SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 1.1, (2026), pp. 2492-2506
Open Access. Online & Print

DOI: 10.5281/zenodo. 121126175

THE PPDAC MODEL AND STUDENTS' DATA ANALYSIS COMPETENCE: A CASE STUDY IN VIETNAM

Huynh Gia Bao[1*] and Bui Ngoc Phuong Chau[2]

[1]School of Education, Can Tho University, Can Tho City, Viet Nam, hgbao@ctu.edu.vn
https://orcid.org/0009-0001-6822-679X
[2]Da Nang University of Science and Education, Vietnam, bnpchau@ued.udn.vn
https://orcid.org/0009-0006-4838-1756

Received: 29/09/2025
Accepted: 22/12/2025

Corresponding Author: Huynh Gia Bao (hgbao@ctu.edu.vn)

ABSTRACT

Developing students' data analysis skills is a major priority in Vietnam's 2018 general education curriculum. This research investigates the effectiveness of the PPDAC framework (Problem–Plan–Data–Analysis–Conclusion) in strengthening secondary students' data literacy. An eight-week mixed-method case study was conducted with 240 Grade 10 and 11 students from three high schools located in a province and in Ho Chi Minh City. Results showed that students' overall competency increased by 36%, with mean scores improving from 5.8 to 7.0 (t (239) = 11.62, p < 0.001). The improvement appeared in the Analysis stage, where skills in statistical interpretation, identifying patterns, and drawing evidence-based conclusions more than doubled. Survey responses indicated that 84% of students believed the PPDAC approach made statistics more relevant to real-life contexts, and 76% reported enjoying the use of digital tools to organize, visualize, and interpret data. Teachers observed growth in students' inquiry abilities, collaborative learning, and critical thinking across classroom tasks. Overall, integrating PPDAC with digital technologies provides a promising direction for advancing competency-based education in Vietnamese schools. The findings suggest that structured inquiry, iterative problem solving, and technology-enhanced data practices can support students' development of stronger analytical skills and more confident data-driven reasoning.

KEYWORDS: Data Analysis Competence, Digital Competence, High School Education, PPDAC, Vietnam.

1. INTRODUCTION

1.1. Educational Context in Vietnam

Vietnam is changing its education system with The General Education Curriculum 2018 (GEC 2018) the previous curriculum was based on knowledge memorization whereas the new curricula will cover the key competencies and principles that approve the use of knowledge by the students in practical forms. The movement is consequently a transition of knowledge to maturing student ability to study themselves and solve problems in a fast changing world. Two important competencies, in particular, are valued - both digital competence and data analytic competence. Digital competence is the ability to access, apply and assess digital technology in an efficient and effective manner; and data analytic competence is the ability to assess, acquire, analyze, and interpret data for valuable information and insights. Both competencies are needed for students to fully participate in our data-driven culture, while evidence-based analysis is an increasing form of decision-making in business, education, and politics. For example, students could utilize survey data to engage in a study examining the impact of social media use on academic performance thus demonstrating a true application of data-driven thinking.

Educational institutions provide access to the most up-to-date research and technological data, enabling countries to train skilled staff. Universities play an important role as a source of talent and knowledge generation. The division of higher education system in a country and the science and technology system can lead to poor or wrongful outcomes. Universities equip students with high-quality professional knowledge that guarantees their success in the future (Koyuncuoglu, 2022). The implementation of Fundamental and Comprehensive Education Reform (FCER) in three primary schools with diverse geographical and socioeconomic settings was studied using a qualitative multiple-case study approach. Among methods, interviews, observations, and document analysis were observed, the results indicated that although early adherence to reform instructions were useful to create a momentum and cost savings, new approaches like learner-centered pedagogy were often minimized, especially in underprivileged settings.

The ideas of CAST-based, like Zone of Feasible Practices, Zone of Expected Practices, and Buffering Zone, are used to explain the dynamics of reform. One of them is contextual heterogeneity, low generalizability, and inconsistent school adaptability (Ho, 2024).

Primary education The Statistical Reasoning (SR) model presented was a four-dimensional interdisciplinary project-based learning (PBL) framework. The paradigm applies longitudinal skill development through the SR dimensions, which are Describing, Organizing, Representing, and Analyzing Data and horizontal collaboration across Mathematics, Science, IT, and Language Arts. It is modeled to include realistic low-carbon action projects in the world, which enables full-scale participation of PPDAC cycle. Such accomplishments are systematic SR skill development and transdisciplinary integration. This has limitations such as theoretical design that has not been empirically tested and no numerical performance data, which requires further testing in classroom settings to determine its efficacy (Wang and Shen, 2025). An interdisciplinary project-based approach integrating statistics and computer science education was used to introduce Data Science in secondary schools by CODAP. The approach focused on analyzing multivariate datasets to help students learn statistical concepts and "new skills" for dealing with big data. Outcomes include increased pupil involvement with actual-world data and actual inquiry skills. The numerical results were not reported. The limitations include preliminary deployment without long-term evaluation, reliance on instructor assistance, and the need for broader curricular integration to measure scalability and sustainable growth in skills (Budde et al., 2020).

Vietnam offers an inviting opportunity to address the study. National policies, such as Resolution 29-NQ/TW (2013) Education Strategy for the 21st Century, and Resolution 52-NQ/TW (2019) Education Digital Transformation, focus on how to equip students with technology and data competence as they prepare students for the Fourth Industrial Revolution (Schwab, 2016).

In addition, the Ministry of Education and Training (MOET) supports this with their digital transformation approach in education (2021-2025, vision to 2030), which encourages progress in integrates digital tools and approaches to provide learning experiences that are accessible, flexible and personalized.

This policy environment provides rich grounds on approaches to instructions such as PPDAC (Problem-Plan-Data-Analysis-Conclusion) which integrate theoretical learning with practical implementation as developing skills in a knowledge-based, data-driven world.

1.2. Problem Statement

Regardless of policy views, there remain obstacles in the Vietnamese high school students in terms of statistical reasoning and data analysis to real life scenarios, no longer memorizing formulae. Indicatively, when requested to interpret a statistical graph, or analyze survey data, many of these students do not know how to formulate a research question, develop a suitable data collection procedure, or evaluate evidence to reach conclusions within the data. As noted in international studies (Wild and Pfannkuch, 1999; Hattie, 2009), the capacity of students to work with data should be developed with systematic problem solving exercises that reflect real scientific research. In this context, PPDAC model first created in statistics learning provides a systematic and student-focused strategy to help learners proceed through problem identification to making inferences.

1.3. Research Objectives

- The overall purpose of the research is to investigate how the use of the PPDAC model may improve the data analysis skills of a high school student in the realities of the current educational reform in Vietnam. To this end, the research is informed by a number of objectives that are connected.
- First, it aims to theorize the pedagogical potential of the PPDAC model by analyzing its conceptual basis and connecting it to the existing competency-based education models. In this way, the study provides a strong justification as to why PPDAC, which started as a tool in statistics education, can be suitably applied to the Vietnamese secondary school curriculum.
- Second, the research will explore the interaction and achievement of Vietnamese high school students throughout the steps of the PPDAC cycle viz., Problem, Plan, Data, Analysis, and Conclusion as applied to both statistical exercises and interdisciplinary projects. This goal emphasizes the necessity to know not only the cognitive abilities of students, but also their capacity to extrapolate learning to other contexts.
- Third, the study has been elaborated in the survey base in a classroom, designed in Bình Dương Province and Ho Chi Minh City. The results are expected to clarify possibilities as well as the challenges of implementing PPDAC in different learning environments to contribute to the indication base of educational reform in Vietnam.
- Conclusively; the study will provide a set of recommendations concerning pedagogy and policy that will disrupt the practice of PPDAC in the Vietnamese secondary education. This is in recommendation to teachers, curriculum developers and policy-makers on how they should intentionally integrate inquiry-based models in their practice taking cognizant of systemic elements like digital infrastructure and teacher professional development.

Research question How can the PPDAC model be utilized to enhance the competence of students in data analysis in Vietnamese high schools and what does it imply to pedagogy and policy?

2. LITERATURE REVIEW

2.1. The PPDAC Model in Education

Wild and Pfannkuch (1999) PPDAC model are one of the classical paradigms of guidance analytical research and promoting factual thinking. It breaks down the research process into five processes that are interdependent. Phases: Problem, plan, data, analysis, and conclusion, evidently presented in Figure 1). Contrary to linear models of research, PPDAC involves cyclical and iterative interaction, thus enabling the learners to revisit previous levels where they face new visions or challenges.

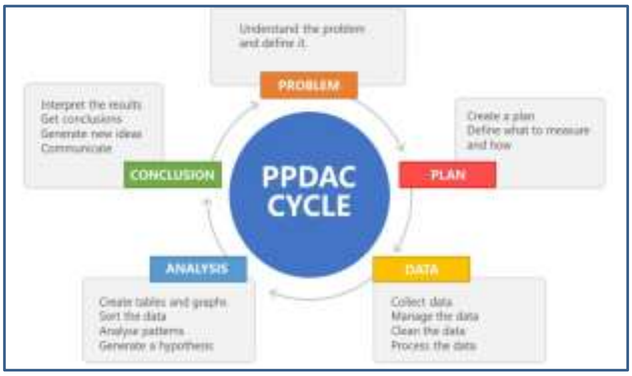

Figure 1 .The PPDAC Model.

Figure 1 illustrates PPDAC Cycle, an organized data search model. The Problem is presented in the Framework first. Next, then, a Plan succeeds to decide the measurement. The data is subsequently gathered, coded and purged. The Framework Analysis is the stage in which data sorting, information that will be used in a graph, and pattern adherence is part of it. Lastly, Conclusion is to make interpretations, illustrate findings, and innovate.

The PPDAC model fits the OCEL.AI storytelling paradigm well since it offers a structured way to incorporate research, investigation and interpretation into machine learning. It aspires students to be systematic when tackling data tasks encouraging them to think critically and make decisions that are thought-provoking. Putting storytelling within this framework a student not only develop more skills in data management, but also the capacity to articulate an idea.This planned but flexible technique increases engagement, confidence, and fosters transferable abilities in STEM and non-STEM situations (Li et al., 2023).

2.1.1. Problem

At this earlier level, students are directed to explain a clear and researchable question based on credible, real world situations. The power of this step is that it transforms the students change in form of vague inquisition to clear investigations which can be empirically resolved. For example, instead of enquiring "Do students like using social media?" a PPDAC-guided preparation would refine this into "What is the relationship between hours spent on social media per day and students' mathematics performance?".

2.1.2. Plan

Once students identify a research issue, they plan a process of collecting and organizing data, which includes selecting what variables to consider and how to select the appropriate participants, and accounting for barriers they had encountered in the past. The planning phase fosters critical thinking and foresight through ensuring the design of the research will support the research purposes once the data has been collected.

2.1.3. Data

At this stage in the process, students begin to collect, record, and organize data to analyze later. Online tools such as Google Forms, Excel or a learning management system allow new users to create surveys, gather responses and maintain the integrity and validity of their data with simplicity. Organized data can include coding variables, cleaning up any errors, and abstracting data to analyze and further analyze it.

2.1.4. Activity

This step involves statistical thinking and digital skills that help students reveal the patterns, trends, and relationships in the data. It may require descriptive statistics (e.g. measures of central tendency, percentages, graphs), inferential statistics (e.g. correlations, regressions), or visual analytics, depending on the level of understanding needed. Importantly, the affordances provided by contemporary software, specifically, render this step of the processes as the most interactive and engaging, and learners have the opportunity to experiment with the variability of representing the same set of data in each case.

2.1.5. Conclusion

The final chapter involves reviewing findings based on the initial research question. Students are also challenged to relate their findings to possible implications as well as reflect on limitations and offer possible recommendations to future research or action. This step is of higher order thinking because students derive meaningful knowledge out of evidence.

The difference between PPDAC and linear representations is that it is cyclical and iterative. When students experience difficulties or new understandings, they typically return to earlier steps, modifying their research problem after they establish limits collecting data, or revising their analysis plan after uncovering anomalies with the data set. Their cyclical engagement embodies the realities of professional researchers, engaging students as agents towards the creation of knowledge, rather than a passive consumer. Along with its uses in statistics, PPDAC has been recognized by researchers as usable to teach science, as a part of project-based learning, and as a means to collaborate across disciplines or subject areas (Pfannkuch & Wild, 2004; Chance, 2018). In the science classroom, PPDAC provides an organized way to plan an investigation and draw conclusions based on data. While in project-based learning, PPDAC has provided a systematic organization for the inquiry cycle across disciplines that support student decision-making and working through real-world problems.

2.2. Data Literacy, Competence and PPDAC Skill

Based on the OECD (2019), data literacy is an important component of a Digital Competency in the information society. Students are expected to be able to manage digital information and use critical thinking skills when evaluating, analyzing, and applying data to problem solving. In Vietnam, GEC 2018 states that discipline/courses in mathematics and science should help to develop students' problem solving and data analysis skills.

Holbrook and Rannikmae (2009) examine that the use of digital tools in science education supports students' capacity for reasoning from data, inquiry and critical thinking. Limniou and Papadopoulos

(2014) also demonstrated that online learning environments and specific software (e.g., PhET, ChemDraw, GeoGebra) were positively associated with self-directed learning and creativity in students' data activities.

Empirical studies indicate that teaching with the PPDAC model improves students' ability to design investigations and interpret statistical results. For instance, Pfannkuch (2011) found that secondary students using PPDAC performed better in identifying variables, designing surveys, and interpreting variability. In Asian contexts, research is emerging on adapting PPDAC to fit cultural and curricular conditions (Chan, 2019). However, there is limited evidence from Vietnam, where curriculum reform is relatively recent.

A Design-Based Research (DBR) approach was used to develop a framework for evaluating teachers' statistical lesson planning. The model consisted of seven criteria which covered learning intentions, task aspects and constructivist approach. Professional development workshops allowed teachers to use statistical literacy concepts, update lesson plans and enhance the structure via repeated feedback. The findings reflected an improved planning and closer intertwining of actual data and representations. Although, it was evident that teacher planning skills increased, the results were not quantitative and were qualitative, which limited the possibility of generalization (Tran et al., 2023). Cores skills of interpretation, relevance assessment, protection, contextual adaptation and integration are emphasized in semi-structured interviews with industry partner methods. Success stories also demonstrate a pragmatic skills model and the need to involve universities and industries in the curriculum development process. Quantitative outcomes were not provided, so their applicability and validity to non-stakeholder scenarios were limited (Taş, 2024).

The study uses a design-based research-based curriculum intervention that uses scientific thinking, modeling, and simulation-based inference. The plan aims at a data-driven approach in order to enhance statistical knowledge and practice among tertiary students. There is an increase in conceptual understanding and bridging between theory and practice among students and teacher ratings. The suggested improvements concern higher involvement and statistical literacy; however, numerical results are limited to qualitative impressions, requiring a wider scope of numerical evidence in different learning contexts. Both the students and educators state significant achievements in feedback. Students have reported a better conceptual grasp of statistical concepts, and have developed a better sense of the general usefulness of statistics to real life. The teachers indicated a higher rate of engagement and a more productive approach to bringing the element of statistical rationale into disciplinary contexts by the learners. The findings highlight the possibility of curriculum changes that involve the use of technology-mediated, simulation-based learning activities as a tool to underscore the bridging role between theory and practice (Gehrke et al., 2021). In order to examine the effect of augmented reality on data analysis of the engineering students as well as solving problems, which is a skill taught to them, a quasi-experimental and mixed method was used. The intervention consisted of teacher and student sessions on the design of learning objects (LOs) to be used in their courses. Besides, a questionnaire was designed to investigate the positions of students on the effectiveness of augmented reality to support their learning processes. To advance the methodological rigor, the statistical tests were conducted through Wilcoxon to proving the significance of the performance differences and multiple correspondence analysis (MCA) to explore the relationships between variables considered related to AR use and academic achievement. The results demonstrated quantifiable enhancements in the performance, especially higher EGEL examination scores, as proof of the positive effect of AR on competency development. Learners were also more interested and motivated because they found augmented reality to be an effective tool in their school-related preparation. The successes of the intervention can be quantified and qualitatively, and engagement and satisfaction should be mentioned. Nevertheless, it is restricted by factors like small sample size of two engineering projects and use of short term results without a longitudinal follow up. To confirm these promising findings, more extensive, multi-institutional studies are needed to guarantee that they can be generalized (Zamora-Antuñano et al., 2022).

The literature review and environmental scan were applied to find and synthesize undergraduate and master research capabilities. The results ascertained 7 key competencies, namely critical appraisal, information synthesis, decision making, problem solving, data collecting, data analysis, and communication, broadly built in both thesis and non-thesis paths. Despite the absence of direct comparisons of types of programs, skill profiles were analogous. Numical data on prevalence were not reported. Variables definitions of research skills across studies and limited empirical comparisons of program formats are also limitations to generalizability (Vieno et al., 2022). The pre-service mathematics teachers (PSMT), in Egypt, were studied qualitatively to unravel the difference between statistical theory and teaching practice. Grossman (1990) methodology was used to evaluate the national academic policies, and focus group discussions with PSMTs provided analytical insights. The most significant development was the adoption of a constructivist-informed Microteaching Lesson Study (MLS) approach, which enhanced awareness of the topic of pedagogy in statistics education. Although qualitative studies reported better teacher perceptions, limitations consist of absence of numerical results and generalization outside the Egyptian context (Elbehary, 2020).

2.3. Research Gap

While foreign research shows that the PPDAC model improves data literacy and inquiry-based learning, no previous Vietnamese studies have looked at its use in schools. This lack of localized evidence emphasizes the need to investigate how PPDAC might be tailored to Vietnam's competency-based education reforms. This confirms that no prior Vietnamese studies on PPDAC.

3. METHODOLOGY

3.1. Research Design

This research adopted a case study combined with quasi-experimental classroom intervention, which is suitable for exploring how pedagogical models affect students' competencies in real educational settings. Both qualitative and quantitative methods were employed to capture the multidimensional nature of students' data analysis competence. Figure 2 depicts the qualitative components (classroom observations, teacher journals, and student comments) and quantitative components (pre- and post-tests, surveys, and rubric-based assessments) were combined in the PPDAC paradigm. It focuses on the instructional stages Problem, Plan, Data, Analysis, and Conclusion as well as the accompanying student learning activities, instructor support, and data collecting points for assessing data analysis ability.

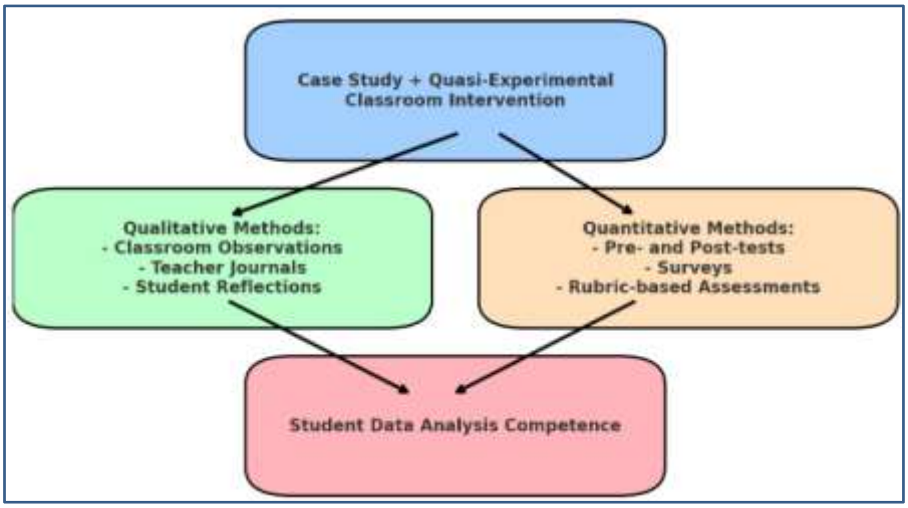

Figure 2: Research Design.

Qualitative components included classroom observations, teacher journals, and student reflections to provide insights into learning processes during PPDAC-based activities.

Quantitative components consisted of pre- and post-tests, surveys, and rubric-based performance assessments to measure changes in students' competence levels.

3.2. Research Context and Participants

The sampling included three high schools in Bình Dương Province and Ho Chi Minh City, two areas that reflect fast developing school settings in Southern Vietnam.

Participants were sampled purposively (Grade 10 and Grade 11 classrooms) to identify those who had studied mathematics and science modules in previous years pertaining to the PPDAC approach. Classes were selected so that both genders and academic achievement levels could be represented equally. Randomization was not feasible because of the school schedule constraints but we tried to balance the characteristics of participants across the sites.

Participants Grade 10 and Grade 11 classes consisting of 240 students (127 female and 113 male).

Teachers 6 teachers of mathematics and science who had attended a brief training workshop on PPDAC and digital tools in the past.

Duration The intervention was 8 weeks long, and it was a part of the mathematics statistics course and project-based learning in science.

A mixed-method design was chosen because it enables for both quantitative assessment of competence improvements by tests, surveys, and rubrics) and qualitative analysis of learning processes (by assessments, documents, and comments). Combining both strands improves validity by allowing triangulation and providing a more complete knowledge than either strategy alone.

Bình Dương Province and Ho Chi Minh City were selected to reflect rapidly emerging educational contexts in Southern Vietnam, with diverse levels of urbanization, access to digital tools, and educational resources. This diversity enables for an analysis of the PPDAC model's performance across diverse educational contexts, as well as insights regarding possible scaling in a variety of settings.

3.2.1. Instruments and Tools

Pre- and post-tests It is aimed at assessing the level of skills of students to locate research questions, manipulate datasets, and explain statistical graphs.

Rubrics It was created to evaluate the performance at every phase of the PPDAC cycle (Problem, Plan, Data, Analysis, and Conclusion). Every rubric had a 3-level scale (Basic - Proficient - Advanced) (Table 1 and 2)

Table 1: Data Analysis Competence Framework (PPDAC).

Component capacity	Behavioral manifestations (Criteria)
1.Problem Formulation (P)	1.1.Identify real-world issues 1.2.Formulate measurable questions
2.Planning & Data Collection (P)	2.1.Design data collection plans 2.2.Use digital tools for data gathering
3.Data Processing (D)	3.1.Clean and code raw data 3.2.Manage datasets with software
4.Data Analysis & Interpretation (A)	4.1.Apply statistical methods 4.2.Interpret trends and relationships
5.Conclusion & Communication (C)	5.1.Draw evidence-based conclusions 5.2.Present findings with digital tools

Table 2: Rubric for Assessing Students' Data Analysis Competence (PPDAC Model).

Indicator	Basic	Proficient	Advanced
Identify real-world issues	Needs guidance to recognize issues	Identifies meaningful issues with support	Independently identifies complex, relevant issues
Formulate measurable questions	Struggles to frame researchable questions	Frames clear questions with some support	Consistently formulates precise, independent questions
Design data collection plans	Plans incomplete or vague	Designs feasible plans with essential details	Designs comprehensive, valid, systematic plans
Use digital tools	Rarely uses tools effectively	Uses tools with moderate accuracy	Integrates multiple tools efficiently
Clean and code data	Cannot clean or code properly	Cleans and codes with reasonable accuracy	Applies advanced coding reliably
Manage datasets	Struggles with software	Manages moderate datasets	Masters large/complex datasets
Apply statistical methods	Uses methods incorrectly	Applies appropriate methods correctly	Selects and applies advanced methods effectively
Interpret trends	Interpretations incomplete/inaccurate	Interprets trends with reasonable accuracy	Provides deep, contextual insights
Draw conclusions	Lacks evidence	Draws data-supported conclusions	Synthesizes insightful, evidence-based conclusions
Present findings	Struggles to present	Presents clearly with tools	Creates impactful, multimodal presentations

Survey questionnaire Adaptation of OECDs student digital competence framework to survey selfperceived data handling skills.

Digital tools Google Forms (creating a data collection tool), Microsoft Excel (cleaning and visualization data), GeoGebra SPSS (basic statistical analysis), and (mathematical modeling).

3.2.2. Validity and Reliability

The pre- test and post-tests, rubrics, and questionnaires were developed based on existing frameworks, particularly, the OECD having student digital competency model. To ensure, expert review was done by experienced mathematics and science teacher's content authenticity. The reliability of the tests was determined by conducting a pilot study on 30 students of a similar class and rubrics, and Cronbachs of above 0.80 insisted on satisfactory internal reliability.

3.2.3. Teaching Intervention

The intervention consisted of one theme of the project that linked mathematics with real-life situations Letter: Social media and learning outcomes. Students examined the hypothesis that hours spent on social media per day were related to mathematics test scores (Table 3).

Table 3; Social Media and Learning Outcomes (PPDAC Model).

PPDAC Stage	Teacher Activities	Student Activities
Problem	Introduce context & guide research question framing	Discuss, define, and agree on research question
Plan	Introduce survey tools; guide questionnaire design	Design online questionnaire; select variables
Data	Monitor collection; check errors	Collect survey responses; clean and prepare dataset
Analysis	Demonstrate Excel/SPSS analysis; guide interpretation	Compute statistics; test relationships (e.g., Pearson correlation)
Conclusion	Guide report writing and visualization	Write report; present findings using slides/posters

3.2.3. Data Analysis Procedures

- Pre- and post-test scores were analyzed by Descriptive statistics.
- Paired sample t-tests were performed to identify the value of the improvement in competence of students.
- Rubric scores were added and compared among PPDAC stages.
- The thematic coding of qualitative data, based on observations and reflections, was performed to determine patterns in student learning behaviors.

The topic of social media and learning outcomes was selected due to its applicability to the daily lives of students data analysis task significant and engaging. In addition, it allowed exploring quantitative correlations applying the easily accessible data and relating mathematics and science curricular goals. This project is an example of the multiple real-life situations in which the PPDAC approach can be implemented.

4. RESULTS

4.1. Improvement in Students' Overall Competence

The comparison of the pre- and post-test scores revealed that the students developed the overall data analysis skills significantly in terms of means, standard deviation, paired t-test results with 2 tails, and the d effect size (Table 4). Mean pre-test results: 5.8, SD = 1.2; post-test results: 7.9, SD = 1.3. This provides readers with information on score fluctuation.

Table 4: Pre-test and Post-test Comparison.

Statistic	Value
Mean pre-test score	5.8 / 10
Mean post-test score	7.9 / 10
Paired t-test result	$t(239) = 11.62$
Significance level (p-value)	$p < 0.001$

This shows statistically significant change which implies that the PPDAC-based intervention had a quantifiable effect increased influence on data analysis skills of students. P-values are used to identify whether the change is statistically significant, effect sizes demonstrate the size of the benefit and the reader is able to understand the intervention better practical impact. In this experiment, the gain in the overall data analysis skill of students acquired a Cohen d of 0.75, pointing to a very large effect. The reporting of the effect size will provide a measurable determination of the benefit seen, beyond statistical significance ($p < 0.001$). Means are not variability in students. Having SDs educates readers regarding the allocation of outcomes and the presence of constant progresses among respondents or a small number of best performers. Table 4 has added standard deviations. It is inappropriate to use such words as improved or substantial shift qualitative. Consider precise percentages before and after the intervention in order to measure the change, and the results available, accurate and consistent. Basic level students declined to 42 and on down to 15. Excellent and Superior improved to 48 and 60 percent and 10 and 25 percent respectively (Figure 3). Cronbach's alpha means that the survey instrument has internal consistency, and student responses are a correct measure of the survey was not random, giving the findings credibility. The survey questionnaire was evaluated on reliability and gave a Cronbachs alpha of 0.82, which implies good internal consistency. Quantitative data are the baseline scores, post-test scores, and effect sizes, stage proficiency percentages, and reliability in the survey. Instructor reflections and qualitative findings also encompass instructor reflections Open-ended student responses that focus on learning habits, teamwork and confidence.

4.2. Performance across PPDAC Stages

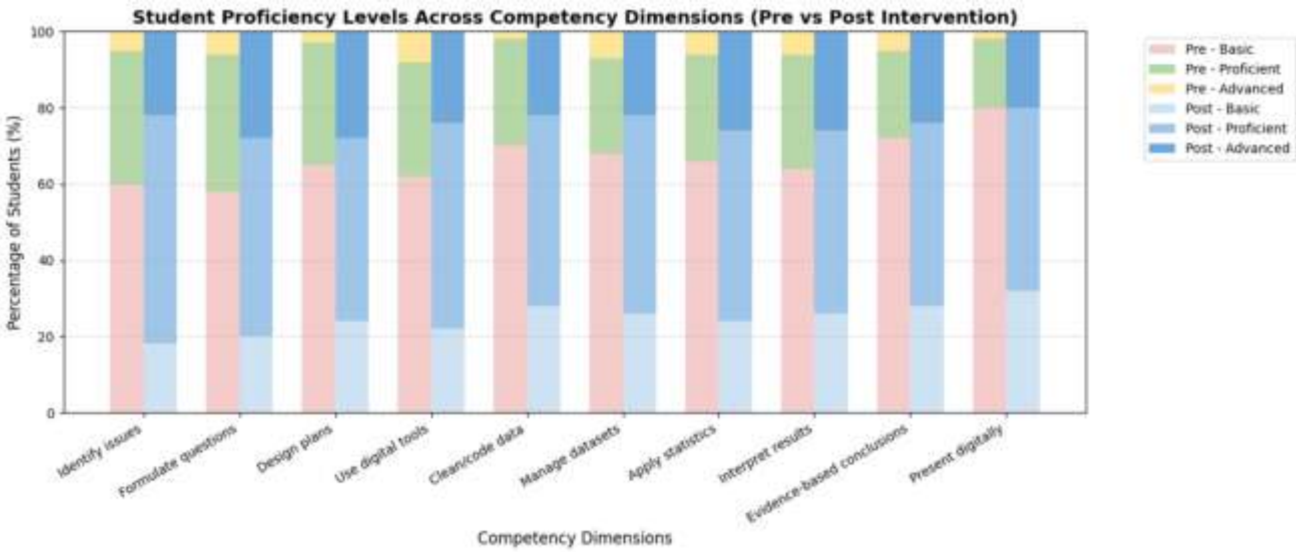

Figure 3: Student Performance across Ten Data Analysis Competence Indicators Before and After the PPDAC-Based Intervention. Percentages Are Shown for Three Proficiency Levels.

The findings imply a comprehensive change in student performance in all ten indicators. After the PPDAC There were fewer students left at the Basic level, and the proportions at the Proficient and Advanced levels were lower grew significantly, especially in statistical analysis, interpretation, and drawing of evidence-based conclusions. This insinuates that the framework significantly developed technical and higher-order data analysis skills.

4.2.1. Students' Perceptions of PPDAC

The results of the surveys showed that students became affirmative towards the PPDAC framework. Most of them (84%) said that PPDAC provided a clearer picture of how learning statistics apply in real life, a factor that strengthens the linkage between classroom activities and real world uses. Moreover, 76% of students stated that the implementation of digital tools made data collection and analysis processes more interesting and more available. Furthermore, 68 percent of them were more confident in their ability to present their findings to their peers, which shows better competence and increased self-efficacy in communication. These findings imply that PPDAC enhanced not only technical competencies but also motivated and created confidence in statistical inquiry in the students.

4.2.2. Teachers' Reflections

Reflections of the teachers reflected the positive aspects in terms of the student performance. In particular, the teachers mentioned that PPDAC was a concise, logical system to assist the students in the inquiry-based learning at every level of formulating a problem to the solution. The teachers also commented that students were more active and cooperated and assumed more responsibility in group projects, which allude that the framework can have potential in developing these abilities of collaboration and independent thinking. Regardless of these positive remarks, the educators observed that it was not always easy to make sure that the entire group of students will achieve the same level of proficiency in using the digital tools, since not all learners can accomplish this on their own without support and scaffolding. It underlines that training was needed that would: allow a more differentiated provision of PPDAC and specialized digital literacy training post-implementation of PPDAC. The quantitative results comprise both pre and posttest scores; effect sizes; percentages by the stage; and reliability of surveys. Qualitative results will comprise instructor comments and open-ended answers of students, which will demonstrate the perception of learning habits, team work, and learning confidence.Include allusions to self-determination theory (Deci and Ryan, 2000) and prior research findings on intrinsic motivation to demonstrate how the effective organization of the PPDAC model, the decision making concerning the formulation of problems, and group based project-based activities might have facilitated the engagement of the students, their confidence in their own abilities, and their persistence in learning data analysis.

The reflection of teachers in Knowledge Creating Communities (KCC) suggested that the confidence and data literacy teaching capability were significantly increased. Some respondents noted that their feeling of power was slowly replacing the feeling of being nervous about giving information initially. Consideration of peers provided an opportunity to discuss a range of matters and solutions and professionals of the university learned more about the topic. Teachers were also exposed to PPDAC as an organized and systemic approach to data literacy education. Identifying relevant issues provided students with authentic entry points while careful sequencing ensured appropriate datasets and classroom activities for the students. The analysis step allowed students to exercise interpretation abilities while teachers facilitated debate and dissemination of findings. Finally, reflecting on the findings helped link statistical knowledge to actual-life decision-making (Robertson et al., 2023)

5. DISCUSSION

The results demonstrate that the PPDAC model is a highly effective pedagogical model to develop competence in data analysis for Vietnamese high school students. The pre and post-test improvement of 5.8 to 7.9 ($t(239) = 11.62$, $p < .001$) in post-test score and rubric-based performance across ten indicators provides strong evidence that PPDAC advances both foundational and advanced skills. PPDAC modified the process of statistical inquiry to manageable stages that relate abstract concepts (e.g. define variables, correlation, and meaning) into tasks with relevance in the real world. In a Vietnamese context that has traditionally involved rote learning, prioritizing recall over reasoning, the PPDAC structuring into manageable stages was valuable.

The analysis by stages of the course indicated that students progressed notably in the Analysis stage, with their ability to interpret graphs, calculate descriptive statistics and to use correlation/regression more than doubling after the intervention. The Problem and Plan phases revealed substantial improvements, with students progressing away to general and vague questions to specific and quantifiable ones with systematic research designs. Data stage also showed substantial progress, as students were able to successfully gather, code, and organize datasets with digital tools, including Google Forms and Excel, which are becoming crucial skills in the digital transformation plan of Vietnam. Despite these positive results recorded in the Conclusion stage, especially in evidence-based reasoning, digital presentation was a comparative weakness since less than 60 % of students were found to have attained proficiency, and closer integration of digital communication and visualization competencies is required.

The PPDAC cycle should be improved to make it stronger in clarity, efficiency, and applicability of data-driven projects. During the Problem phase, ambiguity is minimized by defining measurable questions that are related to workable objectives. Planned stages have the advantage of premeditated variable choice, standardized information-gathering steps, and computer-based or computer-assisted instruments that enhance predictability. At the phase of Data, sophisticated technologies such as sensors, simulations, and web sites allow more detailed datasets with aiding accuracy. The use of mathematical tools, statistical software, or machine-learning uncovers more dramatic trends in analytical steps. Lastly, the visualization tools, automated reporting, and verification techniques enhance the presentation of the results and allow the cycle to be refined iteratively (Trienens et al., 2024).

PPDAC was iterative, and it was necessary: students revisited questions and plans in case of ambiguous data or sampling difficulties, which is reflective of the real world of research (Pfannkuch & Wild, 2004; Chance, 2018). The data collected in surveys indicated a good perceived value with 84% confirming real-world relevance, 76% using digital tools, and 68% citing confidence increment. In sum, PPDAC was highly flexible to Vietnamese secondary education, which supports data literacy and digital skills highlighted in the Digital Transformation Strategy (MOET, 2021).

5.1. *Integration of Digital Tools*

The convention of digital tools was dangerous in growing involvement and capability throughout the PPDAC process. These tools assisted students in organizing, analyzing, and synthesizing data, as well as presenting their findings in straightforward visual formats. In addition to developing technical skills, the tools supported digital literacy and prompted greater confidence in dealing with datasets and communicating data-based findings. This convergence parallels the OECD (2019) assert that the two constructs, data literacy and digital competence are interconnected. In Vietnam, where the Ministry of Education and Training (MOET, 2021; 2022) encourages digital and competency-based education transformation, application of technologies in PPDAC enhanced performance whilst meeting national policy objectives of competencies required in the 21 st century, innovation and applied learning. In general, the findings indicate that electronic. The resources are not only supplements, but the main enrichers of the inquiry-based learning, providing the students with practical knowledge and desire to engage in statistical analysis wholly. The introductory Data science course to high school students manages to work around the PPDAC cycle and electronic resources, including CODAP. Students began by identifying relevant concerns within the survey environment (Problem), such as patterns in leisure activities and media use. During the planning phase (Plan), they developed research topics and data exploration methodologies based on the dataset's structure. Students used CODAP to access and alter the JIM-PB survey data (Data), allowing for hands-on involvement with real-world information. The platform offered visualization, filtering, and summarization, allowing students to analyze trends, correlations, and distributions. Finally, students developed evidence-based interpretations and presented their results (Conclusion), highlighting the importance of making decisions based on data (Frischemeier et al., 2021). The incorporation of PPDAC with electronic resources not only aided in the systematic development of statistical reasoning, but it also improved students' technical literacy, inquiry skills, and comprehension of how computer programs can support genuine data exploration and significant insights in a sequential, iterative learning process.

The image-based data collection process involves the use of digital tools with the PPDAC cycle to enhance the ability of pre-service teachers to analyze. At the Problem stage, the members came up with photo-based investigative questions. In the Plan, they determined what factors to eliminate and the manner in which to group the observations. Data was captured in the Data phase, grouped and organized using common online documents, and observational judgments and inferential interpretations were combined. In Analysis, variables were compared and classified and measured by the participants to find associations and trends digitally. Lastly, structured observations were also learned, highlighting the conflict between subjective interpretation and objective evaluation, and how digital tools help to improve PPDAC engagement on an iterative basis (Kazak, 2025).

The issue of equity has been included and it has emphasized the impact the access to digital devices and a stable internet has on student engagement. Next we shall examine how the digital divide could impact the scalability of PPDAC, and why the policies that offer equitable access to every student, especially in a rural or underserved school, are important. Our results include certain allusions to the Digital Transformation Strategy of MOET (2021, 2022) and national competency-based curriculum objectives. We show that integrating PPDAC and digital solutions aligns with the national objectives of innovation, digital literacy, and skills of the 21st century, and give policymakers practical recommendations.

5.2. Development of Transversal Skills

In addition to analyzing data, the PPDAC framework also helped in the creation of transversal competencies that are crucial to the learning of the 21 st century. Teamwork was encouraged with group projects involving a negotiation process, division of labor, and provision of constructive peer feedback. Thinking was enhanced through critical analysis of the student's quality of their data, challenged the representativeness of their samples and discussed other possible causes of observed trends. The necessity to report the results in digital formats, improved communication skills which challenged students to describe their findings in a clear and convincing way to fellow students and teachers. These outcomes echo the meta-analysis by Hattie (2009) which proved that inquiry-based and project-based methods have powerful effects improved student performance and motivation.

5.3. Challenges and Limitations

Though the outcomes were encouraging, several difficulties were evident in the application of the PPDAC model. First, the digital divide was overcome by the majority of students, but the absence of personal devices, and/or a trusted one. Their digital data collection and analysis was hindered by the internet connection. Second, teacher preparedness created challenges; several teachers should have undergone a broader training of PPDAC pedagogy and statistical software application since the first workshops provided were too brief to make the teachers feel ready. Third, the constraints of time reduced profoundness. The eight weeks was not enough time to enable students to take some more adventurous data analysis methods like hypothesis test or regression analysis. Lastly, the cultural background of Vietnamese classrooms that are traditionally teacher-centered entailed that not all the students at first were able to own the problem formulation process. These issues signify that although PPDAC has great potential, it will need systemic scaled-up to use advocates, especially in professional growth and equal access to digital assets.

5.4. Implications for Policy and Practice

The research provides a number of implications on effective competency-based education in Vietnam.The findings have profound policy and practice implications to enhance online Self-Regulated Learning (SRL) in open universities. At the policy level, education institutions must integrate these SRL skills in their curricula and support systems by making them a formal part of the online education. This includes scheduled learning programs, workshops, and orientation courses that will educate students on how to manage their time, organize their work, and find assistance. Equality issues could also be addressed by policy, especially in female and aged students, who reported lower achievement, through special support and other learning options. In practice, professors and instructional designers should use organized instruction to gradually increase autonomy while providing feedback on SRL-related actions. Learning Management ystems (LMS) could include data and self-monitoring instruments to assist students in tracking their development, reflecting on learning objectives, and adjusting techniques (Santoso et al., 2022).

Conducting quantitative assessments of Workplace Psychosocial Hazards (WPH) research is vital for future policies and practices that enhance worker well-being. In regards to policies, regulators and state/federal bodies should prioritize advancing WPR into national workplace health and safety polices (United Nations, 2020). Policies should not just extend toward personal safety related to injury or disease, but should also extend inclusive of a focus on stress, mental health, and other psychosocial issues that are commonly understudied in many contexts. It may make sense for regulations, individually or collectively, to require some form of psychological risk assessment, reporting and prevention protocols within compliance for employees. Also, the fact that the evaluation of policies and the long-term prosperity connects the various stakeholders with the WPR specialists should have a common comprehension of the management operations so that the management of WPR in the organization could result in long-term productivity and earned equity in all areas.

Practically, employers ought to incorporate psychological risk management in their WSHM systems. These cover proactive monitoring of mental health evidence, development of training activities of managers and employees and enabling organisational cultures that eradicate stigma of psychosocial ailments. This should also treat shift work, activity balance, and job design that are not being studied but are very essential in prevention. The development of collaborative efforts between researchers, companies, and policymakers will ensure that evidence-based initiatives are adopted successfully in the workplace (Dhakal and Mahmood, 2025).

6. CONCLUSION

This research demonstrates that the application of the PPDAC paradigm to teaching mathematics and science in Vietnamese high schools enhances the data analysis skills of students to a greater extent. The intervention demonstrates that the PPDAC is applicable and efficient in the Vietnamese context, which aligns with the General Education Curriculum 2018 (GEC 2018), Digital Transformation Strategy of MOET (2021-2025), and other global standards of 21st century skills development.

6.1. Policy and Practice Implications

- Teachers: Professional development in PPDAC education and the real use of digital tools.
- To Curriculum Designers: Incorporate PPDAC into mathematics and multidisciplinary modules in order to develop subject-specific and transversal competencies.
- To policymakers: Invest on the digital infrastructure and equal access to technology to facilitate inquiry based, technology enhanced learning.

6.2. Future Research Directions

- Perform longitudinal research to determine sustainability of competence increases with time.
- Explore the adoption of advanced statistical methods (regression and testing hypothesis) into PPDAC.
- Exhibit how digital literacy can be institutionalized within competency-based education, particularly in schools that are under resourced. Overall, this paper acknowledges that PPDAC is a scalable and sustainable model that

fills the gap between the traditional curriculum objectives and emerging competency-based models that enable students to not only equip themselves, but also conduct data analytics analysis in the real world.

REFERENCES

Budde, L., Frischemeier, D., Biehler, R., Fleischer, Y., Gerstenberger, D., Podworny, S., & Schulte, C. (2020). Data science education in secondary school: How to develop statistical reasoning when exploring data using CODAP. IASE Roundtable.

Central Committee of the Communist Party of Vietnam. (2013). Resolution No. 29-NQ/TW on fundamental and comprehensive innovation in education and training. Hanoi: Author.

Chance, B. (2018). Thinking and reasoning with data: A review of research. Journal of Statistics Education, 26(2), 71–83. https://doi.org/10.1080/10691898.2018.1436999

Deming, W. E. (1986). Out of the crisis. MIT Press.

Dhakal, S. P., & Mahmood, M. N. (2025). A scientometric analysis of three decades of research on workplace psychosocial hazards: Implications for policy and practice. Journal of Safety Research, 93, 79–89. https://doi.org/10.1016/j.jsr.2025.02.011

Elbehary, S. G. A. (2020). Teacher education of statistics from theory to practice. Journal of Applied Research in Higher Education, 12(5), 857–869. https://doi.org/10.1108/JARHE-06-2019-0141

Frischemeier, D., Biehler, R., Podworny, S., & Budde, L. (2021). A first introduction to data science education in secondary schools: Teaching and learning about data exploration with CODAP using survey data. Teaching Statistics, 43, S182–S189.

Gehrke, M., Kistler, T., Lübke, K., Markgraf, N., Krol, B., & Sauer, S. (2021). Statistics education from a data-centric perspective. Teaching Statistics, 43, S201–S215.

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Ho, L. H. (2024). Implementing education reforms: A multiple case study of Vietnamese primary schools (Doctoral dissertation, University of Glasgow).

Holbrook, J., & Rannikmae, M. (2009). The meaning of scientific literacy. International Journal of Environmental and Science Education, 4(3), 275–288.

Kazak, S. (2025). Data-ing process with image-based data: Variable identification and generation. ZDM–Mathematics Education, 57(1), 61–74. https://doi.org/10.1007/s11858-025-01656-5

Koyuncuoglu, D. (2022). Analysis of digital and technological competencies of university students. International Journal of Education in Mathematics, Science and Technology, 10(4), 971–988. https://doi.org/10.46328/ijemst.2583

Li, Y., Wang, Y., Lee, Y., Chen, H., Petri, A. N., & Cha, T. (2023). Teaching data science through storytelling: Improving undergraduate data literacy. Thinking Skills and Creativity, 48, 101311. https://doi.org/10.1016/j.tsc.2023.101311

Limniou, M., & Papadopoulos, N. (2014). The development and evaluation of a 3D computer-generated learning environment for chemistry education. Journal of Science Education and Technology, 23(2), 178–191. https://doi.org/10.1007/s10956-013-9462-8

Ministry of Education and Training. (2018). Comprehensive general education program. Hanoi: Vietnam Education Publishing House.

Ministry of Education and Training. (2022). Digital transformation strategy for the education and training sector for the period 2021–2025, with a vision to 2030. Hanoi: Author.

OECD. (2019). OECD skills outlook 2019: Thriving in a digital world. OECD Publishing. https://doi.org/10.1787/df80bc12-en

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning and thinking (pp. 17–46). Springer.

Politburo. (2019). Resolution No. 52-NQ/TW on a number of guidelines and policies for proactive participation in the Fourth Industrial Revolution. Hanoi: Author.

Robertson, J., Abaci, S., Linklater, H., Farrell, K., & Kanwal, J. (2023). Building teachers' confidence and competence in teaching data literacy through knowledge creating communities.

Santoso, H. B., Riyanti, R. D., Prastati, T., S, F. T. H., Susanty, A., & Yang, M. (2022). Learners' online self-regulated learning skills in Indonesia Open University: Implications for policies and practice. Education Sciences, 12(7), 469. https://doi.org/10.3390/educsci12070469

Schwab, K. (2016). The Fourth Industrial Revolution. World Economic Forum.

Taş, E. (2024). Data literacy education through university–industry collaboration. Information and Learning Sciences, 125(5/6), 389–405. https://doi.org/10.1108/ILS-06-2023-0077

Tran, D., Nguyen, A. T. T., Nguyen, D. T., Ta, P. M. T., Huynh, B. T., Phan, P. T., Pham, N. T., & Nguyen, N. H. T. (2023). Planning for developing students' statistical literacy: A research-informed framework development. Vietnam Journal of Education, 74–81. https://doi.org/10.52296/vje.2023.227

Trienens, D., Schöppner, V., Krause, P., Bäck, T., Tsi-Nda Lontsi, S., & Budde, F. (2024). Method development for the prediction of melt quality in the extrusion process. Polymers, 16(9), 1197. https://doi.org/10.3390/polym16091197

Vieno, K., Rogers, K. A., & Campbell, N. (2022). Broadening the definition of 'research skills' to enhance students' competence across undergraduate and master's programs. Education Sciences, 12(10), 642. https://doi.org/10.3390/educsci12100642

Wang, Y., & Shen, Q. (2025). A theoretical framework for interdisciplinary instructional design targeting statistical reasoning. iEducation, 1(2), 30–48. https://doi.org/10.71204/as5s4879

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Zamora-Antuñano, M. A., Barros-Baertl, R., Tovar-Luna, B., González-Gutiérrez, C. A., Mendez-Lozano, N. E., & Cruz-Perez, M. Á. (2022). The use of augmented reality to strengthen competence in data analysis and problem solving in engineering students at the Universidad del Valle de México. Education Sciences, 12(11), 755. https://doi.org/10.3390/educsci12110755