icrrd
Quality Index Research Journal
INTERNATIONAL CENTER FOR RESEARCH AND RESOURCES DEVELOPMENT article
ISSN Number: 2773-5958, https://doi.org/10.53272/icrrd, www.icrrd.com

The Effectiveness of a Simulation-Based Training Program to Learn Mathematics and Statistics

Said Ali Al -Hadhrami[1*], Mohammed Nasir Al Riyami[2]

[1] *Department of Mathematical and Physical Sciences, University of Nizwa, Oman*
[2] *Department of Education and Cultural Studies, University of Nizwa, Oman*

**Corresponding author; Email: saidhadhrami@unizwa.edu.om*

Received: 05 June 2024 **Accepted:** 25 August, 2024
Revision: 10 July 2024 **Published:** 01 September 2024. **Vol**-5, **Issue**-2

Cite as: Hadhrami, SA.; Al Riyami, MN. (2024). The Effectiveness of a Simulation-Based Training Program to Learn Mathematics and Statistics. *ICRRD Journal, 5*(3), 154-170.

Abstract*:* Simulation-based learning has become popular approach that mimics real-world environment and provides students practical insights. In this paper, we report the results of a training program implemented at University of Nizwa in Oman. The program subjected 30 students from statistics specialization to some mathematical and statistical topics by using Monte Carlo simulation. The program was carried out for 48 hours and the students' performance was assessed and their opinions were taken through a questionnaire to evaluate the training from their perspective. The results demonstrate high students' performance. Factor analysis was used to extract some factors from several measurable variables in the questionnaire. The main extracted factors are: usefulness, enjoyment, satisfaction, ease, ability and availability. These factors are then used to build two regression models to predict the "intention to use simulation" and the "improvements in learning skills"

Keywords: Monte Carlo simulation, correlation, regression, factor analysis, mediation analysis.

Introduction

In some cases, statistical concepts, theories and ideas are misunderstood by students or hard to figure out the meaning and importance. However, Computer simulation is one of the techniques that used to deepen the understanding, develop better statistical reasoning and add fun to learning. Although simulation is used to solve hard problems or those which have no analytical solutions, it may also be used to solve even simple problems. Simulation is implemented in various fields such as health care, aviation, agriculture, economic, mathematics, physics, chemistry, statistics and others. Monte Carlo simulation uses random sampling and statistical modeling to estimate mathematical functions and mimic the operations of complex systems.(Harrison RL,2010). It is useful computational tools that provide efficient and practical algorithims to solve a wide range of scientific problems (Ronad & Franklin Simulation can be used in education as a teaching schemes and the learner is given chance to design and explore different scenarios, rather than just do calculations. In addition, simulation can make abstract science phenomena more accessible and visible to learners and it helps them visualize the phenomenon that might otherwise be difficult to depict. Moreover, it can animate dynamic changes in scientific processes that are difficult to infer from static illustrations found in the textbooks. Furthermore, this approach promotes better comprehension, allowing learners to figure out the fundamentals without becoming overwhelmed by complexity.

Literature review

Monte Carlo simulation relys on repeated random sampling to compute results of interest. It helps teach basic concepts. Several researches recommended using simulation to help both teachers and students to understand statistics concepts and theory especially the difficult or abstract concepts. Dambolena (1986) gave examples of the use of computer simulation to help students with difficult statistical concepts. Simulation is used to teach various topics in statistics, including sampling distribution of different statistics, central limit theorem, confidence intervals, hypotheses testing, and the power of tests (West & Ogden, 1998). Gordon et al. (1989) described the use of computer graphics simulations to enhance student understanding of sampling distributions of proportions, difference of sample means, difference of sample proportions, and the distribution of sample variances. Webster, Elaine (1992) compared five selected statistical software packages on their performance on enhancing statistics instructors' ability to teach difficult topics. They concluded that teaching is enhanced by appropriate software. Aberson et al. (2000) studied the efficiency of a tutorial to illustrate sampling distribution and central limit theorem. They let students draw samples from different distributions and observe the shapes of sampling distributions. The learner rated the tutorial as easy to use and understand. Mills (2002) also recommended simulation as an effective teaching strategy for teaching statistics concepts. Garfield, et al. (2007) concluded that providing applications to students can increase interaction, improves knowledge retention and promotes an enjoyable atmosphere for learning, rather than relying on theorems and memorization. Ted Hodgson and Maurice Burke (2001) used simulation as an instructional tool that can promote a deep conceptual understanding of statistics. Schützler and Gaschler (2002) explored preferences and efficiency of learners through interactive simulations versus static pictures in acquiring statistics knowledge of Cohen's d and standard normal distribution. Reidar et al. (2007) developed a simulation tool to illustrate the process of sampling and the relationship between the population from which samples are drawn and the distribution of a sample mean. The book by Thomas & Jeffrey(2013) includes a wide range of topics related to Monte Carlo simulation and used R code to illustrate the examples. They examined bias, efficiency, and uncertainty in a visual way**.** Genevieve & Carter (2013) presented a guide to Microsoft Excel for performing Monte Carlo simulations exercises. They showed through simulation that least squares estimators in regression are unbiased and constructed confidence interval estimates of least squares estimators. Stephen & Alan (2014) presented animations that are story-based, that can be useful to teach sampling distributions in the introductory statistics classroom. Braun, et al.(2014) emphasized that teaching statistics with numerical examples and simulation is an effective way to help students crystallize their knowledge about statistical concepts. Sigal & Chalmers (2016) demonstrated Monte Carlo Simulation designs for a classroom and how simulation acquainted students with complex statistical concepts. Rui Manuel da Costa Martins (2018) used the Birthday problem to allowstudents understand the basic reasoning behind simulation and explore its potential. Mikuláš Gangur & Milan Svoboda (2018) used Microsoft Excel to generate random numbers and used them for presenting Bayes theorem. The study of Timur (2019) aimed to investigate the learning environment supported by game and simulation. Simulations were used for visualization a large number of experiments. The games were used to evaluatethe basic probability knowledge of the prospective teachers. Chernikova,O et al.(2020) studied simulation-based learning and concluded that simulations are effective method to facilitate learning of complex skills. Alex & Michael (2020) introduce a new tool in R to help students explore probability calculations through simulation in a fun context. They presented some sample activities for helping students better understand and make in-game probabilistic decisions. Jorge Antonio et al.(2021) used the popular children's game, Chutes and Ladders, to illustrate Monte Carlo simulation.This progression allows to explore the nature of chance in board games through the aid of computer simulation. Thiesmeier et al. (2024) discussed the importance of embracing statistical thinking in public health research and education and the role of simulation to strengthen statistical reasoning. Batanero et al. (2024) developed a synthesis of the most outstanding research on the teaching and learning of probability in the past years.

Objectives of the study

The objectives of the study are:
1. Investigating the efficiency of the simulation – based teaching of statistics and mathematics.
2. Extracting some significant factors that influence the simulation scheme in teaching mathematics and statistics.
3. Building regression models from the extracted factors to predict the level of learning skills and predict the intention to use simulation in learning statistics and mathematics.

Significance of the study

The study will contribute to the field of education especially for higher education as follows:

1. Responding to one of the modern trends in education which is teaching statistics through simulation.
2. Providing numerical evidence of the effect of some factors on the simulation- based teaching.
3. Encouraging teachers to adopt simulation- based approach to convert some statistical theory to numerical approach using computer simulation.
4. Assisting to design and develop curricula by adopting simulation -learning approach.
5. It may open the way for researchers to conduct more studies of such teaching scheme and develop new teaching methods.

The limits of the study

The limitations of the study were as follows:

1. Objective limits: It was limited to learning some statistical and mathematical topics through simulation.
2. Human limitations: It was limited to the students of statistics at university of Nizwa in fall 2023 in Sultanate of Oman.
3. Time limitations: It was conducted in fall 2023 in the period from September 2023 to December 2023.

The methodology

The study designed a training program to teach some statistical and mathematical topics by Monte Carlo simulation. The study took place at University of Nizwa in Oman in the period from September 2023 to December 2023.

The majority of participants finished at least 100 credit hours out of 132 credit hours of their academic plan before undergoing this simulation program. They are of different levels with average cumulative GPA 2.88 and standard deviation of 0.49. The minimum cumulative GPA of the participants is 2.03 and maximum cumulative GPA is 3.79.

The participants studied 4 hours every week for 15 weeks including exams periods. The sessions took place at a computer lab facilitated with R program. Their performance was assessed by exams and assignments and their opinions were taken by a google form questionnaire.

The training program included the following topics:

1- Learning how to use R studio program. This include basic commands, loops, if statement as well as ready syntax.
2- Learning different methods of generating random numbers from different statistical distributions.
3- Using the loops in R to reproduce samples from different distributions which then can be plotted to notice the distribution of those statistics. Students then use the goodness of fit techniques to detect the significance of the distribution.
4- Investigating the behavior of different statistics and their properties such as biasness, consistency and relative efficiency.
5- Illustrating Central limit theorem as well as some popular theorems and inequalities by generating large samples from different distributions.
6- Approximating probability, expectation and standard error of different estimators.
7- Approximating the values of integrals as well as the values of special constants; π, e and ϕ.
8- Visualizing different distribution with different values of parameter and depicting sampling distributions.
9- Understanding the role of the distribution's parameters in controlling the distributions and check whether they are location, scale, shape or rate parameters.
10- Investigating the behavior of the distribution when the sample size gets increased and observe the limiting distribution. For instance, several distributions get closer to the normal distribution under some conditions such as increasing degrees of freedom or sample size.
11- Constructing confidence intervals. Confidence intervals were obtained for several statistics such as the mean, proportion, the variance, the difference of means, difference of proportions and the ratio of two variances. It gives chance to see whether the interval encloses the true parameter or not. It also gives chance to see whether the interval encloses the true parameter. One can calculate the relative frequency of the cases when the parameter is falling in the confidence interval and how sample size effects on the length of the interval.

Validity and reliability of the questionnaire

Validity of an instrument is a measure of how well the measuring instrument performs its function (Anastasi and Urbina, 1997). Factor analysis can be used to check the validity. It explores the relationships between the items of the survey and identify the total number of dimensions represented on the survey (Knekta E, et al. 2019). In this study, factor analysis was carried out and extracted the following factors: usefulness, enjoyment, satisfaction, ease, ability and availability. The details will be discussed later in section 9.3.

To ensure the reliability of the questionnaire, the researchers calculated the Cronbach's alpha coefficient for the factors, using the SPSS program. The scale's reliability coefficients are summarized in Table 1

Table 1 Cronbach's reliability

section	Cronbach
usefulness	0.952
enjoyment	0.903
satisfaction	0.871
ability	0.832
ease	0.843
availability	0.616
skills	0.955
intention	0.932

All the values of Cronbach coefficient are above 0.8 except for one factor" availability" which has value 0.616. This ensure the questionnaire is reliable.

Statistical treatments

Descriptive statistics along with inferential statistics are considered in this study by using SPSS program.

1. The reliability of the questionnaire was checked by Cronbach's coefficient and factor analysis was used to ensure the validity.
2. Factor analysis summarizes many items, observable variables from the sample to small number of factors. These factors are used to investigate regression and mediation analysis. Bartlett's test for sphericity is used to test a hypothesis that the variables are uncorrelated (the correlation matrix is identity matrix). When the hypothesis is rejected (p-value < 0.05), then factor analysis is meaningful.
3. In factor analysis, unrotated factor loadings are difficult to interpret but rotation simplifies the structure of loadings and hence allows easy interpretation. Several rotation methods are available and, in this study, varimax rotation was adopted which may make it easy to interpret the factors and helped naming them.
4. Regression analysis will be investigated by considering "intention to use simulation" and "improved skills" as dependent variables and the extracted factors from factor analysis are the exploratory variables. The aim of this analysis is to build the best regression model that can explain and predict the dependent variable from the exploratory variables.
5. Mediation analysis explores any mediate variable between the dependent and independent variables. To analyze mediation, we followed Baron & Kenny's (1986) scheme and used Sobel test to test mediation.

Results and Discussion

Evaluation of the program is an essential component of the programme implementation as it facilitates an understanding of the unforeseen or underlying dynamics within a given educational programme, which facilitates a sharper focus on ways to make improvements (Patton 2011). Worthen & Sanders (1987) justified evaluation of any program as it assists the program and then either continue, modify, expand, or terminate it. It also revealed the needs for the programme, its feasibility, review the activity and resources, and hence make improvement.

Academic Performance

Different assessment methods were used to measure the students learning performance in this study. This included exams and assignments. Table 2 illustrates the topics covered by simulation-based learning program and the percentages of students who assimilated the topics from their point of view.

Table 2 The topics covered in the training and the students' rate of comprehension

Topics learned by simulation	Comprehension rate
Obtaining empirical Probabilities	100%
Generating random numbers	100%
Drawing density functions	100%
Constructing standard normal distribution table	100%
Approximating sampling distribution	96.7%
Obtaining probability from distributions	96.7%
Approximating the expectation of random variables	96.7%
Approximating integration	96.7%
Approximating the variance of random variables	96.7%
Empirical illustration of central limit theorem	93.3%
Investigating the role of parameters in the distribution	86.7%
Approximating some popular numbers (π, ϕ, e)	86.7%
Approximating the maximum likelihood estimate	83.3%
Checking the relations among distributions	73.3%
Constructing confidence interval	66.7%
Testing hypotheses	60%
Overall	87.92%

From Table 2, the overall percentage of understanding the topics is 87.92% which is high percentage. Students fully understand the topics: obtaining empirical Probabilities, generating random numbers, drawing density functions and constructing standard normal distribution table. Second group of topics with 96.7% includes approximating sampling distribution, obtaining probability from distributions, approximating the expectation of random variables, approximating integration and approximating the variance of random variables. However, 2 topics got less than 70% and they are constructing confidence interval, hypotheses testing. These topics were the last topics covered and needed more time and effort to understand but students had no sufficient time to study these topics thoroughly.
Figure 1 demonstrates the distribution of the final results of the training program. It shows that more than 42.4% of the students got excellent marks (90 or above), about 25% got very good marks (80 - 89), about 27% got good marks (70-79), 6% got grade pass (60-79) and no failure (less than 60). The overall average is 81.63 out of 100. The results showed very good performance of the students when learning statistics and mathematics based on simulation.

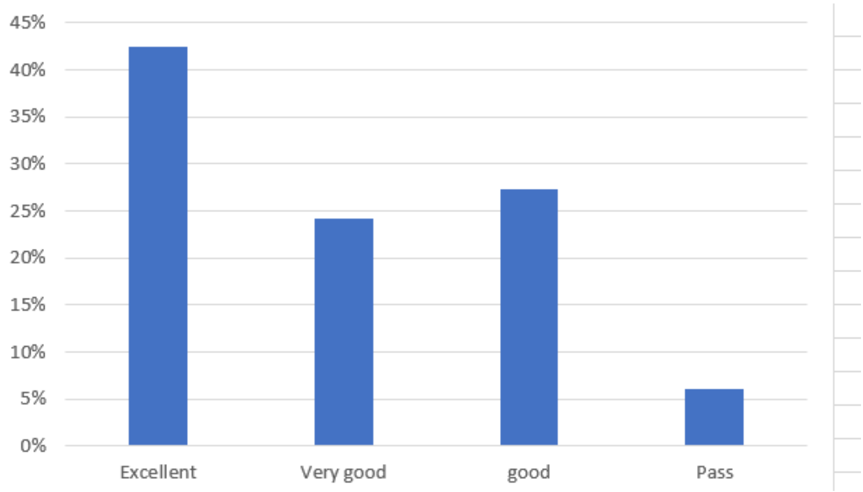

Figure 1: The students' performance in the training

Descriptive Statistics for the Questionnaire

A five-point Likert scale was used to judge the responses of the students' opinion on the training program. The students responded to the items using five-point scale as follows: strongly agree (5), agree (4), Neutral (3), disagree (2), and strongly disagree (1). To interpret the results, the criterion in Table 3 is used

Table 3: The upper and lower bounds of 5-point Likert scale

The average	Degree of consent
1-1.8	Very little
1.8-2.6	Little
2.6-3.4	Medium
3.4-4.2	Large
4.2-5	Very Large

Descriptive statistics including the mean and the standard deviation were calculated and summarized in Table 4. The table demonstrates large rate for most of the items with over all mean 3.713 and standard deviation 1.088. The averages of the items ranged between 4.23 and 2.73. They are satisfied most to the item "The instructor clarifies the programs" and are satisfied least to the item " I did not face problem writing programs."

Table 4: Descriptive statistics about the items of the questionnaire

	Items	mean	Standard deviation	Degree of agreement
1	It solves hard problem (like solving complicated integral)	4.07	0.98	Large
2	I found simulation useful to understand statistics.	3.93	1.23	Large
3	I think simulation is a valuable tool.	3.93	0.91	Large

4	It helps me understand concepts of statistics.	3.90	1.06	Large
5	I found simulation useful to understand mathematics.	3.87	1.07	Large
6	I am aware of the use of simulation.	3.467	1.137	Large
7	I like simulation and I wish to learn more.	3.967	0.999	Large
8	I find it fun.	3.933	1.015	Large
9	I enjoy doing simulation.	3.667	1.061	Large
10	I wish I knew simulation earlier.	3.967	1.033	Large
11	I wish to learn more about simulation	3.867	1.008	Large
12	I wish we had longer class in simulation	3.400	1.221	Large
13	I wish it was taught parallelly with all statistics courses.	3.833	1.234	Large
14	I can develop programs to solve my own problem	3.367	1.066	Large
15	Simulation is user friendly	3.667	0.922	Large
16	It is easy for me to use.	3.500	1.009	Large
17	I am always eager for simulation class	3.933	0.980	Large
18	I did not face problem writing programs.	2.733	1.015	Medium
19	I did not face problem with the computer machine	2.733	1.015	Medium
20	Computer programs are accessible at a satisfactory level of speed.	3.467	1.008	Large
21	I can access the program from the university and home.	3.900	1.185	Large
22	There are adequate facility at the university to enable students to use the program.	3.467	1.106	Large
23	The R program is always available(it is always offline.)	3.133	1.196	Medium

Table 4 (continue): Descriptive statistics about the items of the questionnaire

	Items	mean	Standard deviation	Degree of agreement
24	The instructor clarifies the programs.	4.233	1.104	Very large
25	The instructor clarifies the idea before doing simulation.	4.000	1.114	Large
26	I can build programs to solve problems using simulation.	3.267	1.230	Medium
27	I can think of new cases where we can apply simulation.	3.500	1.137	Large
28	Simulation offers a wide range of opportunities to practice complex skills.	4.000	1.083	Large
29	It helps develop learning by discovery.	4.000	0.910	Large
30	It keeps me active during the class (like using computer.	3.933	1.172	Large
31	Using simulation would improve my performance in learning.	3.900	0.960	Large
32	Simulation keeps me active all the time	3.833	1.147	Large
33	It improves my learning.	3.800	1.064	Large
34	It helps me develop critical thinking (like changing the values of parameters.	3.767	1.194	Large
35	Using simulation enhances my productivity in learning.	3.733	1.081	Large
36	I encourage my friend to take simulation course.	3.933	1.048	Large
37	I intend to solve problems using simulation.	3.833	1.206	Large

38	I need to attend more courses on simulation before I can use it on my own.	3.833	1.147	Large
39	I intend to continue learning more about simulation.	3.700	1.179	Large
40	I intend to search for more applications of simulation.	3.567	1.278	Large
	Overall	**3.713**	**1.088**	**Large**

Factor Analysis of the Independent Variable

Factor Analysis is a method for modeling observed variables in terms of underlying unobservable (latent) factors. In this section, factor analysis was carried out by principal component analysis along with varimax rotation and the results are summarized in Table 5. There are 7 factors extracted with eigenvalues greater than 1. Therefore, the analysis extracted 7 factors which explain 82.991 % of the total variance. The last column provides communalities which present the proportion of the variance that the 7 factors explain for each variable. The bold loadings show which variables have high loadings for each factor. In addition, the last row shows the eigenvalues which present the total variance explained by given factor. In order to name the factors, we put the variables with high loads in the factor together.

Table 5: The extracted factors, loadings, communalities, eigenvalues and variance percentage.

	Loadings							
variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	**communalities**
1	.053	**.809**	.151	.061	.130	.110	.151	**.735**
2	-.024	**.933**	.049	.139	.175	.014	.103	**.934**
3	.202	**.738**	.192	-.136	.011	.404	-.162	**.830**
4	.048	**.888**	.268	.182	-.054	.190	-.007	**.935**
5	.250	**.908**	-.054	.131	.140	-.001	.079	**.934**
6	.379	**.562**	.118	.115	.037	.233	-.133	**.560**
7	**.734**	.044	.405	.322	.137	.203	.084	**.876**
8	**.785**	.200	.281	.130	.331	.071	.087	**.875**
9	**.691**	.352	.269	.190	.210	.106	.098	**.774**
10	**.836**	.160	.014	.015	.114	.039	.121	**.754**
11	**.828**	.113	.255	.353	.091	.039	.146	**.919**
12	**.737**	-.030	.118	.329	-.053	.150	.179	**.724**
13	.543	.095	.242	**.666**	.012	.148	.078	**.833**
14	.331	.289	.328	**.656**	.174	.110	.152	**.796**
15	.340	.188	.039	**.479**	.257	.307	.236	**.598**
16	.370	.214	.421	**.672**	.091	-.140	.036	**.841**
17	.399	.331	**.605**	-.242	.310	.154	-.054	**.816**
18	-.124	-.170	**-.799**	-.204	-.058	-.126	-.408	**.909**
19	-.124	-.170	**-.799**	-.204	-.058	-.126	-.408	**.909**
20	.285	.116	**.842**	.294	.081	.001	-.004	**.896**
21	.372	.064	**.656**	.178	.247	.314	-.251	**.826**
22	.216	.050	.144	.188	.122	.122	**.817**	**.803**

23	.386	.023	.238	-.051	-.442	-.027	**.588**	**.751**
24	.221	.178	.167	.165	**.882**	.142	.024	**.934**
25	.488	.218	.318	.058	**.701**	.031	.038	**.883**
26	.038	.256	.158	.035	.302	**.862**	-.028	**.928**
27	.221	.231	.109	.112	-.096	**.809**	.211	**.834**
eigenvalues	**5.412**	**4.763**	**3.912**	**2.392**	**2.122**	**2.085**	**1.721**	
variance percent	**20.046**	**17.642**	**14.489**	**8.860**	**7.861**	**7.721**	**6.372**	

Table 6: summary of the extracted factors with suggested names.

Factor	Label	High Loading Variables (items)	mean	Standard deviation
1	Enjoyment and attitude	7,8,9,10,11,12	3.800	1.056
2	usefulness	1,2,3,4,5,6	3.861	1.137
3	satisfaction with simulation (satisfaction 1)	17,18,19,20,21	3.353	1.041
4	Ease and ability	13, 14, 15,16	3.592	1.058
5	Satisfaction with instructor (satisfaction 2)	24, 25	4.117	1.109
6	Ability and competence	26, 27	3.383	1.184
7	availability	22, 23	3.300	1.151
overall			**3.629**	**1.105**

Table 6 reveals that averages of the factors are large with overall mean 3.629. The highest average for the " satisfaction 2" with average 4.117 while the lowest is for availability with average 3.3.

Dependent Variables

In addition to the above variables, there are two more variables; "intention to use simulation in learning" and "the skills improved by simulation". The "intention" and "skills" variables will be used as a dependent variable in the regression analysis which will be discussed in the section 9.5.

The Variable "Skills"

The training program aimed to improve students 'skills. Table 7 shows the skills and the rate achieved from the students' point of view. The means of all items are large with overall mean 3.87. The averages ranged between 4 and 3.7.

Table 7 Descriptive statistics of items belong to "skills" variable

	Items	mean	Standard deviation	Degree of agreement
1	Simulation offers a wide range of opportunities to practice complex skills.	4.00	1.08	Large
2	It helps develop learning by discovery.	4.00	0.91	Large
3	It keeps me active during the class (like using computer).	3.93	1.17	Large

4	Using simulation would improve my performance in learning.	3.90	0.96	Large
5	Simulation keeps me active all the time	3.83	1.15	Large
6	It improves my learning.	3.80	1.06	Large
7	It helps me develop critical thinking (like changing the values of parameters).	3.77	1.19	Large
8	Using simulation enhances my productivity in learning.	3.73	1.08	Large
	Overall	**3.87**	**1.075**	**Large**

To apply factor analysis for the items in this section, Barlett's test was significant ($\chi^2 = 249.538, p-value = 0.000$). Kaiser-Meyer-Olkin (KMO) measures sampling adequacy and Factor analysis is appropriate if KMO is at least 0.5. In our data, KMO measure gave value 0.867 indicating meritorious level according to Kaiser and Rice(1974) . The Factor analysis was performed by principal component analysis and the results are given in Table 8. There is only one factor with eigenvalue 6.155. The other factors have eigenvalues less than one. Hence, the analysis extracted only one factor which explains 76.942 % of the total variance. This factor can be put under the name "improvement of skills".

Table 8: loadings, communalities, eigenvalue and the percentage variance for the factor "skills"

	Variables	Loadings	**communalities**
1	Simulation offers a wide range of opportunities to practice complex skills.	**.852**	.726
2	It helps develop learning by discovery.	**.936**	.877
3	It keeps me active during the class (like using computer).	**.929**	.863
4	Using simulation would improve my performance in learning.	**.930**	.865
5	Simulation keeps me active all the time	**.846**	.715
6	It improves my learning.	**.776**	.602
7	It helps me develop critical thinking (like changing the values of parameters).	**.823**	.678
8	Using simulation enhances my productivity in learning.	**.911**	.829
	eigenvalue	6.155	
	Percentage Variance	76.942%	

The Variable "Intention"

Table 9 shows the descriptive statistics of the items belong to the factor "intention". The means ranged between 3.93 and 3.57 with overall mean 3.77 which is of large rate.

Table 9: Descriptive statistics of the items under the name "intention" variable

	items	mean	Standard deviation	Degree of consent
1	I encourage my friend to take simulation course.	3.93	1.05	Large
2	I intend to solve problems using simulation.	3.83	1.21	Large
3	I need to attend more courses on simulation.	3.83	1.15	Large

4	I intend to continue learning more about simulation.	3.70	1.18	Large
5	I intend to search for more applications of simulation.	3.57	1.28	large
	Overall	**3.77**	**1.17**	**large**

Bartlett's test was significant (χ^2 =121.819, $p - value = 0.000$) and the value of KMO value is 0.856 > 0.5 indicating meritorious level. Table 10 illustrates the factor analysis which extracted only one factor with eigenvalue 3.938 >1 while the other factors have eigenvalues less than 1. This factor explains 78.764% of the total variance and is named "intention".

Table 10: loadings, communalities, eigenvalue and the percentage variance for the factor "intention"

		loading	communalities
1	I encourage my friend to take simulation course.	.951	.904
2	I intend to solve problems using simulation.	.920	.847
3	I need to attend more courses on simulation.	.911	.830
4	I intend to continue learning more about simulation.	.829	.687
5	I intend to search for more applications of simulation.	.818	.670
	eigenvalue	3.938	
	Percentage variance	78.764%	

Regression Analysis

The First Model

Let consider "intention" as the dependent variable, and the extracted factors are the independent variables. Using forward method, the following multiple linear regression model is found to be the best significant model among the suggested models.

Intention = (**0.712**) usefulness+ (**0.414**) enjoyment + (**0.275**) satisfaction1+(**0.231**) ease + (**0.186**) ability

The F statistic is 26.714 with p-value =0.000 which indicates significant model. The correlation coefficient between the observed values and predicted values of the dependent variable (intention) is 0.918 and the adjusted coefficient of determination is 0.811. That is 81.1% of the variation in the dependent variable is explained by the model and only 19.9% of the variation is unexplained. Table 11 illustrates the t-test results to test the significance of the regression coefficients. The purpose of t test is to check if each explanatory variable needs to be in the model, given that the others are already there. From the results, all the p-values are less than 0.05 level of significance. Therefore, the proposed model along with all the regression coefficients are significant.

Table 11: The outcomes of t test to test the significance of the regression coefficients

variables	coefficients	t- test	p-value	Correlation		
				Zero-order	Partial	Part
usefulness	0.712	8.968	0.000	0.712	0.873	0.712
enjoyment	0.414	5.215	0.000	0.414	0.722	0.414

Satisfaction 1	0.275	3.463	0.002	0.275	0.569	0.275
ease	0.231	2.912	0.007	0.231	0.503	0.231
ability1	0.186	2.339	0.028	0.186	0.424	0.186

The model is built with standardized variables so that the variables are free of units. The model demonstrates positive coefficients which indicate positive relation. Each regression coefficient in the model tells how much the mean of the dependent variable changes when the independent variable changes by one unit holding other variables constant. Since the regression coefficients are standardized, the most important independent variable is the "usefulness" which has the highest coefficient value. The second important is "enjoyment" which possesses the second highest coefficient and so on. Thus, the variables in the model are in the descending order of importance. In addition, Partial correlation coefficient illustrates the relation between two variables holding the other variables constant. The partial correlation of the variables; usefulness, enjoyment, satisfaction 1, ease and ability are 0.873, 0.722, 0.569, 0.503 and 0.424 respectively.

Diagnostics are used to check whether the assumptions of the model are reasonable. To check the normality of the standardized residual, Kolmogov - Samirnov outcome gave test statistic value 0.153 with p-value 0.071 which indicates that normality of the residuals is not violated. In addition, the extracted factors using varimax rotation have no relation between any two factors as the rotation is orthogonal. In addition, Variance Inflation Factor (VIF) measures of collinearity gave value 1 which is very small. So, there is no collinearity problem in the model. Moreover, to measure how much the model coefficient estimates would change if an observation were to be removed from the data set, Cook's Distance, Mahalanobis and leverage were calculated. There is only one influential point with Cook's distance greater than 1 (Cook's D =1.02353). The same point has leverage value 0.532 which is greater than $\frac{2p}{n}$, p is number of independent variables and n is the sample size. Also, Mahalanobi distance for that point is 15.4286 which is the highest value among all distances. Therefore, only one point may be removed from the data set to get more accurate regression model. However, removing that outlier point from the data set slightly improved the model as its influence was not too much.

The Second Model

Let now consider "skills" to be the dependent variable, and the extracted factors are the independent variables. Using forward method, the following multiple linear regression model is the best significant model among the proposed models.

Skills = (**0.879**) usefulness+ (**0.311**) enjoyment + (**0.204**) ability1 +(**0.141**) satisfy1

The F statistic equals 85.896 with p-value 0.000 which indicates significant model. The correlation coefficient between the observed values and predicted values of the dependent variable (skills) is 0.964 and the adjusted coefficient of determination is 0.919. That is 91.9 % of the variation in the dependent variable is explained by the model. The outcome of the t test given in Table 12 illustrates the coefficients are significant with all p-values are less than 0.05 level of significance.

Table 12: The t test to test the significance of the regression coefficients and the correlations

variables	coefficients	t- test	p-value	Correlation		
				Zero-order	Partial	Part
usefulness	0.879	16.890	0.000	0.879	0.957	0.879

enjoyment	0.311	5.974	0.000	0.311	0.761	0.311
ability1	0.204	3.913	0.001	0.204	0.609	0.204
Satisfaction 1	0.141	2.702	0.012	0.141	0.468	0.141

All the regression coefficients are positive. So, as the independent variables increase, the dependent variable "skills" increases as well. The variables in the model are in the order of importance. That is, among all independent variables, the variable "usefulness" has the highest influence on the dependent variable and "enjoyment" has the second highest important independent variable, then "ability1" followed by" satisfaction1". In addition, Partial correlation coefficient illustrates the relation between two variables holding the other variables constant. The partial correlation of the variables; usefulness, enjoyment, ability 1 and satisfaction 1 are 0.957, 0.761, 0.609 and 0.468 respectively. Moreover, Shapiro-Wilk test of normality gave test statistic value 0.952 with p-value 0.191 which indicates that the residuals are normally distributed. There is no collinearity problem in the model as the Variance Inflation Factor (VIF) measures gave small value, 1. The independent variables are not correlated due to varimax rotation. There is no influential point as all Cook's distance less than 1 (maximum Cook's D = 0.71). So, there is no unusual point.

Mediation Analysis

In some situations, the relationship between the independent and the dependent variable is indirect effect because of the existence of a mediator variable. In regression, when this mediator variable is included in the model, the independent variable is no longer significant while the mediator remains significant. To conclude that a mediator exists or not, one may use Sobel test. This test determines whether the mediation effect is statistically significant. In our study, mediation analysis was carried out for the significantly correlated variables under the assumption that intention is the dependent variable and the independent variables include the extracted factors. Some correlated independent variables were tested for mediation. It is found that the only case of significant mediation is that the "skills" mediated the relationship between "usefulness" and "intention" variables. The results of the simple linear regression show that usefulness is a statistically significant predictor of intention (b =0.831, t =7.910, p-value=0.000). Next, when the mediator, skills, was entered in the regression analysis, usefulness was no longer a significant predictor of intention (b=0.012, t=0.053, p-value=0.958). On the other hand, the mediator, skills, is a significant predictor of intention (b=883, t=3.831, p-value=0.001). Sobel test gave(test statistics=3.684, p-value=0.00023) which is significant. Therefore, "skills" variable mediates the relation between "usefulness" and "intention".

Conclusion

In this study, a training program on learning statistical and mathematical concepts through simulation was implemented to students at University of Nizwa in fall 2023. It turned out that students performed well in the program and showed good attitude toward this scheme. The training program revealed that **s**imulation improves several skills for the student. It improves students' computer skill and programming**.** It also promotes critical thinking, enhances creativity, develops decision making and problem-solving skill. In addition, the students Descriptive Statistics for the Questionnaire

A five-point Likert scale was used to judge the responses of the students' opinion on the training program. The students responded to the items using five-point scale as follows: strongly agree (5), agree (4), Neutral (3), disagree (2), and strongly disagree (1). To interpret the results, the criterion in

Table 3 is used. Students can learn by trial and error and enhances discovery learning which leads to better understanding and long-term retention than does traditional approach. Moreover, students learn best when they actively construct their own knowledge and build their own computer program. The study conducted a questionnaire survey at the end of the training the descriptive statistics demonstrate high positive attitudes and the factor analysis extracted the following latent factors: usefulness, enjoyment, ease, satisfaction and availability. Regression analysis built significant models. Analysis of mediation was also considered to investigate the possible mediate variables between the independent and dependent variable. It is found that "skills acquainted" variable mediates the variables "usefulness" and the "intention to use simulation". The study recommended using simulation in teaching mathematics and statistics. However, when calling for simulation-based classes it does not mean ignoring the analytical solutions to problems. Both analytical and simulated solutions enhance learning. In addition, students should be provided with enough instruction and guide to be active to discover the solutions.

Funding: The research did not receive financial assistance from any funding entity.

Conflicts of Interest: The author has no conflicts of interest to disclose concerning this study.

Declarations: This manuscript has not been published to any other journal or online sources.

Data Availability: The author has all the data employed in this research and is open to sharing it upon reasonable request.

References

Aberson, C.L., Berger, D.E., Healy, M.R., Kyle, D. & Romero, V.L. (2000). Evaluation of an interactive tutorial for teaching the Central Limit Theorem. *Teaching Psychol.*, **27**, 289–291.

Alex Lyford & Michael Czekanski (2020). Teaching students to estimate complex probabilities of board game events through simulation. Teaching StatisticsVolume 42, Issue 3.

Anastasi, A. & Urbina, S. (1997). Psychological testing. Prentice Hall/Pearson Education.

Baron, R. M. & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 5,* 1173-1182.

Braun, W. J., White, B. J. & Craig, G. (2014). R tricks for kids. *Teaching Statistics*, *36*(1), 7-12.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T. & Fischer, F. (2020). Simulation-Based Learning in Higher Education: A Meta-Analysis. *Review of Educational Research*, *90*(4), 499-541.

Dambolena, I. G. (1986). Using simulation in statistics courses. *Collegiate Microcomputer*, *4*(4), 339-44.

Elena Novak (2013). Effects of simulation-based learning on students' statistical factual, conceptual and application knowledge. Journal of Computer Assisted Learning 30(2).

Harrison, RL. (2010). Introduction To Monte Carlo Simulation. AIP Conf Proc. Jan 5, 2010.

Garfield, J., and Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics," *International Statistical Review*, 75, 372–396.

Genevieve Briand & R. Carter Hill(2013). Teaching basic econometric concepts using Monte Carlo simulations in Excel. International Review of Economics Education. Volume 12, Pages 60-79.

Gordon, Florence S.; Gordon & Sheldon P.(1989). Computer Graphics Simulations of Sampling Distributions. Collegiate Microcomputer, v7 n2 p185-89.

Jorge Antonio Zazueta-Hernández & José Daniel López-Barrientos (2021) From Chutes and Ladders to No Te Enojes: Simulation of two (moral) random paths. Teaching Statistics, Volume 43, Issue 3.

Knekta E, Runyon C & Eddy S. (2019). One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. CBE Life Sci Educ. 2019 Mar;18(1).

Koparan, Timur; Yilmaz & Gül Kaleli(2015). The Effect of Simulation-Based Learning on Prospective Teachers' Inference Skills in Teaching Probability. Universal Journal of Educational Research, v3 n11 p775-786 .

Mikuláš Gangur & Milan Svoboda (2018). Simulation of Bayes' rule by means of Monte Carlo method. Teaching Statistics, Volume 40, Issue 3.

Mills, J. D. (2002). Using computer simulation methods to teach statistics: A review of the literature. Journal of Statistics Education, 10(1).

Patton M. (2011). Developmental evaluation: Applying complexity concepts to enhance innovation and use. New York: Guilford Press.

Reidar Hagtvedt, Gregory Todd Jones & Kari Jones (2007). Pedagogical Simulation of Sampling Distributions and the Central Limit Theorem. Teaching Statistics
Volume 29, Issue 3 p. 94-97.

Ronald W. W. Shonkwiler and Franklin Mendivil (2009). Explorations in Monte Carlo Methods (Undergraduate Texts in Mathematics).Springer.

Rui Manuel & Costa Martins (2018). Learning the principles of simulation using the birthday problem. Teaching Statistics, Volume 40, Issue 3.

Sigal, M. J., & Chalmers, R. P. (2016). Play It Again: Teaching Statistics With Monte Carlo Simulation. Journal of Statistics Education, 24(3), 136–156.

Stephen Turner, Alan R. Dabney (2014). A Story-based Simulation for Teaching Sampling Distributions. Teaching Statistics, Volume 37, Issue 1.

Ted Hodgson and Maurice Burke (2001). On Simulation and the Teaching of Statistics
Teaching Statistics, Volume 22, Issue 3

Thiesmeier, R., Orsini, N., Gracely, E., & Oster, B. (2024). Teaching Statistics in Health Sciences: The Potential of Simulations in Public Health. CHANCE, 37(2), 34–39.

Batanero, C., Álvarez-Arroyo, R.(2024) Teaching and learning of probability. *ZDM Mathematics Education* **56**, 5–17.

Tim P. Morris, Ian R. White & Michael J. Crowther (2019). Using simulation studies to evaluate statistical methods. Statistics in Medicine. Volume 38, Issue 11 p. 2074-2102.

Timur Koparan (2019). Teaching Game and Simulation Based Probability. International Journal of Assessment Tools in EducationVolume: 6 Issue: 2, 235 – 258.

Webster, Elaine (1992). Evaluation of Computer Software for Teaching Statistics. Journal of Computers in Mathematics and Science Teaching, v11 n3-4 p377-91.

West, R. W., & Ogden, R. T. (1998). Interactive demonstrations for statistics education on the World

Wide Web.Journal of Statistics Education, 6(3).
Worthen, B.R. & Sanders, J.R. (1987) Educational Evaluation: Alternative Approaches and Practical Guidelines. Longman Press, New York, 102.