Effectiveness of Digital Simulation-Based Learning Approach in Optimizing Students' Understanding of Queueing Models Using Real-Life Data*

Arin Berliana Angrenani[1*], Dian Kurniati[2], Susi Setiawani[3], Rafiantika Megahnia Prihandini[4], and Ngizatul Afifah[5]

***Abstract*—This study examines indications of the effectiveness of a digital simulation-based learning approach in supporting students' understanding of queueing models using real-life data. A quasi-experimental one-group pretest–posttest design, supported by qualitative interview data, was conducted with 31 undergraduate mathematics education students at the University of Jember. The ExtendSim software was used to create interactive queueing simulations that allowed students to explore parameters such as arrival rate, service rate, and waiting time. Validity and reliability tests were conducted using item–total (Pearson) correlations and Cronbach's alpha, with results indicating high validity ($r > 0.5$, $p < 0.05$) and high internal consistency ($\alpha > 0.80$). A paired *t*-test showed a statistically significant increase in scores within this sample ($t = 8.89$, $p < 0.001$). Students' perceptions of the simulation were highly positive, with an average Likert score of 3.23 (very high). Qualitative interviews further indicated that the simulations helped students visualize queue dynamics and relate theoretical concepts to real-life contexts. There were also indications of increased motivation, engagement, and computational thinking skills; however, these findings are limited by the single-site sample and the one-group study design.**

I. INTRODUCTION

QUEUEING models constitute an important component of applied mathematics, used to analyze service systems across various fields such as healthcare, transportation, industry, and public services. Efficiency in queue management has a direct impact on user satisfaction and operational performance [1]. In the context of mathematics education, understanding queueing models is an essential competency that connects theoretical mathematical structures with real-world decision-making. However, instructional practices in mathematics education often emphasize theoretical derivations without providing sufficient exposure to authentic, data-driven applications [2].

This limited practical exposure can lead to challenges in students' conceptual understanding, particularly when attempting to relate abstract queueing formulas to real operational systems. As a result, a gap frequently emerges between students' theoretical knowledge and their ability to apply queueing concepts in practical contexts. Meanwhile, the advancement of digital technology has created opportunities for integrating interactive learning tools that may help bridge this gap, including simulation-based learning environments [3], [4].

Studies indicate that simulation-based learning can support motivation and higher-order thinking skills, particularly when students engage with computational models in interactive environments. Prior research also suggests that virtual laboratory and computational thinking–based approaches may enhance engagement and conceptual reasoning in STEM fields [5]. Digital simulations allow students to manipulate queueing system parameters using real-world data and observe system behavior dynamically. Through these interactions, students can visualize queue flows, evaluate performance metrics, modify arrival and service rates, and conduct scenario-based explorations. Despite these benefits, studies focusing specifically on the use of digital simulation for teaching queueing models in mathematics education remain limited.

To address this gap, the present study examines indications of the effectiveness of a digital simulation-based learning approach in supporting students' understanding of queueing models using real-life data, including examples such as gas stations, shopping centers, and restaurants. A one-group pretest–posttest design is used to identify changes in learning outcomes within this specific cohort after the implementation of digital simulations. Additional qualitative insights are gathered through classroom observations and interviews to complement the quantitative findings. The study aims to contribute to the development of contextual, technology-enhanced learning approaches in applied mathematics that align with the demands of 21st-century education.

II. METHODS

A. Queueing Theory

Queueing theory is a field of applied mathematics that examines service systems involving waiting lines. It helps analyze how services operate and how resources can be managed efficiently [6]. This theory is widely used to support decision-making in sectors such as healthcare, transportation, industry, and public services. A queueing model describes a system based on several elements, including how customers arrive, how they are served, the number of servers, system capacity, and the discipline used to determine the order of service. From these elements, key performance indicators can be calculated, such as queue length, waiting time, and server utilization [7].

*This work was supported by LP2M UNEJ

A. B. Angrenani, D. Kurniati, S. Setiawani, R. M. Prihandini, N. Afifah are with the Department of Mathematics Education, Faculty of Teacher Training and Education, University of Jember, Indonesia

*Correspondending Author: aberlianaa@unej.ac.id

Manuscript received November 10, 2025; accepted December 10, 2025.

A queueing system generally consists of: customers (people, items, or data requiring service), the arrival rate (λ) as the average number of customers arriving per unit time, queue discipline (e.g., FIFO, LIFO, priority rules), servers (resources providing service), the service rate (μ) as the average number of customers served per unit time, system capacity (maximum number of customers allowed in the system), and the number of servers operating simultaneously [8].

Queueing systems are commonly represented using Kendall's notation A/B/s, where A denotes the distribution of arrivals, B denotes the distribution of service times, and s is the number of servers. For example, an M/M/1 model represents exponential (Markovian) arrivals and service times with one server, while M/M/s represents exponential arrivals and service times with multiple servers.

B. Validity and Reliability Testing

Instrument testing was conducted to examine the quality of the questionnaire administered in this study. Validity refers to the degree to which an instrument accurately measures the construct it is intended to measure [9]. Item validity was assessed using item–total correlation, in which each item was correlated with the total score of its respective construct (PU, PEOU, ATU, BI). Reliability testing, on the other hand, evaluates the consistency of an instrument when applied under similar conditions [10]. A reliable instrument produces stable and dependable results.

Validity Test:
H_0: the item is not valid (not correlated with total score)
H_1: the item is valid (correlated with total score)

The Pearson Product–Moment correlation coefficient is defined as:

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} \tag{1}$$

where: r = Pearson correlation coefficient, N = number of respondents, X = item score, Y = total score.

TABLE I
INTERPRETATION OF VALIDITY COEFFICIENT

Range of r	Interpretation
$0.80 < r \leq 1.00$	Very High
$0.60 < r \leq 0.80$	High
$0.40 < r \leq 0.60$	Moderate
$0.20 < r \leq 0.40$	Low
$r \leq 0.20$	Very Low

Reliability was examined using Cronbach's Alpha, which measures internal consistency among items within each construct. The Cronbach's Alpha formula is:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_{total}^2}\right) \tag{2}$$

where: α = Cronbach's Alpha, k = number of items, σ_i^2 = item variance, σ_{total}^2 = total score variance.

TABLE II
INTERPRETATION OF RELIABILITY COEFFICIENT

Range of α	Interpretation
$\alpha > 0.90$	Very High
$0.80 < \alpha \leq 0.90$	High
$0.70 < \alpha \leq 0.80$	Moderate
$0.60 < \alpha \leq 0.70$	Low
$\alpha \leq 0.60$	Very Low

C. Normality

Normality testing was conducted to determine whether the data met the assumptions required for parametric analysis. The Shapiro–Wilk test was applied to the difference scores (posttest minus pretest), which is the appropriate assumption check for paired-sample analysis [11]. The Shapiro–Wilk statistic is:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(n+1-i)} - a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \tag{3}$$

where $x_{(i)}$ are the ordered sample values, a_i are constants derived from the expected values of order statistics of a standard normal distribution, and $\bar{x}$ is the sample mean.

D. Paired t-Test

A paired t-test was used to analyze whether there was a statistically significant difference in students' scores before and after the learning intervention. This test compares the mean of the paired differences and is appropriate when the difference scores are normally distributed [12]. The test statistic is given by:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}, \tag{4}$$

where $\bar{d}$ is the mean of the paired differences, s_d is the standard deviation of the differences, and n is the number of paired observations.

III. RESULTS AND DISCUSSION

A Likert-scale questionnaire was used to measure students' perceptions of the digital simulation-based learning activity using ExtendSim. The instrument consisted of four constructs: *Perceived Usefulness (PU)*, *Perceived Ease of Use (PEOU)*, *Attitude Toward Use (ATU)*, and *Behavioral Intention (BI)*. To ensure instrument quality, validity and reliability analyses were conducted.

TABLE III
RESULTS OF VALIDITY TESTING

Indicator	Corr.	Sig.	Desc.	Indicator	Corr.	Sig.	Desc.
PU1	0.797	0.000	Valid	ATU1	0.492	0.006	Valid
PU2	0.718	0.000	Valid	ATU2	0.674	0.000	Valid
PU3	0.881	0.000	Valid	ATU3	0.747	0.000	Valid
PEOU1	0.590	0.001	Valid	BI1	0.782	0.000	Valid
PEOU2	0.590	0.001	Valid	BI2	0.721	0.000	Valid
PEOU3	0.542	0.002	Valid	BI3	0.612	0.000	Valid

Table III shows that all items across the PU, PEOU, ATU, and BI constructs have significance values below 0.05.

This indicates that each item is statistically valid based on item–total correlation criteria. PU items fall within the high to very high correlation range, PEOU and ATU items range from moderate to high, and BI items show high correlations. Overall, all indicators are appropriate for measuring their respective constructs.

TABLE IV
RELIABILITY TEST RESULTS (CRONBACH'S ALPHA)

Construct	Cronbach's Alpha	Description
PU	0.802	High Reliability
PEOU	0.813	High Reliability
ATU	0.801	High Reliability
BI	0.813	High Reliability

Table IV shows that all constructs have Cronbach's Alpha values greater than 0.80, which indicates high reliability and suggests strong internal consistency among items within each construct.

TABLE V
STUDENT PERCEPTIONS TOWARD EXTENDSIM SIMULATION

No.	Indicator	Mean	Description
1	Helps me understand queueing concepts	3.29	Very High
2	Speeds up my learning process	3.16	Very High
3	Clarifies conclusions from simulations	3.19	Very High
4	Easy for beginners to use	3.13	Very High
5	Easy to understand the interface	3.19	Very High
6	Steps in simulation are easy to follow	3.29	Very High
7	I enjoy learning with digital simulation	3.16	Very High
8	Makes learning more engaging	3.29	Very High
9	Increases my motivation to learn	3.23	Very High
10	I want to use ExtendSim again	3.19	Very High
11	I will recommend this method to others	3.29	Very High
12	Open to using simulation in other courses	3.32	Very High
Average		3.23	**Very High**

Table V indicates that students expressed very positive perceptions about the simulation-based learning activity. All items received "very high" mean scores, implying that the students felt the simulation helped them understand concepts, accelerated learning, and made lessons more interesting. The students also reported high motivation and enthusiasm, showing a strong intention to reuse ExtendSim in future learning.

After implementing the simulation-based learning activity using ExtendSim, the performance of the students was evaluated through pre–test and post–test assessments to measure their understanding of queueing theory concepts before and after the intervention. The data collected were descriptively and inferentially analyzed to identify improvements in learning outcomes and determine the statistical significance of the observed changes. Descriptive statistics were used to summarize the central tendency and variability of both pre–test and post–test scores, while inferential analysis was conducted using a paired t-test to examine whether the mean difference between the two measurements was statistically significant.

The descriptive statistics presented in Table VI show a notable improvement in the learning outcomes of the students after implementing the simulation-based learning approach using ExtendSim. The mean score increased from 66.45 on the pre–test to 89.03 on the post–test, indicating a substantial learning gain of approximately 22.6 points. The median score

TABLE VI
DESCRIPTIVE STATISTICS OF PRETEST AND POSTTEST SCORES

Statistic	Pretest	Posttest
N	31	31
Mean	66.45	89.03
Median	65.00	90.00
Minimum	40	65
Maximum	100	100
Standard Deviation	13.65	7.78

also increases from 65.00 to 90.00, confirming a general upward shift in student' performance. In addition, the standard deviation decreased from 13.65 to 7.78, suggesting that the post–test scores were more homogeneous and that the learning outcomes became more consistent across participants. These results simply show that the ExtendSim-based simulation effectively improved students' understanding of queueing theory concepts, reduced performance variability, and enhanced overall mastery of the material. To visualize the difference in score distribution before and after the learning intervention, a box plot was created to compare the spread and central tendency of pre–test and post–test scores. This visualization provides a clear depiction of how simulation-based learning using ExtendSim affected student' performance levels.

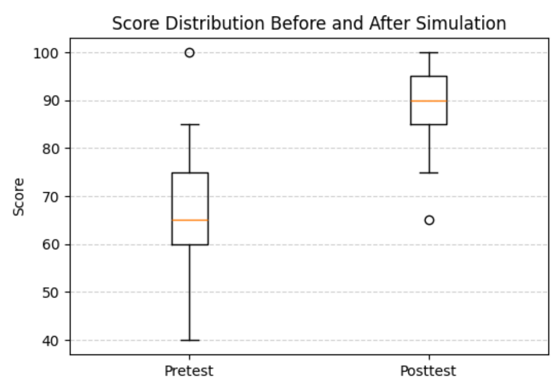

Fig. 1. Boxplot of Pretest and Posttest Score Distribution

As shown in Figure 1, the entire distribution of post–test scores shifted upward compared to the pre–test distribution. The median line of the post–test box is notably higher, and the interquartile range is narrower, indicating that students' scores not only improved but also became more consistent. Few lower outliers appeared in the pre–test data, meanwhile post–test scores clustered near the upper end of the scale, showing reduced variability and higher overall achievement. This visual evidence supports the descriptive findings in Table VI.

Normality testing was performed before the paired t-test. The Shapiro–Wilk test was applied to the difference scores (post–test minus pre–test).

TABLE VII
SHAPIRO–WILK NORMALITY TEST ON DIFFERENCE SCORES

Statistic	W	p-value
Differences (Post–Pre)	0.9576	0.2511

Table VII shows that the difference scores follow a normal distribution ($p = 0.2511 > 0.05$), satisfying the assumption for applying the paired t-test.

TABLE VIII
PAIRED t-TEST FOR PRETEST AND POSTTEST SCORES

t	df	p-value
8.9948	30	5.079×10^{-10}

Table VIII shows a statistically significant difference between pretest and posttest scores ($p < 0.001$). Within the context of this one-group sample, this indicates a statistically detectable increase in students' scores following the learning activity, though causal conclusions cannot be drawn due to the design limitations.

To complement the quantitative findings, interviews were conducted with three students, two with high post–test scores and one with moderate performance, to gain qualitative insights into their learning experiences.

TABLE IX
SUMMARY OF STUDENT INTERVIEW RESULTS

Respondent	Post-Test Level	Statement Summary
M1	High	The simulation helped in understanding queueing flow and in observing how input changes affected system output.
M2	High	The simulation increased interest and supported faster understanding compared to theoretical explanation alone.
M3	Moderate	Initial difficulties were reduced after repeated use, leading to clearer visualization of queueing concepts.

The interview results align with the quantitative patterns. The students described the simulation as helpful for visualizing queue dynamics and for making abstract concepts more concrete. The student with a moderate score also expressed positive change, noting that the practice with simulation made the concept easier to grasp. These qualitative insights suggest indications of increased engagement, motivation, and conceptual clarity within this sample.

An additional observation supporting these indications is the increase in the proportion of students choosing queueing theory as their mini-project topic in the Operations Research course—from 46.53% to 64.52%. This trend may reflect growing interest and perceived relevance of the topic following the simulation-based activity.

IV. CONCLUSIONS

The findings of this study indicate positive patterns of improvement in students' understanding within the scope of the one-group pretest–posttest design used. Validity and reliability analyses showed that all questionnaire items met the criteria for item–total validity and demonstrated high internal consistency ($\alpha > 0.80$). The paired t-test revealed a statistically significant increase in scores within this sample after using the simulation ($p < 0.001$). Qualitative interview results supported the quantitative findings by indicating that students experienced clearer conceptual understanding, higher motivation, and greater interest when engaging with the simulation-based activities.

Within the limitations of this design, these results suggest that integrating ExtendSim as a digital learning tool may support students' engagement and conceptual understanding in applied mathematics, particularly queueing theory. However, several limitations should be noted. The one-group pretest–posttest design does not permit strong causal inference due to the absence of a comparison group. The sample size was relatively small and drawn from a single institution, limiting generalizability. In addition, the psychometric evaluation of the questionnaire was restricted to item–total correlation and internal consistency without further validation procedures.

Future research may employ larger and more diverse samples, include control or comparison groups, and incorporate more comprehensive psychometric analyses to strengthen the robustness and generalizability of the findings.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Institute for Research and Community Service (*Lembaga Penelitian dan Pengabdian kepada Masyarakat* – LP2M), University of Jember, for providing research funding and support for this study. The assistance and facilities provided by LP2M greatly contributed to the successful completion of this research project.

REFERENCES

[1] Y. Moteki, "Factors Determining Customer Satisfaction With Counter Services of Local Government Offices: On-Site Survey Examining Service Quality at Higashihiroshima City Hall, Japan", 2021. https://doi.org/10.21203/rs.3.rs-1118972/v1

[2] N. Mishra and P. Aithal, "Modern Multidisciplinary Education: Challenges and Opportunities of Modern Learning Pedagogy", International Journal of Case Studies in Business, IT, and Education, p. 269-280, 2023. https://doi.org/10.47992/ijcsbe.2581.6942.0319

[3] Q. Ren, Y. Wang, Q. Zheng, L. Ye, X. Zhou, & L. Zhang, "Survey of Student Attitudes Towards Digital Simulation Technologies at A Dental School in China", European Journal of Dental Education, vol. 21, no. 3, p. 180-186, 2016. https://doi.org/10.1111/eje.12198

[4] Y. Shi, "Investigation and Analysis on The Effect of Digital Teaching Means in Western Universities", Journal of Innovation and Development, vol. 6, no. 1, p. 50-58, 2024. https://doi.org/10.54097/ncx7e648

[5] T. Wu, E. Sarwono, & Y. Huang, "Incorporating computational thinking into virtual laboratories to enhance learning motivation, engagement, and higher-order thinking skills", Journal of Computer Assisted Learning, vol. 41, no. 2, 2025. https://doi.org/10.1111/jcal.70017

[6] J. Liu, Y. Zhou, & J. Chen, "Customer Segmentation and Ex Ante Fairness: A Queueing Perspective", Production and Operations Management, vol. 32, no. 10, p. 3246-3265, 2023. https://doi.org/10.1111/poms.14033

[7] J. Wang and J. Li, "A Single Server Retrial Queue With General Retrial Times and Two-Phase Service", Journal of Systems Science and Complexity, vol. 22, no. 2, p. 291-302, 2009. https://doi.org/10.1007/s11424-009-9164-8

[8] I. Adan, I. Kleiner, R. Righter, & G. Weiss, "FCFS Parallel Service Systems and Matching Models", Performance Evaluation, vol. 127-128, p. 253-272, 2018. https://doi.org/10.1016/j.peva.2018.10.005

[9] S. Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta, 2019.

[10] I. Ghozali, *Aplikasi Analisis Multivariate dengan SPSS*. Semarang: Badan Penerbit Universitas Diponegoro, 2018.

[11] S. Santoso, *Validitas & Reliabilitas dalam Riset*. Malang: Universitas Negeri Malang Press, 2017.

[12] L. Herdiman, I. Priadythama, T. Rochman, H. Setyanto, & D. Setiawan, "Evaluation of Work Posture When Assembling A Study Chair in Industrial Engineering Design Practicum 2", E3S Web of Conferences, vol. 465, p. 02026, 2023. https://doi.org/10.1051/e3sconf/202346502026