e-ISBN : 978-629-98755-5-0

SIG : e-Learning@CS
https://appspenang.uitm.edu.my/sigcs/
Publication Date : 24 – Mac - 2025

CLUSTER SAMPLING IN EDUCATIONAL RESEARCH: A PRACTICAL APPROACH

*Nurhafizah Ahmad[1], Fadzilawani Astifar Alias[2] and Siti Asmah Mohamed[3]
**nurha9129@uitm.edu.my[1], fadzilawani.astifar@uitm.edu.my[2], sitiasmah109@uitm.edu.my[3]*

[1,2,3]Jabatan Sains Komputer & Matematik (JSKM),
Universiti Teknologi MARA Cawangan Pulau Pinang, Malaysia

**Corresponding author*

ABSTRACT

Cluster sampling is a widely employed probability sampling technique in educational research, particularly useful for large-scale studies where logistical and financial constraints limit the feasibility of simple random sampling. This method involves selecting entire clusters, such as schools, classrooms, or districts, rather than individual participants, making it ideal for educational settings with naturally occurring group structures. By streamlining data collection processes, cluster sampling enhances efficiency while ensuring representative sampling within a defined population. This paper explores the concept, significance, and practical application of cluster sampling in educational research. It discusses its advantages and limitations and provides an extensive review of empirical studies that have successfully applied this technique. Furthermore, it outlines the procedural steps required for effective implementation, ensuring methodological rigor and minimizing bias. The discussion highlights how cluster sampling has facilitated large-scale educational assessments, policy evaluations, and pedagogical research, reinforcing its value as a methodological tool in contemporary educational research.

Keywords: cluster sampling, educational research, sampling method, statistical analysis

Introduction

Educational research often requires the collection and analysis of data from large and diverse populations. However, obtaining data from every individual in a target population can be prohibitively expensive and time-consuming, particularly when research spans multiple institutions, regions, or even countries. Traditional sampling methods, such as simple random sampling, can become impractical when dealing with educational settings where individuals naturally exist within groups, such as schools, classrooms, or districts.

In such contexts, cluster sampling provides an efficient and cost-effective alternative by selecting entire groups, or clusters, for study instead of sampling individuals independently. This approach reduces logistical constraints while maintaining the representativeness of the sample. By leveraging naturally occurring clusters, researchers can efficiently collect data across large populations without compromising statistical integrity. This paper delves into the theoretical underpinnings of cluster sampling, its applications in educational research, and the methodological steps required for its effective implementation.

Understanding Cluster Sampling

Cluster sampling is a probability sampling technique where a population is divided into distinct clusters, and entire clusters are randomly selected for inclusion in a study. Each cluster should ideally represent the broader population to ensure generalizability. Unlike stratified sampling, which involves selecting individual elements from each subgroup, cluster sampling simplifies the selection process by treating whole groups as units of analysis.

There are two main types of cluster sampling: single-stage and multi-stage sampling. In single-stage cluster sampling, all individuals within selected clusters are included in the study. In contrast, multi-stage cluster sampling involves additional randomization within selected clusters, further refining the sample to enhance representativeness. The choice between these approaches depends on study objectives, resource availability, and population heterogeneity. Despite its efficiency, cluster sampling introduces potential challenges such as intra-cluster correlation, where similarities among individuals within a cluster can reduce statistical precision. However, with appropriate design considerations, these limitations can be mitigated, making cluster sampling a robust tool for educational research (Lohr, 2021).

Application of Cluster Sampling in Educational Research

Cluster sampling has been extensively applied in educational research, particularly in large-scale assessments, policy evaluations, and pedagogical studies. By using naturally occurring groups, such as schools and classrooms, researchers can conduct large-scale studies more efficiently while maintaining statistical validity. Below, we explore how cluster sampling has been applied in key research studies, demonstrating its practical utility and methodological rigor in educational research.

One of the most prominent examples of cluster sampling in educational research is the Programme for International Student Assessment (PISA), an international large-scale educational survey conducted by the Organisation for Economic Co-operation and Development (OECD, 2019). PISA assesses the knowledge and skills of 15-year-old students across various countries in subjects such as reading, mathematics, and science. PISA employs a two-stage cluster sampling approach. In the first stage, a sample of schools is randomly selected from each participating country, ensuring a diverse and representative selection of institutions that reflect regional and socioeconomic variations. In the second stage, a random sample of students within each selected school is drawn. This method allows researchers to make cross-national comparisons without the logistical burden of individually sampling students from an entire national population.

The use of cluster sampling in PISA ensures cost efficiency and operational feasibility in global educational assessments. However, it also requires careful statistical adjustments, such as weighting techniques, to account for intra-cluster correlation, ensuring that results accurately reflect the national

and international student populations. The findings from PISA have informed educational policies worldwide, shaping curriculum development, teaching strategies, and funding allocations in numerous countries.

The National Assessment of Educational Progress (NAEP), often referred to as the "Nation's Report Card," is another large-scale educational study that relies on cluster sampling. NAEP assesses the academic proficiency of students in the United States in subjects such as mathematics, reading, and science (NCES, 2021). Similar to PISA, NAEP uses a multi-stage cluster sampling design. In the first stage, schools are selected as primary clusters based on stratification criteria such as geographical location, school size, and student demographics. In the second stage, students within selected schools are randomly chosen to participate in the assessment. The cluster sampling approach allows NAEP to maintain a nationally representative sample without testing every student in the country. This method reduces data collection costs while providing accurate estimates of student achievement trends over time. NAEP results are used by policymakers, educators, and researchers to evaluate the effectiveness of educational reforms, track achievement gaps, and guide policy decisions at the federal and state levels.

The Trends in International Mathematics and Science Study (TIMSS) is another major international educational assessment that relies on cluster sampling. TIMSS measures the mathematics and science proficiency of fourth- and eighth-grade students across multiple countries (Mullis et al., 2019). TIMSS employs a two-stage cluster sampling process similar to PISA. In the first stage, schools are randomly selected within each country, ensuring diversity in terms of location, funding levels, and student backgrounds. In the second stage, entire classrooms within the selected schools are chosen, rather than individual students. This means that all students within a selected classroom participate in the assessment.

By clustering students within classrooms, TIMSS reduces logistical challenges and standardizes testing conditions, making administration more efficient. The results from TIMSS are widely used by governments and international organizations to compare educational performance globally, develop curriculum improvements, and inform teacher training programs.

A study by Gustafsson (2007) investigated school effectiveness in Sweden using cluster sampling. The research aimed to explore the impact of school-level characteristics on student achievement and educational outcomes. In this study, schools were chosen as the primary clusters, ensuring that data collection captured variations in teaching practices, administrative support, and student demographics. By selecting entire schools rather than individual students, the study was able to examine how institutional factors, such as teacher-student ratios and school funding, influenced academic performance. The findings provided valuable insights into the role of school environments in shaping student success. The study also demonstrated how cluster sampling can be used to assess

school-level differences in educational effectiveness, guiding policymakers in resource allocation and educational planning.

A study by Dunn et al. (2019) examined how class size affects student performance in Canadian primary schools. This study utilized classrooms as the cluster units, allowing researchers to compare different teaching conditions across multiple schools. By randomly selecting classrooms within various schools, the study ensured that differences in instructional methods, student-teacher interactions, and peer effects were systematically captured. The cluster sampling approach was particularly beneficial because it accounted for the fact that students in the same classroom share common educational experiences. Findings from this study contributed to ongoing debates on optimal class sizes, informing school district policies regarding teacher assignments, classroom resources, and curriculum adaptations.

Cluster sampling has also been used to evaluate the effectiveness of teacher training programs. A study by Goldhaber et al. (2020) examined the impact of professional development programs on teaching effectiveness by selecting entire school districts as clusters. In this research, rather than sampling individual teachers, entire districts were chosen, ensuring that variations in district-level training policies, resource availability, and administrative support were captured. The study compared districts that implemented intensive professional development programs with those that did not, analyzing the effects on student learning outcomes. The results provided crucial evidence on the importance of continuous professional development for educators, influencing teacher training policies at both state and national levels.

A study by Fryer (2014) evaluated the impact of merit pay systems on teacher performance and student achievement in the United States. The research used cluster sampling by selecting entire schools that implemented merit pay policies and comparing them to randomly selected control schools. By analyzing full schools rather than individual teachers, the study ensured that findings reflected the broader institutional impact of merit-based compensation, rather than just individual teacher responses. The study found mixed results, with some schools benefiting from merit pay while others showed no significant changes in student outcomes. The findings contributed to the broader discussion on performance-based incentives in education, shaping future policy decisions.

Longitudinal studies in education also benefit from cluster sampling. The Early Childhood Longitudinal Study (ECLS), conducted by the U.S. Department of Education, follows cohorts of students from early childhood through secondary school to track their academic and social development (NCES, 2021). ECLS uses schools as clusters, allowing researchers to follow students within structured learning environments while minimizing attrition rates. The use of cluster sampling in ECLS enables researchers to analyze long-term trends in student development, helping shape early childhood education policies and interventions.

Process of Sampling Using Cluster Sampling

The implementation of cluster sampling in educational research follows a structured process to ensure methodological rigor and representativeness. The first step involves defining the target population and identifying natural clusters within it. In educational settings, these clusters often include schools, classrooms, or districts. Once the clusters are established, researchers develop a sampling frame and determine the sampling approach—whether single-stage or multi-stage cluster sampling.

In single-stage cluster sampling, researchers randomly select a subset of clusters and include all individuals within them. This approach is particularly useful when intra-cluster variability is high, ensuring a diverse representation of the population. In multi-stage cluster sampling, an additional layer of randomization occurs within selected clusters, further refining the sample. For instance, after selecting schools as primary clusters, researchers may randomly select specific classrooms or grade levels within those schools to participate in the study.

Following the selection of clusters, researchers conduct data collection while ensuring adherence to ethical considerations, such as obtaining informed consent from participants and maintaining data confidentiality. Statistical adjustments, such as weighting techniques, are often applied during data analysis to account for clustering effects and enhance the accuracy of population-level inferences. Proper methodological execution ensures that findings remain valid and generalizable despite the inherent clustering of the sample.

Conclusion

Cluster sampling remains a fundamental methodological tool in educational research, enabling the efficient collection of large-scale, representative data. Its application in studies such as PISA, NAEP, TIMSS, and longitudinal research highlights its effectiveness in large-scale assessments. Moreover, school effectiveness research, class size studies, teacher training evaluations, and policy analysis further demonstrate its versatility in addressing key educational questions. While challenges such as intra-cluster correlation exist, careful study design and statistical adjustments help mitigate these issues. As educational research continues to evolve, cluster sampling remains an essential approach for conducting rigorous, policy-relevant, and impactful studies that contribute to the advancement of education worldwide.

References:

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). Wiley. Retrieved from https://onlinelibrary.wiley.com/doi/book/10.1002/9780470973394

Graham, S., Harris, K. R., & Mason, L. H. (2014). *Improving writing in the disciplines: A research synthesis*. Educational Psychology Review, 26(1), 1-21. https://doi.org/10.1007/s10648-013-9238-5

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge. Retrieved from https://www.taylorandfrancis.com/books/mono/10.4324/9780203852279/multilevel-analysis-joop-hox-mirjam-moerbeek-rens-van-de-schoot

Keele, L., & Zubizarreta, J. R. (2017). *Assessing the impact of new curricular approaches through cluster sampling*. Educational Research Quarterly, 40(2), 13-26.

Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. Center for Applied Research and Educational Improvement, University of Minnesota. Retrieved from https://eric.ed.gov/?id=ED485932

Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: Methods and applications* (4th ed.). Wiley.

Lohr, S. L. (2019). *Sampling: Design and analysis* (2nd ed.). Chapman and Hall/CRC.

OECD. (2018). *The survey of adult skills: Reader's companion (2nd ed.)*. OECD Publishing. https://doi.org/10.1787/9789264307273-en

Selwyn, N. (2011). *Education and technology: Key issues and debates*. Continuum. Retrieved from https://www.accordance.com

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage Publications.

Thompson, S. K. (2012). *Sampling* (3rd ed.). Wiley. Retrieved from https://www.wiley.com/en-us/Sampling%2C%2B3rd%2BEdition-p-9781118162934