control simStudents without a special learning needs label, across the 55 teachers. The testing of RQ1 then consisted of conducting a series of unpaired T-tests for each of the outcome measures of Tip Ratio, Hand Ratio, and Accommodation Effectiveness, testing mean differences on each of three outcome measures for the portion of the simulated student class labeled as having special needs versus simStudents without a special learning needs label.

Intervention

Each of the participants completed two instructional videos showing how to both use the simulation and read and use the report/feedback data. Following the video modules, participants completed three teaching modules, reteaching five times in each. 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. The data reported in this paper includes only Module 1 data which included five simulated sessions of at least 15 minutes each for a minimum of 75 minutes. As shown in Table 1, the modules were selected for different grade bands to be appropriate for the level the

participants taught.

Table 1. *Modules Completed by Classroom Teachers*

Grade Level Band	Module Content
Elementary teacher modules	Everyone's a Helper Who, Me? A Scientist What is Empathy?
Middle School Modules	Examining Identity and Assimilation Cliques in Schools Media Consumers and Creators
High School Modules	Showing Empathy Sounds of Change Why Local Elections Matter

At the end of each simulation session, participants received graphical feedback displaying degree of success at promoting academic, emotional and equity performance in the simulated class overall, as well as feedback regarding the degree of suitability of the instructional activities selected for each individual simulated student in the class. Figure 1 illustrates an example of the graphical classroom interface overlayed with a profile of one selected student, as well as a graphical, real time, academic performance report for that instance in time. Participants receive the more extensive summative feedback after completing at least 15 minutes of teaching a specified lesson (module) with a specific set of 12 simulated students.

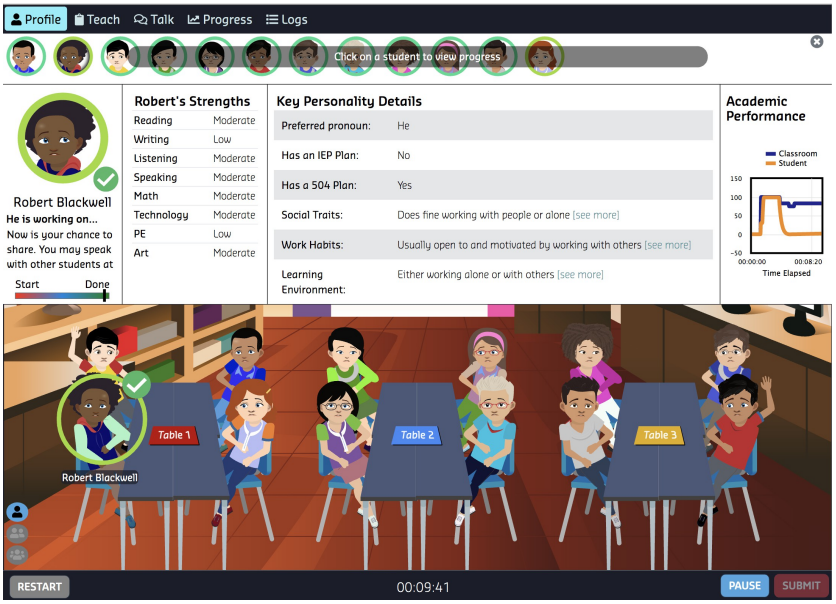

Figure 1. SimSchool classroom highlighting student profiles.

Study Participants

Fifty-five (55) teachers from K-12 schools in California and Texas participated in simulation-based teaching experiences. Teachers were instructed to complete the three modules as assigned based on whether they were elementary, middle school, or high school teachers. For this study data from Module 1, completed after the first two instructional video modules, were gathered across five teaching repetitions of class sessions with feedback provided after each session. Participating teachers included 40 females (72.7%) and 15 males (23.7%). The ethnicity of the participants included 3 (5%) Asian, 3 (5%) Black/AA, 14 (25%) Hispanic or Latinx, 34 (62%) White and 1 (3%) did not reply. Teaching level representation was 12.5% elementary level (n = 7), 25.0% middle school (n = 14) and 62.5% high school (n = 35). The majority of the teachers were female, and the majority taught at the high school level.

Outcome Measures and Analysis Techniques

The three outcome measures used in this study were:

a) Tip Ratio, computed for each teacher, was produced by dividing the number of targeted actions completed by one teacher for Module 1, divided by number of tips provided by the system to the teacher. Each simStudent in this study had an accompanying Tip Ratio recorded by the system.

b) Hand Ratio, computed for each teacher, was produced by dividing the number of hand raise responses completed by a teacher, divided by number of hand raises produced by a particular simStudent. Each simStudent in this study had an accompanying Hand Ratio computed by the system.

c) Accommodations Effectiveness, computed for each teacher, was produced by dividing the number of appropriate actions taken by the number of declining states an individual simStudent displayed. Each simStudent in this study had an accompanying Accommodation Effectiveness ratio computed by the system. Note that missing data appeared more frequently for this measure because it was common that a simStudent might have no declining states during a 15-minute unit of instruction. Declining states data are captured by the system when a student declines in their learning for a defined period of time.

The three measures listed above were used as dependent variables in the analyses performed for this study. Also worthy of detailed description are the independent variables used in the analyses. The three types were each coded as 0 = not present or 1 = present for each simStudent in the class of each teacher. The first type was English Language Learner (ELL), which was a designation randomly assigned to two of the simStudents in each class of 12. The second type was a specifically identified disability accommodation of the category known in the US as Individual Education Plan (IEP) or Title 504. Examples included were anxiety disorder, spectrum disorder, and attention deficit (ADHD). Each was assigned to one or more simStudents in the class with a list of accommodations and/or modifications displayed in the accompanying student profiles. Thus, there were four simStudents with labels for disability accommodations in each simClass and two with labels assigned as ELL. T-test analyses were then completed for each of the following coded independent variable test situations: a) Group mean differences in tip, hand, and accommodation ratios for any simStudent who was labeled with any special needs accommodation (treatment1), versus all other simStudents without any special needs label (comparison1); b) Group mean differences in tip, hand, and accommodation ratios for any simStudent who was labeled with any specific disability accommodation (treatment2), versus all other simStudents without any special needs label (comparison2); and c) Group mean differences in tip, hand, and accommodation ratios for any simStudent who was labeled as ELL (treatment3), versus all other simsStudents without any special learning needs label (comparison3). The results of these analyses are presented in the following section.

Findings

Effect of Labeling a simStudent as Having a Special Learning Need

As shown in Table 2, simStudents labeled as having special learning needs received greater attention by the educators, as measured by Tip Ratio (p =.0005), Hand Ratio (p =.0005), and Accommodation Effectiveness (p = .001). Effect sizes for the three measures ranged in magnitude from d = .35 to d = .50, all of which would indicate that the magnitudes of the differences between the guiding actions taken for simStudents with a special learning needs label versus those with no special learning needs label, based on each of the three outcome measures, all lie in the range of findings that exceed the ES = .3 criterion for magnitudes of effects that would be considered educationally meaningful (Bialo & Sivin-Kachala, 1996).

Table 2. *Contrasts in Simulated Teaching Behaviors Based on Presence or Absence of Any Special Needs Label*

Outcome Measure for Module 1	Mean	N	Std. Dev	Sig	ES
Tip Ratio (number of targeted actions divided by number of tips provided by system, per student)	.1879	329	.237		
Tip Ratio Comparison (No special need label)	.3123	330	.264	.0005	.50
Hand Ratio (number of hand raise responses divided by the number of student hand raises)	.4704	255	.365		
Hand Ratio Comparison (No special need)	.6173	246	.357	,0005	.41

Accommodation Effectiveness (appropriate actions divided by number of student declining states)	.0255	164	.122		
Accommodation Effectiveness Comparison (No special need)	.0936	250	.229	.001	.35

Effect of Labeling a simStudent as Having a Specific Disability Accommodation Need

As shown in Table 3, simStudents labeled as having a disability/accommodation special learning need (IEP/504 in US; ADHD, anxiety disorder, or spectrum disorder) received greater attention by the educators, as measured by Tip Ratio (p =.0005), Hand Ratio (p =.0005), and Accommodation Effectiveness (p = .001). For this analysis, the comparison group included all simStudents without any label as was the case with findings displayed in Table 2, but the treatment group in this case excluded those from Table 1 that were labeled as ELL. Effect sizes for the three outcome measures ranged in magnitude from d = .38 to d = .58, all of which would indicate that the magnitudes of the differences between the guiding actions taken for simStudents with specific learning accommodation labels versus those with no label, based on each of the three outcome measures, all lie in the range of findings that exceed the ES = .3 criterion for magnitudes of differences that would be considered educationally meaningful (Bialo & Sivin-Kachala, 1996). The effect sizes also lie in the range that would be considered closer to "moderate effects" (Cohen's d = .5) rather than "small effects" (Cohen's d = .2) according to guidelines provided by Cohen (1988). Note that the effect sizes reported in Table 2 for each of criterion measures of Tip Ratio, Hand Ratio, and Accommodation Effectiveness were larger than the comparable measure reported in Table 1 – which was based on the broader treatment criteria of any special learning need, including ELL. This implies that the extra attention-getting effect of a specific special need label such as ADHD, anxiety disorder or spectrum disorder may be stronger than being labeled as ELL. This will be further discussed in the next section testing for effect of ELL versus no label among simStudents.

Table 3. *Contrasts in Simulated Teaching Behaviors Based on Presence or Absence of a Specific Disability Accommodation Label (ADHD, anxiety disorder or spectrum disorder)*

Outcome Measure for Module 1	**Mean**	**N**	**Std. Dev**	**Sig**	**ES**
Tip Ratio (number of targeted actions divided by number of tips provided by system, per student)	.1879	329	.237		
Tip Ratio Comparison (No special learning need label)	.3345	220	.275	.0005	.58
Hand Ratio (number of hand raise responses divided by the number of student hand raises)	.4704	255	.365		
Hand Ratio Comparison (No special learning need label)	.6347	141	.404	.0005	.43
Accommodation Effectiveness (appropriate actions divided by number of student declining states)	.0255	164	.122		
Accommodation Effectiveness Comparison (No special learning need label)	.0943	160	.230	.001	.38

Effect of Labeling a simStudent as an English Language Learner (ELL)

As shown in Table 4, simStudents labeled as having an ELL special learning need (English Language Learner) received greater attention by the educators, as measured by Tip Ratio (p =.002), Hand Ratio (p =.002), and Accommodation Effectiveness (p = .002). For this analysis, the comparison group included all simStudents without any label as was the case with findings displayed in Table 2, but the treatment group in this case excluded those from Table 2 that were labeled as having a specific disability accommodation (IEP/504 in US). Effect sizes for the three measures ranged in magnitude from d = .34 to d = .40, all of which would indicate that the magnitudes of the differences between the guiding actions taken for simStudents with ELL labels versus those with no label, based on each of the three outcome measures, all lie in the range of findings that exceed the ES = .3 criterion for magnitudes of differences that would be considered educationally meaningful (Bialo & Sivin-Kachala, 1996). Effect sizes for the test of the importance of an ELL label generally lie in the range that would be considered in between "small effects" (Cohen's d = .2) and "moderate effects" (Cohen's d = .5) according to guidelines provided by Cohen (1988). Note that the average effect size for ELL (Table 4) across the three outcome measures was smaller than for

the comparable measure for specific IEP/504 disability accommodations (Table 3), adding credibility of the implication of findings reported in the previous section that the extra attention-getting effect of a specific special learning needs label such as ADHD, anxiety disorder or spectrum disorder may be stronger than the effect of being labeled as ELL.

Table 4. *Contrasts in Simulated Teaching Behaviors Based on Presence or Absence of English Language Learner (ELL) Label*

Outcome Measure for Module 1	Mean	N	Std. Dev	Sig			ES	
Tip Ratio (number of targeted actions divided by number of tips provided by system, per student)				.1879	329	.238		
Tip Ratio Comparison (No special learning need label)				.2679	110	.234	.002	.34
Hand Ratio (number of hand raise responses divided by the number of student hand raises)				.4704	255	.365		
Hand Ratio Comparison (No special learning need label)				.5940	105	.283	.002	.36
Accommodation Effectiveness (appropriate actions divided by number of student declining states)				.0255	164	.122	.002	.40
Accommodation Effectiveness Comparison (No special learning need label)				.0922	90	.227		

Discussion

Addressing the Research Question

With regard to the primary research question for this study, findings reported in Tables 2 - 4 have confirmed ($p < .05$) that labeling simulated students as having special learning needs results in classroom teachers paying increased, targeted attention to these students during a simulated teaching / learning exercise. This was found to be true when measured by any of a) Tip Ratio (number of targeted actions divided by number of tips provided by system, per student), b) Hand Ratio (number of hand raise responses divided by the number of student hand raises), or c) Accommodation Effectiveness (appropriate actions divided by number of declining states). Effect sizes for the experiment, when any type of special need was combined into a single category, ranged from Cohen's d = .35 to d = .50 (Table 2). Magnitudes of the effect of labeling for a specific disability accommodation such as ADHD, anxiety disorder or spectrum disorder ranged from Cohen's d = .38 to d = .58 for the experiment when only simStudents with specific disability accommodations were included and ELL simStudents were excluded from the analysis (Table 3). Cohen's d ranged from d = .34 to d = .40 for ELL students versus those with no labeled special learning needs, excluding specific disabilities accommodations from the analysis (Table 4). Trends across magnitudes of the effect, depending on which types of special needs are included or excluded, imply that specific disability accommodations such as ADHD, anxiety disorder or spectrum disorder appear to garner the most extensive clustering of teacher focused attention, to a greater extent than if a simStudent is labeled ELL. However, both types of labeling are capable of refocusing teacher attention so that more learning support is directed toward students with any type of identified special learning need, rather than applying equal targeted support efforts to all members of a class. This is interpreted by the project team as a positive finding with respect to the broader project goal of advancing equitable teaching practices based on identified (differing) student needs, rather than simply training each teacher to devote equal time and effort to all students. These may be considered "good biases" or "good tendencies" in the broader context that equity does not necessarily mean equal treatment for all.

Differential Effects of Gender on Responses to Hands Raises.

Nevertheless, some evidence has emerged from more detailed disaggregated subgroup analysis of the data used to produce Tables 2 – 4 that questionable implicit biases may remain among teachers even as they seek to focus help where it is needed. Specifically, through a series disaggregation analyses (not shown) based on gender of the teacher and gender of the simStudents, Tables 5 and 6 were produced to summarize emergent trends.

Table 5 can be viewed as illustrating a second level effect (2 x 2) of variations in the gender of simStudents cross-tabulated with variations in the gender of the teacher. Effect sizes are listed along with p – levels to contrast the magnitudes of the effect of different combinations of gender for teachers and simulated students. As shown in Table 5, both female and male teachers tended to pay more attention to (take actions as a result of) hand raises by simStudents flagged as having special learning needs when the simStudents were labeled male. However, as is also shown in Table 5, female teachers tended to *not* take greater amounts of actions to hand raises to any significantly ($p < .05$) degree for simStudents with special needs – if the simStudents were labeled as female.

Table 5. *Differences in Responses to Hand Raises (HandRatio) by Female and Male Teachers for SimStudents Labeled Female versus Male in the Context of Identified Special Learning Needs*

Teachers	**F simStudents (N = 330)**	**M simStudents (N =330)**	**All simStudents (N = 660)**
F (N = 40)	ES =.24 (p =.103 NS)	ES = .42 (p = .006)	ES = .31 (p = .003)
M (N = 15)	ES = .54 (p = .027)	ES = .76 (p = .005)	ES = .65 (p = .001)
	ES = Cohen's d	p = 2-tailed t	

Differential Effects of Gender on Responses to Specific Learning Needs

Possibly the anomaly shown in Table 5, where female teachers did not to a greater extent ($p < .05$) respond to female simStudents with special learning needs who raised their hands, could have occurred because these teachers had previously studied the profiles of females with special learning needs and were being effective in accommodating female special needs (Table 6) whether simStudents raised their hands or not. As shown in Table 6, male teachers in this study were not significantly ($p < .05$) greater in Accommodation Effectiveness for simStudents identified with specific learning needs, for any of the three disaggregation categories of female simStudents, male simStudents, or when both male and female simStudents were combined. However, female teachers were significantly ($p < .05$) greater in Accommodation Effectiveness for all three disaggregation categories of female, male, and combined simStudents if they had been labeled in student profiles as having specific learning needs. The effect size indicating magnitude of increased attention due to labeling was largest (Cohen's d = .46) for the disaggregated group of female simStudents, the same disaggregated group for which female teachers had been identified as not paying extra attention ($p < .05$) to hand raises from female simStudents in Table 5.

Table 6. *Differences in Accommodation Effectiveness by Female and Male Teachers for SimStudents Labeled Female versus Male in the Context of Identified Special Learning Needs*

Teachers	**F simStudents**	**M simStudents**	**All simStudents**
F (N = 40)	ES =.46 (p =.005)	ES = .36 (p = .027)	ES = .41 (p = .001)
M (N = 15)	ES = .05 (p = .874 NS)	ES = .41 (p = .161 NS)	ES = .20 (p = .333 NS)
	ES = Cohen's d	p = 2-tailed t value	

Avoiding the Assumption of Simplicity in Implicit Bias

Tables 5 and 6 illustrate that there is likely great complexity in tracing the origins of true implicit biases in classroom teaching practices, and in determining which specific ones need to be targeted for remediation versus alternative gender-aligned practices that may take a different path but be equally effective in the end, toward fostering student learning goals. In this specific example, it appears that there may indeed be prototypical tendencies that differ for male and female practicing teachers, for activities such as responding to hand-raising versus detailed study in advance and planning for known special needs of learners.

This latest finding calls into question where the boundaries lie for the concept of implicit bias. One commonly accepted description of implicit bias is that it includes subconscious feelings, attitudes, prejudices, and stereotypes an individual has developed due to prior influences and imprints throughout their lives (Shah & Bohlen, 2023). Using this delineation, findings in this study would appear to fall under the heading of implicit bias.

However, the connotation of the term implicit bias, in the opinion of the authors, is almost always negative, whereas many of the findings reported in this study might simply be considered equitable teaching practices.

Conclusion

The primary research question in this study of whether practicing teachers tend to immerse themselves in a simulated teaching environment in a manner comparable to a real teaching environment – has been answered based on the extent to which they tend to focus increased efforts for learners where there are special needs indicated. The answer is "yes." This provides credible evidence of the fidelity of the simulated teaching environment as perceived by actual teachers, in that teachers take actions in the simulator comparable to what teachers commonly do in a real classroom, to target extra assistance where they perceive it is needed. This finding has reconfirmed preliminary indications from a smaller participant pool (n = 40) (Christensen et al., 2023) and has expanded confirmation of the effect to other types of special learning need labels beyond ELL previously examined, based on a sample of 55 teachers providing data. Findings from this study have also further illustrated the complexity of identifying true biases in teaching practices versus desirable focused efforts by teachers, in the context of recognition that that equity in teaching practices is not always the same as treating as all students equally.

Acknowledgement: This research was funded in part by US National Science Foundation Grant #2118849.

References

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.

Christensen, R. & Knezek, G. (2022). Impacting culturally responsive teaching strategies by decreasing bias through simulation experiences. *Excellence and Innovation in Teaching and Learning, 7*(2), 39 – 56. Doi: 10.3280/exioa2-2022oa15077.

Christensen, R., Knezek, G. & Norton, S. (2023). Using framing effect to detect implicit bias in educators through a simulated teaching environment. In T. Bastiaens (Ed.), *Proceedings of EdMedia + Innovate Learning* (pp. 949-957). Vienna, Austria: Association for the Advancement of Computing in Education (AACE).

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Danielson, C. (1996, 2007). *Enhancing professional practice: A framework for teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

OECD (2009). *Creating effective teaching and learning environments: First results from TALIS – ISBN 978-92-64-05605-3.* https://www.oecd.org/berlin/43541655.pdf

Kovanovic, V., Mazziotti, C., & Lodge, J. (2021). Learning analytics for primary and secondary schools. *Journal of Learning Analytics 8*, 1-5. Doi: 10.18608/jla.2021.7543

Shah, H.S., & Bohlen, J. (2023). *Implicit Bias*. US National Institute of Health National Library of Medicine. PMID: 36944001. https://www.ncbi.nlm.nih.gov/books/NBK589697/#:~:text=Implicit%20bias%20includes%20the%20subconscious,and%20imprints%20throughout%20their%20lives.