Full text open access online (Since 2009)

PRINT: ISSN 0975-1122 ONLINE: ISSN 2456-6322
Int J Edu Sci, 54(2): 249-257 (2026)
DOI: 10.31901/24566322.2026/54.2.1486

A Comparative Study of Traditional and Flipped Classroom on Students' Statistical Reasoning

Anton A. Romero

Central Bicol State University of Agriculture, San Jose Pili, Camarines Sur, Philippines
Orcid: https://orcid.org/0000-0003-4765-2412

KEYWORDS Analysing and Interpreting. Classroom Shortage. Findings and Conclusions. Mastery Learning Theory. Statistical Software

ABSTRACT The intent of this experiment is to investigate how flipped classrooms affect students' statistical reasoning. This study comprised 109 students, wherein the experimental group comprised 51 while 58 students were part of control. The Statistical Reasoning instrument was used to acquire the scores, and the Mann-Whitney and Wilcoxon Signed Rank Test were utilised for analysis. The findings revealed that both groups scores were comparable and both have significantly improved and numerous students achieved analytical levels of reasoning. It was concluded that the flipped classroom is comparable to the traditional method and that improvement depends primarily on the feedback mechanism rather than in-person lectures. Further, this study suggests implementing flipped classrooms in basic education and other undergraduate courses to address the prevailing problem of classroom shortage. To produce more rigorous results, the researchers recommend to reassess the research instrument, improve sample selection techniques, and employ a mixed method approach to further improve the flipped classroom implementation.

INTRODUCTION

Deficiency of classrooms is one of the leading problems facing Philippine government schools. At the beginning of the 2023-2024 school year, the Department of Education (DepEd) publicised that there was a scarcity of 159,000 lecture rooms (Macasero 2023), which is one of the reasons why only 4 out of 1000 Filipino students who graduated from K-12 achieved proficiency level (Hernando-Malipot 2026). The educational system in accordance with the theory of transactional distance will continue to face this problem if students are incapable to embrace independence and responsibility for their own learning (Abuhassna and Yahaya 2018). Hence, this issue also affects the country's State Universities and Colleges (SUCs), and it is likely that it will not get fixed immediately given the Commission in Higher Education's (CHED) 2024 reductions in funding, which cause negative impact on SUCs' capacity to deliver high-quality instruction (Tucay 2023). Since it is unrealistic to build enough buildings to accommodate the increasing number of students enrolled in SUCs, alternative solutions must be considered as a way to at least minimise the problem and improve the quality of instruction.

Flipped classroom is a kind of blended learning and an approach aligned to education 4.0 (Niromand 2023; Páez-Quinde et al. 2023) and employed self-regulated learning (Gumartifa et al. 2026). Furthermore, Ilgaz (2022) found that flipped classrooms are superior to traditional methods among nursing students in their Biostatistics class, while Chien (2018) concluded that students learn better in flipped classrooms than other approaches. Despite numerous studies mentioning the efficacy of flipped classrooms compared to traditional methods, there are few inquiries on its effect on students' statistical reasoning in Biostatistics, which requires them to draw correct findings and contextualise conclusions. Consequently, it is essential to look into how one of the key components of statistical reasoning, that is, analysis and interpretation, is influenced by the intervention.

Investigating the effectiveness of flipped classrooms in the perspective of public universities could help the University officials minimise if not address its perennial issues. Likewise it could address the problem of classroom scarcity as well as improve the learning processes that are crucial to achieve the learning objectives of the course. In particular, each classroom is capable of accommodating more faculty than a traditional one without building more classrooms if limited face-to-face classes employing flipped classrooms are proved to be effective or similar to full face-to-face classes. Finally, it could possibly be beneficial for the university to come up with a policy that restricts in-person instruction to certain courses.

Objectives of the Study

This paper aims to investigate whether the strategy of flipped classrooms is better or comparable to the traditional approach in terms of the statistical reasoning (analysis and interpretation skills) in the biostatistics course. Specifically, it answered the following research questions:

1. What is the level of statistical reasoning of the control and experimental groups?
2. Is there a significant difference between the pretest and posttest statistical reasoning performance of the control and experimental groups?
3. Is there a significant difference in statistical reasoning performance between the control and experimental groups in both the pretest and posttest?

Literature Review

The study's variables flipped classrooms and analysis and interpretation skills (component of statistical reasoning) are discussed in this section. This paper's theoretical foundation, conceptual framework, and hypotheses were discussed as well.

Statistical Reasoning

Students' responses can be categorised as idiosyncratic, transitional, quantitative, and analytical level (Mooney 2002; Jones et al. 2004; Ramadhani et al. 2022). According to Chan et al. (2016), students frequently struggled to analyse and interpret processes, failed to reach the analytical level of reasoning, and were failed to draw accurate findings and contextualized conclusions (Maryati and Priatna 2018; Hasim et al. 2024). Subsequently, it is essential to focus on enabling students improve their analytical and interpretation skills since these are important to generate correct findings and contextualized conclusions.

Flipped Classroom and Statistical Reasoning

The flipped classroom is used as an intervention in biostatistics courses, and has been the subject of numerous investigations in the literature. In comparison with alternative approaches to learning, flipped classrooms are more effective (Peterson 2016; Chien 2018; Farmus et al. 2020; Ilgaz 2022) and also used in other subjects (Yalew 2025). The rate of success had been greater in the year that flipped classes were implemented than in the previous year when they were not (Yañez et al. 2023). In contrast, compared to the Flipped Classroom, students in traditional classes achieved noticeably higher marks (Binsomprasong et al. 2019). The relative usefulness of flipping the class in comparison to other methods of instruction could not be assessed according to Mclaughlin and Kang (2017). Before and after the flipped classroom was implemented in 2012 and 2013, other studies found no difference (McGraw and Chandler 2015; Gillette et al. 2018). The way practitioners use statistical concepts and interpret the results of statistical analysis is known as statistical reasoning (Garfield 2002; Garfield and Ben-Zvi 2009; Zapata-Cardona 2023). In light of this, a significant number of studies do not specifically assess statistical achievement using statistical reasoning.

Conceptual and Theoretical Framework

Figure 1 shows the conceptual and theoretical paradigm of this paper. According to Winget and Persky (2022), mastery learning has a solid theoretical foundation for its edge over traditional methods and is closely associated with modern approaches like flipped classroom, which both leverage feedback to improve learning. This paper is grounded on Benjamin Bloom's Mastery Learning Theory, which is further elaborated upon by Akpan (2020) with the proposition that more lecture time might not result in successful achievement of the learning objective due to the lack of appropriate feedback mechanisms. For a classroom to be regarded as flipped learning, it must conform to the four pillars of flipped learning, particularly calling for instructors to provide students with immediate feedback (Network 2014).

The theory offers a useful theoretical framework for investigating if additional lecture time really enhances students' achievements. In context, the traditional as well as the flipped classroom groups were provided with exact same teaching resources and instruction and feedback from the same instructor. While the lectures in the flipped classroom were asynchronous through the module and video lessons, and face to face lectures were employed in the traditional method. Therefore, the theory's concepts support the proposi-

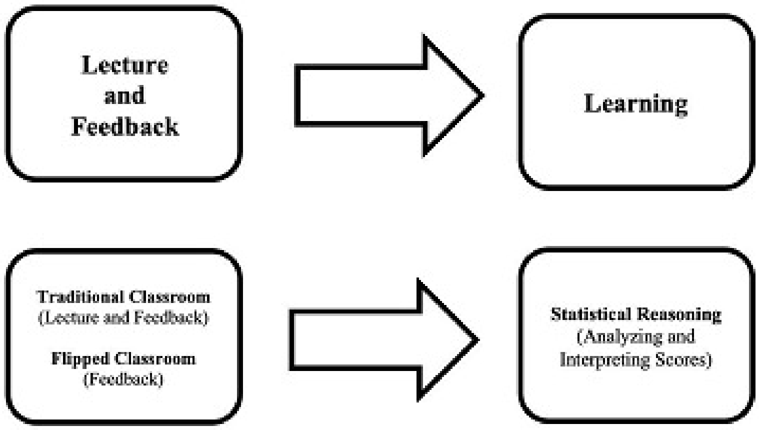

Fig. 1. Conceptual and theoretical paradigm

tion that the feedback mechanism in flipped classrooms, even in the absence of lectures delivered in person could enhance students' analysis and interpretation skills and may be comparable to classrooms that provide lectures in-person.

Hypotheses

Two hypotheses can be made, namely,

1. The scores before and after the traditional and flipped classroom intervention will significantly improve because both interventions have feedback loops.
2. The learning gains of the experimental group are comparable to those of the control group, because according to the theory, more lecture time does not always improve learning.

METHODOLOGY

The research design, subjects of the study, research instrument, data analysis, research process and ethics are discussed in this section.

Research Design

This paper employed a nonequivalent quasi-experimental control group design which is adopted from the book published by Creswell and Creswell (2017). With this approach, the researcher has no or little influence over the four sections' random assignment to the variable under control. The experimental group in this paper used the flipped classroom as intervention, whereas traditional classrooms served as the control group.

Subjects of the Study

This paper has 109 subjects, with 51 for the experimental and 58 for the control group from one of the State Universities in the Philippines. The subjects were undergraduate students enrolled in the Biostatistics course from the four intact classes. Out of the four sections, two were chosen at random to serve as the experimental group, while the remaining served as the group under control. This kind of sampling technique is typically applied in quasi-experimental designs (Leavy 2017), where both groups' pretests for statistical reasoning were comparable at the beginning of the experiment. She emphasised that matching could be employed to create comparable pretest scores, with the control group allotted one pair and assigned the other to the experimental group.

Sampling

One of the two universities in the Philippines' Bicol Region that offered biostatistics classes for veterinary medicine constituted the population for this paper, which included all of the biostatistics students in one of the Universities. Convenience sampling was applied in the second phase in order to guarantee that the two groups were matched prior to the intervention's execution. Cluster sampling has two stages, according to Leavy (2017). Choosing from the identified cluster or incorporating all of the population after the initial step of selection from existing clusters (University). G^* power must be used in the power analysis, as suggested by Creswell and Creswell (2017). In the research they conducted on the effects of flipped classrooms on statistical reasoning, Ramadhani and Evans (2022) found an effect size of -.978. With just 14 and 28 samples needed for each group, respectively in the succeeding study, the actual computed powers using the G^* Power software was .81. That is why 51 samples for the experimental and 58 samples for the control were sufficient for the purpose of this research.

Research Instrument

The present study employed a teacher-made test for statistical reasoning assessment, including only analysis and interpretation skills. There are 6 open-ended questions on the parametric test topic. For assessing the students' responses, the instructor employed a four-point rubric with one point for idiosyncratic reasoning, two points for transitional reasoning, three points for quantitative reasoning, and four points for analytical reasoning. The reliability of the instrument was tested using the Kendall's Tau and percent agreement. Interrater results showed significant agreement between the two experts, evidence in the percent agreement (90% and 100% during the first and second validations) while intrarater results were 80 percent and 90 percent for the 1st and 2nd validator, respectively. All the results using the Kendall's Tau-b showed significant agreement between the two experts for interrater (b=.741 for 1st validation and b=.940 for 2nd validation) and intrarater (rater 1 with t_b=.810 and rater 2 with t_b=.800). To determine the validity of the instrument, another expert evaluated the test of the students using their own criteria, the results showed using spearman rank correlation that the four validation scores from the first two experts has significant relationship with the score provided by the third expert with r-values (.735, .929, .742 and .936) and the p-values were less than .05.

Data Analysis

The quantitative data was analysed utilising JAMOVI version 2.7, and employed the Mann-Whitney U to compare the two groups, while Wilcoxon-Signed Rank tests were applied to examine the within groups. The nonparametric test's effect size was computed using the value of (r). In particular, rank-biserial and Glass rank-biserial correlation for within and between variations, respectively (Tomczak and Tomczak 2014). Sawilowsky (2009) suggested and employed the rule of thumb to determine the effect size and its conversion (Ellis 2010) as shown in Figure 2 and Table 1.

$$d=\frac{2r}{\sqrt{1-r^2}}$$

Fig. 2. Conversion formulas

Table 1: Effect size interpretation

d-value	*Effect size*
2.00	Huge
1.20	Large
.50	Medium
.20	Small
.01	Very small

Students' answers to the six statistical reasoning items were scored on a 1-4 scale. Total scores (6-24) were categorised based on frameworks used in statistics education research, where students' reasoning progresses from idiosyncratic interpretation of data toward analytical understanding of data relationships and interpretation (Jones et al. 2004). Specifically, this study applies logical categorisation to identify the students' level of statistical reasoning by dividing the range by the number of items (6), obtaining 4 to 5 points per category, as shown in Table 2.

Table 2: Statistical reasoning interpretation

Score	*Interpretation*
1 – 9	Idiosyncratic
10 - 14	Transitional
15 – 19	Quantitative
20 - 24	Analytical

Research Procedure and Ethics

This research was implemented in the 1st semester of the academic year 2025-2026. The researcher's first step is to submit a research proposal to the research division of the university. The pretest was administered using a paper and pencil after the participants received a consent letter. In two intact classes, the researcher used the flipped classroom approach, while in two more intact classes, the traditional method was used. The exact same method was used for a posttest of the SRA at the final week of the semester, and the participants' identities confidentiality is ensured. Finally, the findings were presented to the university's research division and invited experts for improvement purposes. The flipped classroom model was executed and served as the intervention to improve students' statistical reasoning (Romero and Angeles 2021; 2023).

RESULTS

The statistical analysis between the two groups, between their performances before and after the interventions and the statistical reasoning of the participants is presented in this section.

Students Levels of Statistical Reasoning

Table 3 reveals the level of reasoning during the students' tests under the group experimental. During the pretest, students achieved 74.5 percent (idiosyncratic), 25.5 percent (transitional) and no one achieved quantitative and analytical reasoning. The percentage of reasoning was 5.9 percent (idiosyncratic), 19.6 percent (transitional), 47.05 percent (quantitative) and 27.45 percent (analytical reasoning). It shows that only 25.5 percent remains in idiosyncratic and transitional levels of reasoning and more than 74 percent achieved quantitative and analytical reasoning.

The control groups' type of reasoning is shown in Table 4. During the pretest, students achieved 72.41 percent (idiosyncratic), 27.59 percent (transitional) and no one achieved quantitative and analytical reasoning. The posttest results show students achieved 1.72 percent (idiosyncratic), 29.31 percent (transitional), 43.10 percent (quantitative) and 25.86 percent (analytical reasoning). It shows that only 31 percent remains in idiosyncratic and transitional levels of reasoning and more than 69 percent achieved quantitative and analytical reasoning.

Pretest and Posttest Comparison

Pretests were administered to the four sections enrolled in Biostatistics during the first week of the semester, and a posttest was conducted during the eighteenth week. As shown in Table 4, the experimental group's posttest mean and median scores were 16.90 and 17.0 respectively, while the pretest mean score was 7.49 and the median was 6.00 under the control group. For the control group, the Z-value was -4.874 while -4.549 for the experimental group and both were significant at .001 with r-values of -1.00. It reveals that the control and experimental group's pretest and posttest results differ significantly, with a large effect size.

Table 3: Experimental group level of reasoning

Level of reasoning	*Pretest*		*Posttest*	
	Frequency	*Percentage*	*Frequency*	*Percentage*
Idiosyncratic	38	74.5	3	5.90
Transitional	13	25.5	10	19.60
Quantitative	0	0	24	47.05
Analytical	0	0	14	27.45
Total	51	100	51	100

Table 4: Pretest and posttest statistical reasoning assessment: Wilcoxon Signed Ranks test

Test/Group	*N*	*Mean*	*Median*	*Z*	*r(ES)*	*Interpretation*
Control						
Pretest	58	7.49	6.00	-4.874***	-1.00	Huge
Posttest	58	16.90	17.00			
Experimental						
Pretest	51	8.10	7.00	-4.549***	-1.00	Huge
Posttest	51	16.70	17.0			

Note: *p<.05, **p<.01, ***p<.001

Control and Experimental Comparison

The pretest data of the two groups were analysed during the second week of the semester. Based on Table 5, the experimental group's mean score was 7.49 with median of 6.00 while the control group mean was 8.10 with median of 7.00. Although the control group's mean and median scores are slightly higher than those of the experimental, the M-W-U test value of 1302 is not statistically significant. It suggests that the pretest scores of the two groups did not differ significantly, with an effect size interpreted as small. Additionally, the experimental group's mean and median posttest scores were 16.9 and 17.0 respectively, whereas the control groups were 16.7 and 17.0. It demonstrates that the experimental group's mean score is a little greater than the other group with a Mann Whitney U value of 1415 and a very small effect size, hence, the difference is not significant in the posttest of the two groups.

DISCUSSION

The current study investigated the effects of traditional and flipped classroom instructional approaches on veterinary medicine students' statistical reasoning in a biostatistics course. Both instructional approaches significantly enhanced students' ability to analyse and interpret statistical results, with no statistically significant difference between the groups. These findings imply that the quality of instructional design has a greater effect on statistical reasoning than the instructional format. Higher-order statistical reasoning can be successfully reinforced in both traditional and flipped classrooms when educational environments promote conceptual understanding, active engagement, and timely constructive feedback. The similar priority given to analyzing statistical results generated by statistical software in contrast to carrying out manual calculations could be an explanation for the comparable improvements. This approach is in line with modern ideas in statistics education, which put more emphasis on evidence-based decision-making, reasoning, and interpretation than on procedural calculation. Statistical software possibly allowed students to focus on analyzing data and making insightful findings by reducing the cognitive demands of computation by hand. In the same way, Alexander et al. (2026) argued that advancements in statistics education generally favor technologically aided educational settings that promote conceptual understanding and real-world data analysis. According to a 2025 systematic review of active learning in biostatistics instruction, students' conceptual understanding and engagement are improved by technology-enhanced and structured learning experiences (Quinapanta Castro et al. 2026).

The absence of a significant advantage for flipped learning contrasts with studies reporting superior outcomes for flipped classrooms (Peterson 2016; Chien 2018; Farmus et al. 2020; Ilgaz 2022; Yañez et al. 2023). However, this difference may reflect variations in instructional design rather than the effectiveness of the flipped model itself. Wu and Cheng (2026) and Kathiravan et al. (2025) mentioned that flipped classrooms generally enhance knowledge acquisition and self-directed learning in medical education, whereas Andrade-Castellanos and Yanowsky-Escatell (2026) reported substantial variation across studies due to differences in implementation. In addition, Czarnowske et al. (2025) and Spaic et al. (2025) concluded that organized learning tasks and adequate instructional guidance are the most effective ways to facilitate flipped learning. The similar benefits seen are consistent with evidence that pedagogical quality is more significant than instructional strategy alone

Table 5: Control and experimental statistical reasoning assessment: Mann-Whitney U test

	Experimental		*Control*			
SRAT	*N*	*Mean*	*Median*	*Z*	*r(ES)*	*Interpretation*
Interpretation						
Pretest	7.49	6.00	8.10	7.00	1302	.135
Small						
Posttest	16.9	17.0	16.7	17.0	1415	.060
Very small						

Note: *p<.05, **p<.01, ***p<.001

because both instructional groups in the present study included opportunities for feedback, statistical software, and guided learning. Because both instructional groups in the present study incorporated guided learning, statistical software, and opportunities for feedback, the comparable improvements observed are consistent with evidence that pedagogical quality is more important than instructional strategy alone. One noteworthy finding was that significant increases in lecture time in the traditional classroom did not result in substantial gains in statistical reasoning. Despite having almost twice as much face-to-face instruction, students in traditional classroom achieved learning outcomes comparable to those in the flipped classroom. This result suggests that instructional efficiency depends more on quality learning experiences than on the quantity of lecture time. In veterinary education, where complex educational programs demand efficient use of classroom time, it also has practical implications. Therefore, well-designed flipped classes may offer flexibility to give more time to practical learning without sacrificing students' ability to reason statistically.

These findings align with Bloom's Mastery Learning Theory, which puts more emphasis on formative assessment, systematic instruction, and corrective feedback than just instructional time (Akpan 2020). In a comparable manner, Hattie and Timperley (2007) discovered that student learning is significantly affected by feedback. By showing that well-designed learning spaces can result in comparable gains in statistical reasoning irrespective of whether instruction is given through traditional or flipped classrooms, the present research extends this findings to veterinary biostatistics. Future research must look at specific elements of an instructional design that optimally support statistical reasoning in health professions education.

CONCLUSION

The findings indicate that students' statistical reasoning in terms of interpretation and analysis dimension has been improved in both traditional and flipped classrooms, notably through the use of statistical software. On the other hand, it implied that the feedback mechanism used in both groups enhanced learning rather than the extended lecture time in the traditional classroom. This is aligned with Bloom's Mastery Learning Theory, which emphasises that learning support and informative feedback, rather than lengthy lectures, are the key to mastery. Additionally, carrying out the flipped classroom model, which is based on the four pillars of flipped learning, aided provide avenues for active learning and continuous feedback. These results thus show that in order to ensure mastery of statistical reasoning skills, effective learning in biostatistics necessitates an effective feedback mechanism and an interactive learning environment.

RECOMMENDATIONS

Despite the comparable effectiveness of both methods, it is suggested to adopt flipped classrooms, as it lessens student-teacher interaction needed and addresses the University's problem on classroom scarcity. Also, the strategy is designed to focus on continuous feedback and is highly student centred. In addition, it is highly encouraged to use statistical software when the goal is to develop the interpretation and analysis skills of the students wherein they can simulate data and focus on interpreting statistical results rather than the computation aspect. Lastly, it is highly recommended to use the flipped classroom in other undergraduate courses, ensuring that they comply with the flipped learning pillars in order to avoid chaos and confusion during the implementation stage.

LIMITATIONS AND FUTURE STUDIES

This investigation has some of limitations, such as the sole emphasis on quantitative databases, the absence of random sampling in the study's subject selection method, descriptor used was based on the logical perspective of the researcher, failure to evaluate the manipulation check measure of variable related to statistical reasoning and the other three components were excluded. Despite the fact that there was evidence that the research instrument was valid and reliable, additional testing must be conducted in order to ensure the preciseness of the data gathered. In addition, the development of the remaining constructs of statistical reasoning must be done in the Biostatistics Course for holistic development of the students in the course. Further, the four pillars of flipped learning must be the basis of a mixed method experimental core design to be able to guarantee that flipped learning takes place in the flipped classroom. Additionally, stratified and random sampling must be used, along with a manipulation check measure based on statistical reasoning to improve the results' generalisability. To maximise the usefulness of flipped classrooms in dealing with an issue of classroom scarcity, further investigation in basic education and undergraduate courses needs to be conducted.

CONFLICT OF INTEREST AND ACKNOWLEDGEMENTS

The author affirms that there is no conflict of interest concerning the publication of this study and expresses sincere thanks to the University's Research Division who provided full financial and time support, to the participants who gave their consent to the use of their data, to experts who validated the flipped classroom module and research instrument, and to God for providing the author with the motivation and critical thinking necessary to carry out this study.

REFERENCES

Abuhassna H, Yahaya N 2018. Students' utilization of distance learning through an interventional online module based on Moore transactional distance theory. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(7): 3043-3052. https://doi.org/10.29333/ejmste/91606

Alexander C, Gaskell J, Davies V 2026. Reflections on the future of statistics education in a technological era. *arXiv preprint arXiv:* 2602.18242. https://doi.org/10.48550/arXiv.2602.18242

Andrade-Castellanos CA, Yanowsky-Escatell FG 2026. Impacto del aula invertida en la educación médica: Una revisión sistemática con metaanálisis [Impact of the Flipped Classroom on Medical Education: A systematic review with meta-analysis]. *Revista medica de Chile*, 154(1): 1-12. https://doi.org/10.4067/s0034-98872026000100001

Akpan B 2020. Mastery learning-Benjamin bloom. In: B Akpan (Ed.): *Science Education In Theory And Practice: An Introductory Guide To Learning Theory.* Cham: Springer International Publishing, pp. 149-162. https://doi.org/10.1007/978-3-030-43620-9_11

Binsomprasong OA, Tanasugarn L, Buraphat P, Homsanit M 2019. Comparison of academic outcomes of epidemiology and biostatistics course between traditional lecture and flipped classroom blended approach in 2nd year medical students: A retrospective cohort study. *Siriraj Medical Journal*, 71(2): 123-126. https://doi.org/10.33192/smj.2019.19

Chan SW, Ismail Z, Sumintono B 2016. A framework for assessing high school students' statistical reasoning. *PLOS ONE*, 11(11): e0163846. https://doi.org/10.1371/journal. pone.0163846

Chien LY 2018. Effects of a flipped classroom for nursing students studying biostatistics: Changes in depth of learning. *Teaching Practice and Innovation*, 1(1): 153.

Creswell JW, Creswell JD 2017. *Research Design: Qualitative, Quantitative, And Mixed Methods Approaches*. Sage Publications.

Czarnowske D, Heiss F, Schmitz T, Stammann A 2025. The effects of flipped classrooms in higher education: A causal m effects of flipped classrooms in higher education: A causal machine learning analysis. *arXiv preprint arXiv:* 2507.10140. https://doi.org/10.48550/arXiv.2507.10140

Ellis PD 2010. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results.* Cambridge, United Kingdom: The Cambridge University Press. https://doi.org/10.1017/cbo9780511761676

Farmus L, Cribbie RA, Rotondi MA 2020. The flipped classroom in introductory statistics: Early evidence from a systematic review and meta-analysis. *Journal of Statistics Education*, 28(3): 316-325. https://doi.org/10.1080/10691898.2020.1834475

Garfield J 2002. The challenge of developing statistical reasoning. *Journal of Statistics Education,* 10(3): 1-12. https://doi.org/10.1080/10691898.2002.11910676

Garfield J, Ben-Zvi D 2009. Developing Students' Statistical Reasoning: Connecting Research And Teaching Practice. *Springer. https://doi.org/10.1007/s11858-009-0176-6*

Gillette C, Rudolph M, Kimble C, Rockich-Winston N, Smith L, Broedel-Zaugg K 2018. A meta-analysis of outcomes comparing flipped classroom and lecture. *American Journal Of Pharmaceutical Education*, 82(5): 6898. https://doi.org/10.5688/ajpe6898

Gumartifa A, Sofendi S, Inderawati R, Vianty M 2026. Designing and evaluating an ethnopedagogy-based-flipped SRL framework for developing EAP reading materials for management students. International J*ournal of Learning, Teaching and Educational Research,* 25(2): 306-329. https://doi.org/10.26803/ijlter.25.2.15

Hasim SM, Rosli R, Halim L 2024. A quantitative case study of secondary school students' level of statistical thinking. *Eurasia Journal of Mathematics, Science and Technology.* https://doi.org/10.29333/ejmste/14358

Hattie J, Timperley H 2007. The power of feedback. *Review of Educational Research*, 77(1): 81-112. https://doi.org/10.3102/003465430298487

Hernando-Malipot M 2026. Filipino students' proficiency plummets to near zero by Senior High School, study finds. *Manila Bulletin*. https://doi.org/10.29333/ejmste/14358

Ilgaz A 2022. Flipped classroom with interactive videos applied to nursing students in the teaching of biostatistics knowledge: A mixed-method study. *Journal of Nursology*, 25(4): 245-253. https://doi.org/10.5152/janhs.2022.2281132

Jones GA, Langrall CW, Mooney ES, Thornton CA 2004. Models of development in statistical reasoning. In: Dani Ben-Zvi, Joan Garfield (Eds.): *The Challenge Of Developing Statistical Literacy, Reasoning And Thinking.* Dordrecht, Netherlands: Springer, pp. 97-117. https://doi.org/10.1007/1-4020 2278-6_5

Kathiravan G, Shilpa Shree J, Sudeep Kumar NK, Thirumavalavan R 2025. Student-centred pedagogical design in veterinary education: A conjoint approach. *International Journal of Educational Sciences,* 51(2): 32-41. DOI: 10.31901/24566322.2024/51-02-1412

Leavy P 2017. *Research Design: Quantitative, Qualitative, Mixed Methods, Arts-Based, and Community Based Participatory Research Approaches.* New York, United States: The Guilford Press.

Macasero R 2023. Philippine Classroom Shortage Rises to 159,000–DepEd. RAPPLER. From <https://www.rappler.com/nation/deped-report-classroom-shortage-school-year-2023-2024/> (Retrieved on 25 January 2024).

Maryati I, Priatna N 2018. Analysis of statistical misconception in terms of statistical reasoning. In: *Journal of Physics: Conference Series*, 1013(1): 012206. 6596/1013/1/012206. https://doi.org/10.1088/174

McGraw JB, Chandler JL 2015. Flipping the biostatistics classroom, with a twist. *Bulletin of the Ecological Society of America*, 96(2): 375-384. https://doi.org/10.1890/0012-9623-96.2.375

Mclaughlin JE, Kang I 2017. A flipped classroom model for a biostatistics short course. *Statistics Education Research Journal*, 16(2): 441-453. https://doi.org/10.52041/serj.v16i2.200

Mooney ES 2002. A framework for characterizing middle school students' statistical thinking. *Mathematical Thinking.* https://doi.org/10.1207/s15327833mtl0401_2

Network FL 2014. The Four Pillars of FLIP. Florida Leadership Network. From <http://www.fliplearning.org> (Retrieved on 24 May 2024).

Niromand F 2023. Evolution of education from web 1.0 to 4.0. *Academic Research In Educational Sciences,* 4(11): 127-133.

Páez-Quinde C, Hernández-Rosales MJ, Robles-Ortega DA, Mendoza-Chavarria V 2023. Flipped Classroom Methodology As A Teaching Strategy In Education 4.0. In: *2023 IEEE Global Engineering Education Conference (EDUCON).* IEEE, Lisbon, Portugal, 27-30 June 2023, pp. 1-5. https://doi.org/10.1109/educon54358.2023.10125150

Peterson DJ 2016. The flipped classroom improves student achievement and course satisfaction in a statistics course: A quasi-experimental study. *Teaching of Psychology*, 43(1): 10-15. https://doi.org/10.1177/0098628315620063

Ramadhani R, Evans B 2022. Measuring studentsâ•™ statistical reasoning abilities using flipped classroom model with SPSS and STATCAL. *Journal of Honai Math*, 5(1): 1-14. https://doi.org/10.30862/jhm.v5i1.246

Ramadhani R, Saragih S, Napitupulu EE 2022. Exploration of students' statistical reasoning ability in the context of Ethnomathematics: A study of the Rasch Model. *Mathematics Teaching Research Journal,* 14(1): 138-168. https://doi.org/10.31764/jtam.v6i3.8375

Quinapanta Castro NI, Choez-A JF, Orbea AF 2026. Active learning in biostatistics education for medical students: A systematic review and meta-analysis. *Cureus*, 18(3): e105637. https://doi.org/10.7759/cureus.105637

Romero AA, Angeles ED 2021. Exploring the pros and cons of implementing flipped classroom in a virtual learning environment in Mathematics 7. *International Journal of Humanities and Social Sciences*, 13(2): 61-75.

Romero AA, Angeles ED 2023. Flipped classroom in a digital learning space: Its effect on the students' attitude toward mathematics. *International Journal of Learning, Teaching and Educational Research*, 22(1): 210-227. https://doi.org/10.26803/ijlter.22.1.12

Sawilowsky SS 2009. New effect size rules of thumb. *Journal Of Modern Applied Statistical Methods,* 8(2): 26. https://doi.org/10.22237/jmasm/1257035100

Spaic D, Bukumiric Z, Rajovic N, Markovic K, Savic M, Milin-Lazovic J, Grubor N, Milic N, Stanisavljevic D, Despotovic A, Bokonjic D, Vladicic Masic J, Janicijevic V, Masic S, Milic N 2025. The Flipped Classroom in Medical Education: Systematic review and meta-analysis. *J Med Internet Res*, 27: e60757. https://doi.org/10.2196/60757

Tomczak M, Tomczak E 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences,* 1(21): 19-25.

Tucay M 2023. *30 SUCs to Suffer Budget Cuts in 2024.* Manila, Philippines: AlterMidya.

Winget M, Persky AM 2022. A practical review of mastery learning. *American Journal of Pharmaceutical Education*, 86(10). https://doi.org/10.5688/ajpe8906

Yalew GD 2025. The impact of flipped classroom instruction on college EFL trainees' reading proficiency, perceptions, and engagement. *Participatory Educational Research*, 12(6): 146-168. https://doi.org/10. 17275/per.25.83.12.6

Yañez AM, Adrover-Roig D, Bennasar-Veny M 2023. Personality, preferences, satisfaction, and achievement in a biostatistics course: Traditional versus flipped classrooms in nursing education. *Education Sciences*, 13(2): 197. https://doi.org/10.3390/educsci13020197

Wu Y, Cheng M 2026. Effectiveness of the flipped classroom in medical education: a systematic review and meta-analysis. *BMC Medical Education*. https://doi.org/10.1186/s12909-026-09526-8

Zapata Cardona L 2023. The role of contexts in supporting early statistical reasoning in data modeling. https://doi.org/10.52041/serj.v22i2.448

Paper received for publication in April, 2026
Paper accepted for publication in July, 2026