



Evaluating a Basic Education Tutorial Support Program: Implementation, Perceived Outcomes, and Institutional Improvement Priorities

DOI: 10.5281/zenodo.21700492

Rodjhun B. Navarro, LPT, PhD

Vice-President for Research, Innovations, and Linkages, La Consolacion College Bacolod
<https://orcid.org/0000-0001-7889-5770> | rnavarro@lccbonline.edu.ph

Rejay Y. Alvarado, LPT, MPM

Head, Research and Creative Works Office, La Consolacion College Bacolod
<https://orcid.org/0000-0003-2662-5749> | rejayalvarado@lccbonline.edu.ph

Mikhail Ivan A. Muchuelas

Research, Innovations, and Linkages Staff, La Consolacion College Bacolod
<https://orcid.org/0009-0006-2959-4414> | mikhail.ivan.muchuelas@lccbonline.edu.ph

Minnette P. Gela, LPT, MAED

Subject Area Chairperson, Filipino, IBED La Consolacion College Bacolod
<https://orcid.org/0009-0007-1335-7884> | minnettegela@lccbonline.edu.ph

Jennifer M. Divinagracia, LPT

Integrated School Faculty, IBED La Consolacion College Bacolod
<https://orcid.org/0009-0008-5435-0539> | jenniferdivinagracia@lccbonline.edu.ph

Abstract

This study evaluated a tutorial support program in an integrated basic education department through a secondary mixed-methods analysis of institutional evaluation data collected in November 2025. The analysis examined personnel perceptions of five institutionally defined areas: program implementation, implementation and delivery, perceived student progress and outcomes, tutor support and development, and program management and overall satisfaction. The dataset comprised 35 complete responses from tutorial teachers, administrators, referring subject teachers, and one respondent with multiple roles. Quantitative responses to 25 Likert-type items were analyzed using descriptive statistics, 95% confidence intervals, internal consistency estimates, and response-concentration analysis. Brief open-ended responses were examined through descriptive qualitative content analysis and integrated with the quantitative results. The instrument showed very high internal consistency overall ($\alpha = .981$), with coefficients ranging from .894 to .974 across the five areas. Respondents generally evaluated the program favorably ($M = 4.32$, $SD = 0.63$). Implementation and Delivery received the highest mean ($M = 4.43$, $SD = 0.59$), whereas Tutor Support and Development received the lowest ($M = 4.25$, $SD = 0.74$). Regular conduct of sessions was the highest-rated practice, while tutor training and guidance received the lowest item mean. Written responses identified focused instructional assistance and remediation as perceived strengths but also raised concerns regarding tutor preparation, monitoring, learner dependence, eligibility policies, subject matching, workload, scheduling, and venue adequacy. The concentration of responses at the upper end of the scale and the absence of direct student outcome measures require cautious interpretation. The findings support continuation of the program with stronger tutor preparation, clearer operational policies, standardized monitoring, and future evaluation using direct evidence of student participation and learning.

Keywords: tutorial support, program evaluation, mixed methods, academic support, basic education, institutional research

Introduction

Educational systems continue to confront persistent differences between the competencies expected of learners and the skills that many students actually demonstrate. Although these gaps predated the COVID-19 pandemic, prolonged disruptions to schooling intensified them and exposed longstanding inequalities in access to effective



instruction, individualized assistance, and supportive learning environments. International monitoring shows that achievement declined sharply across many education systems between 2018 and 2022, with the Organisation for Economic Co-operation and Development reporting an unprecedented average decrease of 15 points in mathematics and 10 points in reading among participating OECD countries. Global learning-recovery frameworks consequently emphasize identifying learners' current achievement levels, prioritizing foundational competencies, adapting instruction to demonstrated needs, and providing additional support for those who have fallen behind (OECD, 2023a; World Bank et al., 2022).

The challenge is particularly consequential in the Philippines. In the 2022 Programme for International Student Assessment, only 16% of Filipino 15-year-olds reached at least the baseline proficiency level in mathematics, while 24% reached the baseline in reading and 23% in science. These proportions were considerably below the corresponding OECD averages of 69%, 74%, and 76%. Philippine performance in the three domains also remained statistically similar to the country's 2018 results, indicating that the learning difficulties evident before and during the pandemic had not been substantially reversed by 2022 (OECD, 2023b). Although international assessments do not capture every operational area of school quality or the circumstances of individual institutions, they underscore the urgency of providing timely and targeted assistance to learners who have not mastered grade-level competencies.

One increasingly prominent response to this need is tutorial support, broadly understood as supplementary instruction provided to individual learners or small groups to address identified difficulties, reinforce essential competencies, or extend opportunities for guided practice. Tutorial support may be delivered by classroom teachers, designated tutors, paraprofessionals, peers, volunteers, or technology-mediated systems. It may occur during the school day, after classes, during weekends or vacation periods, and through face-to-face, telephone-based, or online arrangements. This diversity makes tutoring adaptable, but it also means that the term encompasses interventions that differ substantially in frequency, duration, tutor qualifications, group size, instructional focus, and degree of curricular alignment. Tutoring should therefore not be treated as a single, uniform intervention whose effects can be assumed regardless of how it is designed and implemented.

The empirical case for well-designed tutoring is nevertheless strong. In a systematic review and meta-analysis of experimental studies involving prekindergarten to Grade 12 learners, Nickow et al. (2024) found a pooled positive effect of approximately 0.29 standard deviations on learning outcomes. Effects tended to be larger when tutoring was conducted by teachers or paraprofessionals, provided in the earlier grades, scheduled at least three times per week, and integrated into the school day. Earlier syntheses similarly found positive effects for tutoring in literacy, mathematics, and other areas, particularly when instruction was closely connected to learners' assessed needs and the regular curriculum (Ritter et al., 2009; Slavin et al., 2011). These findings establish tutoring as a promising academic-support strategy, but they also show that its potential depends on specific design and delivery conditions rather than on the mere existence of a tutorial program.

Recent guidance has consequently distinguished ordinary or occasional tutoring from high-impact or high-dosage tutoring. High-impact tutoring generally involves frequent sessions, sustained participation over several weeks or months, small tutor-to-student ratios, consistent tutor-learner relationships, use of student performance data, alignment with classroom instruction, and structured materials or routines. Evidence-informed design principles suggest that effective programs commonly provide at least three sessions per week over a minimum period of approximately 10 weeks, although the appropriate duration and session length may vary by learners' age, subject, and level of need (Robinson et al., 2024). Such specifications are important because low-frequency, loosely organized, or primarily homework-completion arrangements may be labeled "tutoring" without providing the intensity or instructional coherence associated with stronger outcomes.

Tutoring also has an equity operational area. Properly targeted support can provide additional learning time to students who have had fewer opportunities to master foundational knowledge, including learners from disadvantaged backgrounds, those at risk of failure or dropping out, and those requiring more individualized pacing. OECD reviews identify tutoring and mentoring as school-level strategies that can promote achievement, participation, belonging, and educational inclusion when directed toward learners who need them most (OECD, 2023c; OECD, 2025). Yet the same programs may reproduce inequity when access depends on parental resources, informal teacher discretion, student self-selection, or the capacity to attend sessions outside regular school hours. Questions about who is



selected, who participates consistently, and whose needs receive priority are therefore central to evaluating tutorial support.

This tension between potential and implementation is increasingly relevant in Philippine education policy. Republic Act No. 12028, or the Academic Recovery and Accessible Learning Program Act, established a national intervention intended to help struggling learners attain required competencies, particularly in reading, mathematics, and science. The law institutionalized tutorial support as part of the country's response to learning gaps, while subsequent Department of Education initiatives expanded remediation, learning camps, literacy programs, and targeted forms of tutoring. These developments signal that tutoring is no longer merely an optional school activity; it is becoming an important component of national learning-recovery and acceleration efforts. At the same time, national authorization does not by itself ensure effective implementation at the school level, where tutor preparation, scheduling, instructional materials, learner identification, supervision, and resource availability ultimately shape program quality (Republic Act No. 12028, 2024).

Emerging Philippine initiatives also illustrate the variety of forms that tutorial support can take. For example, a Department of Education partnership implemented telephone-based mathematics tutoring for Grades 3 to 5 learners, combining weekly instructional calls with messages and involving more than 1,000 learners and 265 teachers across 90 schools during its initial 2024 implementation (Department of Education, 2024a). The expansion of the *Tara, Basa!* Tutoring Program similarly reflects growing institutional interest in structured literacy assistance for struggling learners (Department of Education, 2024b). These initiatives demonstrate the feasibility of personalized and targeted support, including through alternative delivery modes. However, they also reinforce the need to examine whether program models remain responsive to local contexts, whether tutors have the necessary preparation, and whether institutional systems can sustain implementation beyond initial or externally supported cycles.

A central issue is implementation quality, referring to the extent to which a program is delivered in a manner consistent with its intended objectives, essential components, and operating procedures. Implementation quality may include adherence to schedules, appropriateness of learner selection, tutor competence, instructional dosage, availability of materials, coordination among personnel, monitoring of sessions, and responsiveness to contextual constraints. Research on educational and social interventions has repeatedly shown that programs with credible theoretical or empirical foundations may produce weak or inconsistent results when implementation is fragmented, under-resourced, or poorly monitored (Durlak & DuPre, 2008). Conversely, implementation should not be understood as rigid compliance alone. Schools often need to adapt interventions to learners' ages, staffing patterns, schedules, and available resources. The evaluative challenge is therefore to distinguish constructive adaptation from changes that weaken the program's essential instructional mechanisms.

Tutor support and development are especially important within this implementation process. Tutors are expected not only to possess content knowledge but also to diagnose misconceptions, scaffold difficult tasks, adjust explanations, provide feedback, monitor understanding, and gradually promote learner independence. Without clear guidance, tutoring may become an extension of ordinary teaching, an unsystematic review session, or a venue for completing assignments on behalf of students. Excessive assistance may produce short-term task completion while inadvertently encouraging dependency and limiting the development of self-regulated learning. At the same time, demanding individualized instruction without adequate preparation, manageable workloads, suitable spaces, instructional materials, and administrative support places unrealistic expectations on tutors. Program evaluation must therefore attend to both tutor accountability and the institutional conditions that enable tutors to perform their roles effectively.

The assessment of tutorial outcomes presents another conceptual and methodological difficulty. Academic support programs are often considered successful when teachers observe better grades, participation, confidence, motivation, study habits, or classroom behavior. These perceptions are valuable because teachers and administrators directly observe learners across instructional settings and may identify changes that standardized measures overlook. However, perceived outcomes are not equivalent to objectively verified effects. Favorable ratings may be influenced by respondents' involvement in the program, institutional loyalty, social desirability, expectations of improvement, or limited access to comparative data. Strong claims about effectiveness require direct evidence such as pre-intervention and post-intervention achievement measures, attendance and completion records, comparison conditions, or longitudinal student data. In their absence, evaluation findings should be presented as stakeholder judgments about implementation and perceived contribution rather than as proof of causal impact.



This distinction highlights the role of program evaluation as both an accountability and organizational-learning process. Program evaluation is the systematic collection and interpretation of evidence concerning a program's design, implementation, outcomes, value, and possibilities for improvement (Fitzpatrick et al., 2011; Stufflebeam & Coryn, 2014). Unlike research that is primarily concerned with testing a universally applicable hypothesis, institutional evaluation may be undertaken to support decisions about continuation, modification, resource allocation, staff development, or quality assurance. Its value lies not only in determining whether stakeholders are satisfied but also in identifying why particular components function well, where implementation breaks down, and what feasible changes can be introduced. Evaluation is most useful when its findings are sufficiently rigorous to be credible and sufficiently contextualized to guide action.

Institutional surveys are commonly used for this purpose because they can capture the experiences of personnel who plan, deliver, coordinate, and monitor programs. Quantitative ratings can identify broad patterns across program operational areas, whereas open-ended responses can reveal concerns that are concealed by favorable averages. Yet institutional evaluation surveys also have limitations. Small populations restrict statistical power; researcher-developed instruments may lack extensive validation; respondents may cluster their answers toward the positive end of rating scales; and employees may hesitate to criticize an institutional program directly. These concerns do not make the data unusable, but they require transparent reporting, cautious interpretation, examination of item-level variation, and the integration of qualitative comments that may explain or qualify the numerical results.

The relationship between quantitative and qualitative evidence is particularly important when evaluations produce high overall ratings alongside concrete recommendations for change. A program may be considered valuable and functional while still exhibiting weaknesses in training, monitoring, student selection, scheduling, or logistical support. Treating favorable means as evidence that the program requires little improvement would ignore the diagnostic purpose of institutional evaluation. Conversely, emphasizing isolated complaints without considering the generally positive assessment would misrepresent stakeholders' overall judgment. A mixed-methods approach can address this tension by examining where the two forms of evidence converge, where qualitative responses explain lower-rated areas, and where narrative concerns complicate apparently strong quantitative results.

Despite substantial evidence on the average effects of tutoring, much of the literature has emphasized experimental, externally funded, or highly structured interventions in which student achievement is the principal outcome. Less attention has been given to the everyday implementation of routinely administered school-based tutorial programs, particularly to how teachers and administrators experience their scheduling, learner selection, tutor preparation, monitoring, logistical support, and management. This gap is particularly relevant to institutional research because interventions that perform well under controlled conditions must still be translated into workable arrangements within actual schools. Examining routine implementation can identify whether a program possesses the organizational conditions needed to sustain academic support, even when direct causal evidence of student impact is not yet available.

Within the institution examined in this study, the Tutorial Program operates as a supplementary academic-support mechanism for learners identified as needing additional assistance beyond regular classroom instruction. Students may be referred by subject teachers or other authorized school personnel, while tutorial teachers provide focused assistance in identified learning areas. The program is intended primarily to address learning gaps, reinforce essential competencies, and support learners who experience difficulty meeting classroom expectations. However, the available institutional dataset did not contain complete records of tutorial dosage, group size, attendance, duration of participation, or the specific criteria used for entry and completion. The present evaluation therefore focused on how personnel involved in the program perceived its implementation and operation rather than on whether the program met established high-dosage tutoring standards.

Against this background, the present study conducted a secondary mixed-methods analysis of an internal evaluation of the Tutorial Program of an Integrated Basic Education Department undertaken in November 2025. It examined personnel perceptions of five institutionally defined evaluation areas: program implementation, implementation and delivery, perceived student progress and outcomes, tutor support and development, and program management and overall satisfaction. It also analyzed brief written responses concerning successful practices, implementation concerns, and recommendations for subsequent program cycles. Stakeholder ratings were treated as evidence of how the program was experienced by personnel involved in its delivery and oversight, not as direct proof of student



impact. By integrating quantitative patterns with descriptive qualitative evidence, the study sought to identify operational strengths, clarify institutional improvement priorities, and demonstrate how routinely collected evaluation data may be used to support evidence-informed program development.

Research Questions

This study aimed to evaluate the Tutorial Program of the Integrated Basic Education Department using secondary data collected through an internal program evaluation survey in November 2025. Specifically, it sought to answer the following questions:

1. How do personnel involved in the Tutorial Program evaluate the program in terms of:
 - 1.1 program implementation;
 - 1.2 implementation and delivery;
 - 1.3 perceived student progress and outcomes;
 - 1.4 tutor support and development; and
 - 1.5 program management and overall satisfaction?
2. Which specific aspects of the Tutorial Program are perceived as its strongest features and its principal areas for improvement?
3. What descriptive categories emerge from the respondents' open-ended accounts regarding:
 - 3.1 the most successful aspects of the program;
 - 3.2 the areas requiring improvement; and
 - 3.3 the recommendations for the next implementation cycle?
4. How do the qualitative findings explain, support, or qualify the quantitative evaluation results?
5. What exploratory descriptive variations are observed in respondents' evaluations of the Tutorial Program according to:
 - 5.1 role in the program; and
 - 5.2 length of involvement in the program?

Because the role and experience categories were small and unequal, the fifth research question was treated as exploratory and was not used to test statistically significant group differences.

Methodology

Research Design

The study employed a secondary mixed-methods program evaluation design to examine the implementation quality, perceived outcomes, and improvement priorities of the Tutorial Program of the Integrated Basic Education Department. The dataset was originally generated through an internal evaluation conducted in November 2025 and was subsequently analyzed for research and institutional learning purposes. The study was considered secondary research because the researchers analyzed data that had already been collected for routine program monitoring rather than gathering new responses specifically for the present investigation. A convergent mixed-methods orientation was adopted because the institutional survey contained both numerical ratings and open-ended responses collected during the same administration. The quantitative and qualitative data were analyzed separately and then integrated during interpretation to determine where the findings converged, complemented one another, or revealed tensions between generally favorable ratings and specific implementation concerns. This approach was appropriate because mixed-methods program evaluation can provide both an overall assessment of program functioning and contextual explanations for observed patterns (Creswell & Plano Clark, 2018; Fitzpatrick et al., 2011).

The study was descriptive and utilization-focused rather than experimental. It did not seek to establish that participation in the tutorial program caused changes in student achievement, behavior, or motivation. Instead, it examined how the personnel involved in the delivery and oversight of the program perceived its implementation and contributions. Consequently, ratings relating to academic performance, confidence, failure rates, study habits, and classroom behavior were treated as perceived student outcomes, not as direct measures of program impact. This distinction was necessary because the dataset did not include pre-intervention and post-intervention achievement scores, comparison groups, student-level academic records, or longitudinal outcome measures.



Research Locale and Program Context

The evaluation was conducted in the Integrated Basic Education Department of La Consolacion College Bacolod, a private educational institution in Bacolod City, Philippines. The department implements a Tutorial Program as part of its academic-support and continuous-improvement processes. The program is intended to provide additional instructional assistance to learners requiring remediation, reinforcement of competencies, or focused guidance beyond regular classroom instruction. Personnel involved in its operation include tutorial teachers who provide direct assistance, subject teachers who identify or refer learners, and administrators who oversee implementation.

The available secondary dataset documented personnel perceptions of the program but did not include complete information on the number of participating students, frequency and duration of sessions, tutor-to-student ratios, attendance, subject-specific dosage, or student-level academic records. Consequently, the study evaluated the perceived implementation and management of the program rather than its compliance with a particular high-dosage tutoring model or its causal effect on student achievement.

No student names, grades, diagnostic scores, or other direct student-level identifiers were included in the analytical dataset.

Respondents, Sample Size, and Data-Inclusion Technique

The secondary dataset contained 35 complete survey records from personnel involved in the Tutorial Program. The respondents consisted of 23 tutorial teachers, six administrators, five subject teachers who had referred students to the program, and one respondent who indicated multiple roles as a tutorial teacher, referring subject teacher, and administrator. In terms of length of involvement, 13 respondents had participated for less than one semester, 10 for one semester, four for two to three semesters, and eight for more than three semesters.

All 35 completed survey records available in the institutional dataset were included in the secondary analysis rather than selecting a sample from the existing records. Because the available records represented completed responses rather than independently verified documentation of the entire eligible staff population, the study did not assume that the 35 respondents constituted a statistical census of all personnel who could have participated. Claims of representativeness were therefore limited to the respondents included in the November 2025 evaluation. No completed response was excluded from the quantitative analysis because all records contained responses to the 25 rating-scale items. All 35 respondents also provided entries to the three open-ended questions, although the depth and substantive content of their answers varied.

For the principal analysis, the responses were examined collectively because the study focused on the institutional evaluation of the program. Role and length-of-involvement classifications were used only for exploratory descriptive summaries. No strong inferential conclusions were drawn from subgroup comparisons because the categories were small, unequal in size, and included one respondent who selected multiple roles.

Research Instrument

The data were collected through the researcher-developed IBED Tutorial Program Evaluation Survey—Teacher Version, administered electronically through Google Forms. The instrument consisted of respondent-profile questions, 25 closed-ended evaluation statements, and three open-ended questions. The closed-ended items were organized into five intended operational areas, with five items assigned to each operational area:

1. Program Implementation, which examined the clarity of objectives, student-selection criteria, coordination, scheduling, and administrative orientation;
2. Implementation and Delivery, which assessed the regularity of tutorial sessions, relevance of learning materials, responsiveness to learning gaps, learning environment, and program monitoring;
3. Perceived Student Progress and Outcomes, which covered perceived changes in academic performance, confidence, class participation, failure or incomplete grades, study habits, discipline, and motivation;
4. Tutor Support and Development, which examined tutor training, workload, evaluation, collaboration, and administrative or logistical support; and



5. Program Management and Overall Satisfaction, which assessed organization, communication, responsiveness to academic and behavioral concerns, personalized learning support, and general satisfaction.

Respondents rated each statement using a five-point Likert-type scale: 1—Strongly Disagree, 2—Disagree, 3—Neutral, 4—Agree, and 5—Strongly Agree. For the original institutional evaluation, composite means were interpreted as follows: 4.50–5.00, Highly Evident; 3.50–4.49, Evident; 2.50–3.49, Moderately Evident; 1.50–2.49, Slightly Evident; and 1.00–1.49, Not Evident. These categories were retained only as institutional descriptors and were not treated as empirically validated thresholds.

Although individual Likert-type items are ordinal, means and standard deviations were used cautiously to summarize the institutionally grouped items and were interpreted alongside response distributions and qualitative evidence (Boone & Boone, 2012).

The qualitative portion contained three questions asking respondents to identify: (a) the most successful aspects of the program, (b) areas requiring improvement, and (c) recommendations for the next tutorial cycle. These questions were intended to contextualize the numerical ratings and identify operational issues that might not be adequately captured through fixed-response items.

Validity and Reliability of the Instrument

The instrument was developed for institutional evaluation based on the objectives and operational components of the Tutorial Program and on literature concerning tutorial support, implementation, perceived student outcomes, tutor development, and program management. Its 25 items were organized into five intended operational areas corresponding to the aspects the institution sought to evaluate. However, the organization of items into these areas did not by itself establish content or construct validity. Documentary records available for the secondary analysis did not show that the instrument had undergone an independent expert content-validity procedure, pilot testing, factor analysis, or prior construct-validation study. The results were therefore interpreted as findings from an institutionally developed evaluation instrument rather than from a standardized or fully validated measurement scale.

Internal consistency was examined using Cronbach's alpha. The complete 25-item instrument demonstrated very high internal consistency, $\alpha = .981$. The five intended operational areas likewise showed strong internal consistency: Program Implementation, $\alpha = .894$; Implementation and Delivery, $\alpha = .906$; Perceived Student Progress and Outcomes, $\alpha = .974$; Tutor Support and Development, $\alpha = .931$; and Program Management and Overall Satisfaction, $\alpha = .971$. These coefficients indicate that the items within each operational area were answered consistently. However, exceptionally high alpha coefficients may also suggest substantial item overlap or redundancy rather than stronger validity. The coefficients should therefore be interpreted as evidence of internal consistency, not as proof of construct validity (Tavakol & Dennick, 2011). Factor analysis was not conducted because the sample of 35 was insufficient for a defensible assessment of the instrument's latent structure. Accordingly, the reported coefficients indicate response consistency within the institutionally defined item groupings but do not confirm that the five areas represent distinct psychometric constructs.

Data-Gathering and Data-Preparation Procedures

The original institutional evaluation was administered through Google Forms in November 2025. Personnel involved in the Tutorial Program were invited to complete the online questionnaire based on their experience with the program. The form collected role information, length of involvement, responses to the 25 Likert-type statements, and comments concerning program strengths, weaknesses, and recommendations. The resulting spreadsheet was retained as part of the institution's internal program-evaluation records.

For the secondary analysis, the Google Forms spreadsheet was exported in Microsoft Excel format and reviewed for completeness and consistency. Direct identifiers, including names and institutional email addresses, were separated from the analytical file and were not used in computing results or interpreting qualitative responses. The researchers checked the dataset for duplicate records, missing quantitative responses, values outside the intended one-to-five range, inconsistent category labels, and responses that could inadvertently identify individual personnel or students.



Minor inconsistencies in category wording, such as the spelling of one-semester involvement, were standardized without changing the substantive meaning of the responses.

The quantitative and qualitative portions were then separated for analysis. Composite scores for each operational area were computed as the mean of the five items assigned to each intended operational area. The overall evaluation score was computed as the mean of all 25 quantitative items. Open-ended responses were de-identified before coding. The two forms of evidence were subsequently brought together through a joint interpretation process in which lower- and higher-rated items were compared with the descriptive categories identified in respondents' written comments.

Statistical Treatment of Quantitative Data

Frequencies and percentages were used to summarize respondents' roles and length of involvement in the Tutorial Program. For each quantitative item, the mean, standard deviation, response distribution, and verbal interpretation were examined. Composite means and standard deviations were calculated for the five operational areas and for the entire instrument. Where appropriate, 95% confidence intervals were computed to communicate the precision of the estimated means. Items were ranked to identify the most favorably evaluated program practices and the comparatively lower-rated areas that could serve as institutional improvement priorities.

Cronbach's alpha was calculated for the overall instrument and each five-item operational area. The frequency of maximum ratings and the proportion of respondents attaining the highest possible composite scores were also examined to assess upper-end response concentration. This analysis was important because a strong concentration of responses at the upper end of the scale could restrict variability and make numerical differences appear smaller than the qualitative concerns expressed by respondents.

Descriptive summaries were also generated according to role and length of involvement. These comparisons were treated as exploratory and were not used to establish statistically significant differences between groups. Inferential tests, multiple regression, factor analysis, and structural equation modeling were not employed because the small and unequally distributed respondent groups, restricted score variation, and highly correlated operational-area scores would make such analyses unstable and potentially misleading.

The open-ended responses were analyzed using descriptive qualitative content analysis, which was considered appropriate because most entries were brief survey comments rather than extended qualitative narratives. Responses consisting solely of "none," "not applicable," "yes," punctuation, or equivalent non-substantive entries were retained in the dataset but excluded from coding. The substantive responses were read repeatedly and segmented into meaningful units. Initial codes were assigned to statements concerning program contributions, implementation concerns, and recommendations. Multiple codes were permitted when a response addressed more than one program issue. Related codes were then grouped into broader descriptive categories, consistent with established procedures for qualitative content analysis (Elo & Kyngäs, 2008; Schreier, 2012).

Coding was conducted by one researcher and subsequently reviewed by a second member of the research team for clarity, consistency, and agreement with the original responses. A coding matrix was maintained to document the relationship among original responses, assigned codes, descriptive categories, and representative excerpts. Because the responses were generally short, the analysis was intended to summarize recurring content rather than generate an interpretive theory or claim qualitative saturation.

Mixed-methods integration occurred after the quantitative and qualitative analyses had been completed. Quantitative results were compared with the descriptive qualitative categories to determine whether the written responses confirmed, explained, or qualified the numerical patterns. Integration was presented through a joint display linking selected quantitative findings, corresponding qualitative evidence, and their institutional implications. The integration did not treat the qualitative responses as independent confirmation of program effectiveness; rather, it used them to clarify how respondents interpreted strengths and concerns reflected in the survey ratings.

Ethical Considerations



The original institutional evaluation was reviewed by the Institutional Ethics Review Board of La Consolacion College Bacolod under Protocol No. 202526-118-1 dated November 22, 2025. Before the dataset was used for secondary research and publication, authorization for the secondary use of the de-identified data was obtained from the Institutional Ethics Review Board. The final analysis was conducted only after this authorization had been confirmed.

The analysis followed the principles of confidentiality, data minimization, and restricted access. Names, email addresses, and other direct identifiers were removed from the working dataset before analysis. Respondents were represented by non-identifying codes, and findings were reported in aggregate form. Any excerpts from the open-ended responses were screened and minimally edited where necessary to remove details that could identify a respondent, learner, class, or subject assignment. The de-identified dataset was stored in a password-protected institutional location accessible only to authorized researchers. Because the student-outcome items reflected personnel perceptions rather than direct student data, the manuscript consistently avoided presenting these ratings as verified evidence of student achievement or causal program impact.

Results

Respondent Characteristics

The secondary dataset contained 35 complete responses from personnel involved in the Integrated Basic Education Department Tutorial Program. Most respondents were tutorial teachers ($n = 23$, 65.7%), followed by administrators ($n = 6$, 17.1%) and subject teachers who referred students to the program ($n = 5$, 14.3%). One respondent (2.9%) reported multiple roles as a tutorial teacher, referring subject teacher, and administrator.

In terms of program experience, 13 respondents (37.1%) had been involved for less than one semester, 10 (28.6%) for one semester, four (11.4%) for two to three semesters, and eight (22.9%) for more than three semesters. All respondents completed the 25 quantitative items, resulting in no missing values in the rating-scale data.

Table 1

Profile of the Respondents ($N = 35$)

Characteristic	<i>n</i>	%
Role in the Tutorial Program		
Tutorial teacher	23	65.7
Administrator	6	17.1
Subject teacher who referred students	5	14.3
Multiple roles	1	2.9
Length of involvement		
Less than one semester	13	37.1
One semester	10	28.6
Two to three semesters	4	11.4
More than three semesters	8	22.9

Internal Consistency of the Instrument

The 25-item evaluation instrument demonstrated very high internal consistency, with an overall Cronbach's alpha of .981. The five intended operational areas also showed strong internal consistency, with coefficients ranging from .894 for Program Implementation to .974 for Perceived Student Progress and Outcomes.



These coefficients indicate that the items within each operational area were answered consistently. However, the very high coefficients, particularly for the overall instrument and several operational areas, may also reflect substantial similarity among items or response homogeneity. The coefficients were therefore interpreted as evidence of internal consistency rather than proof of construct validity.

Table 2
Internal Consistency of the Tutorial Program Evaluation Instrument

Operational Area	Number of items	Cronbach's α
Program Implementation	5	.894
Implementation and Delivery	5	.906
Perceived Student Progress and Outcomes	5	.974
Tutor Support and Development	5	.931
Program Management and Overall Satisfaction	5	.971
Overall instrument	25	.981

Overall Evaluation of the Tutorial Program

The Tutorial Program received a generally favorable evaluation, with an overall mean of 4.32 (SD = 0.63, 95% CI [4.11, 4.54]). Based on the institution's predetermined interpretation scheme, this value corresponded to the "Evident" category. Because the category labels were developed for internal evaluation, the present analysis prioritized the reported means, standard deviations, confidence intervals, and response distributions in interpreting the findings.

Implementation and Delivery received the highest composite mean ($M = 4.43$, $SD = 0.59$), followed by Program Implementation ($M = 4.40$, $SD = 0.60$). Perceived Student Progress and Outcomes obtained a mean of 4.29 ($SD = 0.69$), while Program Management and Overall Satisfaction obtained a mean of 4.26 ($SD = 0.78$). Tutor Support and Development had the lowest composite mean ($M = 4.25$, $SD = 0.74$), although it remained within the Evident range.

The 95% confidence intervals were calculated from respondent-level composite scores using the t distribution with 34 degrees of freedom. The intervals reflect the limited precision associated with the sample of 35 respondents. The rankings were interpreted descriptively because no inferential tests of differences among operational-area means were conducted and the confidence intervals showed considerable overlap.

Table 3
Overall Evaluation of the Tutorial Program by Operational Area

Operational Area	M	SD	95% CI	Institutional Interpretation	Rank
Implementation and Delivery	4.43	0.59	[4.23, 4.63]	Evident	1
Program Implementation	4.40	0.60	[4.19, 4.61]	Evident	2
Perceived Student Progress and Outcomes	4.29	0.69	[4.05, 4.52]	Evident	3
Program Management and Overall Satisfaction	4.26	0.78	[3.99, 4.52]	Evident	4
Tutor Support and Development	4.25	0.74	[3.99, 4.50]	Evident	5
Overall	4.32	0.63	[4.11, 4.54]	Evident	—



Note. The verbal categories were predetermined for institutional evaluation and should not be interpreted as empirically validated thresholds.

Program Implementation

Program Implementation obtained a composite mean of 4.40 ($SD = 0.60$), classified as Evident under the institutional interpretation scheme. The fairness and appropriateness of the criteria for selecting students obtained the highest mean in this operational area ($M = 4.51$, $SD = 0.66$), interpreted as Highly Evident. Twenty-one respondents, or 60.0%, selected Strongly Agree for this item.

The clarity and alignment of the program objectives with the IBED mission received a mean of 4.46 ($SD = 0.74$). Scheduling and grouping of tutorial sessions obtained a mean of 4.43 ($SD = 0.65$), while coordination between subject teachers and tutorial teachers obtained a mean of 4.37 ($SD = 0.73$).

Administrative orientation and guidance before implementation received the lowest mean in the operational area ($M = 4.23$, $SD = 0.81$), although it remained within the Evident category.

Table 4
Evaluation of Program Implementation

Indicator	<i>M</i>	<i>SD</i>	Institutional Interpretation
Selection criteria were fair and appropriate	4.51	0.66	Highly Evident
Objectives were clear and aligned with the IBED mission	4.46	0.74	Evident
Scheduling and grouping were efficient	4.43	0.65	Evident
Coordination between subject and tutorial teachers was effective	4.37	0.73	Evident
Administrators provided adequate orientation and guidance	4.23	0.81	Evident
Composite mean	4.40	0.60	Evident

Implementation and Delivery

Implementation and Delivery was the highest-rated operational area ($M = 4.43$, $SD = 0.59$). The regular conduct of tutorial sessions and adherence to the assigned schedule obtained the highest item mean in the entire instrument ($M = 4.63$, $SD = 0.55$), interpreted as Highly Evident. Twenty-three respondents, or 65.7%, selected Strongly Agree for this item.

The tutorial environment's promotion of focus and engagement obtained a mean of 4.46 ($SD = 0.74$). The relevance and suitability of learning materials and the extent to which tutorial sessions addressed students' learning gaps both obtained means of 4.43.

Monitoring and feedback from program coordinators received the lowest mean in this operational area ($M = 4.20$, $SD = 0.76$). Seven respondents selected Neutral, while 28 selected either Agree or Strongly Agree.

Table 5
Evaluation of Implementation and Delivery

Indicator	<i>M</i>	<i>SD</i>	Institutional Interpretation
Sessions were conducted regularly and followed the schedule	4.63	0.55	Highly Evident



Indicator	M	SD	Institutional Interpretation
Tutorial environment promoted focus and engagement	4.46	0.74	Evident
Learning materials were relevant and suitable	4.43	0.74	Evident
Sessions addressed students' learning gaps	4.43	0.65	Evident
Monitoring and feedback from coordinators were adequate	4.20	0.76	Evident
Composite mean	4.43	0.59	Evident

Perceived Student Progress and Outcomes

The Perceived Student Progress and Outcomes operational area obtained a composite mean of 4.29 ($SD = 0.69$), classified as Evident under the institutional interpretation scheme. The highest-rated indicator was the perception that tutorial sessions helped reduce failure rates or incomplete grades ($M = 4.37$, $SD = 0.65$). The perceived contribution of the program to classroom discipline and motivation followed with a mean of 4.34 ($SD = 0.68$).

The program's perceived contribution to effective study habits obtained a mean of 4.29 ($SD = 0.75$), while noticeable academic improvement obtained a mean of 4.23 ($SD = 0.77$). Better confidence and class participation received the lowest mean in the operational area ($M = 4.20$, $SD = 0.76$).

These results represent respondents' perceptions of student progress. The dataset did not include student achievement records, pretest-posttest measures, or comparison data; therefore, the findings were not interpreted as direct evidence of program impact.

Table 6

Evaluation of Perceived Student Progress and Outcomes

Indicator	M	SD	Institutional Interpretation
Sessions helped reduce failure rates or incomplete grades	4.37	0.65	Evident
Program improved classroom discipline and motivation	4.34	0.68	Evident
Program encouraged effective study habits	4.29	0.75	Evident
Students showed noticeable academic improvement	4.23	0.77	Evident
Students demonstrated better confidence and participation	4.20	0.76	Evident
Composite mean	4.29	0.69	Evident

Tutor Support and Development

Tutor Support and Development obtained the lowest operational area mean ($M = 4.25$, $SD = 0.74$), although it remained within the Evident category. The manageability of tutors' workload and assigned number of students obtained the highest mean in this operational area ($M = 4.37$, $SD = 0.81$). Opportunities for collaboration among tutors and other teachers obtained a mean of 4.34 ($SD = 0.73$).

The availability of administrative and logistical support and the fairness of tutor feedback and evaluation both obtained means of 4.26. Sufficient training and guidance for tutorial teachers obtained the lowest mean among all 25 items ($M = 4.00$, $SD = 0.94$). For this item, one respondent selected Strongly Disagree, nine selected Neutral, 13 selected Agree, and 12 selected Strongly Agree.



The comparatively lower mean and larger standard deviation indicate greater variation in respondents' experiences of tutor preparation.

Table 7

Evaluation of Tutor Support and Development

Indicator	<i>M</i>	<i>SD</i>	Institutional Interpretation
Workload and number of assigned students were manageable	4.37	0.81	Evident
Opportunities for collaboration were adequate	4.34	0.73	Evident
Feedback and evaluation of tutors were fair	4.26	0.85	Evident
Administrative and logistical support was available	4.26	0.82	Evident
Tutorial teachers received sufficient training and guidance	4.00	0.94	Evident
Composite mean	4.25	0.74	Evident

Program Management and Overall Satisfaction

Program Management and Overall Satisfaction obtained a composite mean of 4.26 ($SD = 0.78$), classified as Evident under the institutional interpretation scheme. The program's contribution to the school's philosophy of personalized learning support received the highest mean in this operational area ($M = 4.34$, $SD = 0.76$). Communication among teachers, parents, and administrators followed with a mean of 4.29 ($SD = 0.79$).

The organization and efficiency of program implementation obtained a mean of 4.26 ($SD = 0.78$), while the program's perceived capacity to address academic and behavioral concerns obtained a mean of 4.23 ($SD = 0.91$).

Overall satisfaction with program outcomes and management obtained the lowest mean in the operational area ($M = 4.17$, $SD = 0.86$). Twenty-seven respondents selected either Agree or Strongly Agree. Eight respondents did not express agreement: seven selected Neutral and one selected Disagree.

Table 8

Evaluation of Program Management and Overall Satisfaction

Indicator	<i>M</i>	<i>SD</i>	Institutional Interpretation
Program reinforced personalized learning support	4.34	0.76	Evident
Communication was open and constructive	4.29	0.79	Evident
Program was well-organized and efficiently implemented	4.26	0.78	Evident
Program addressed academic and behavioral concerns	4.23	0.91	Evident
Respondents were satisfied with outcomes and management	4.17	0.86	Evident
Composite mean	4.26	0.78	Evident

Highest- and Lowest-Rated Indicators

The item-level analysis identified the regular conduct of tutorial sessions as the highest-rated indicator, followed by the fairness and appropriateness of student-selection criteria. Clear objectives and an engaging tutorial environment were tied at a mean of 4.46.



The lowest-rated indicator was tutor training and guidance. Overall satisfaction ranked second lowest, while monitoring and feedback and students' confidence and participation were tied at a mean of 4.20.

Table 9

Selected Highest- and Lowest-Rated Indicators

Indicator	<i>M</i>	<i>SD</i>	Relative position
Sessions were conducted regularly and followed the schedule	4.63	0.55	Highest
Student-selection criteria were fair and appropriate	4.51	0.66	Second highest
Objectives were clear and aligned with the IBED mission	4.46	0.74	Tied among highest
Tutorial environment promoted focus and engagement	4.46	0.74	Tied among highest
Tutorial teachers received sufficient training and guidance	4.00	0.94	Lowest
Overall satisfaction with outcomes and management	4.17	0.86	Second lowest
Monitoring and feedback from coordinators were adequate	4.20	0.76	Tied among lowest
Students demonstrated better confidence and participation	4.20	0.76	Tied among lowest
Administrators provided adequate orientation and guidance	4.23	0.81	Among lowest

Upper-End Response Concentration

Responses were strongly concentrated at the upper end of the five-point scale. Fourteen respondents (40.0%) obtained the maximum composite score of 5.00 for Program Implementation, Implementation and Delivery, and Perceived Student Progress and Outcomes. Twelve respondents (34.3%) obtained the maximum score for Tutor Support and Development, while 13 (37.1%) attained the maximum score for Program Management and Overall Satisfaction. Twelve respondents (34.3%) selected the maximum response across all 25 items. These distributions indicate substantial upper-end response concentration and reduced variability, which limits the instrument's capacity to distinguish between favorable and exceptionally favorable evaluations.

Descriptive Variations According to Respondent Role

Descriptive differences were observed across the original role categories. Within the available sample, tutorial teachers reported the highest overall mean ($M = 4.46$), followed by administrators ($M = 3.99$) and subject teachers who referred students ($M = 3.94$). The single respondent who reported multiple roles had an overall score of 5.00.

Tutorial teachers recorded their highest mean for Implementation and Delivery ($M = 4.50$), followed by Program Implementation ($M = 4.49$). Administrators also rated Implementation and Delivery most favorably ($M = 4.33$), while their lowest means were recorded for Tutor Support and Development and Program Management and Overall Satisfaction (both $M = 3.77$). Referring subject teachers recorded their lowest means for these same two operational areas (both $M = 3.80$).

The one-person multiple-role category is presented for transparency but was not interpreted as a meaningful subgroup. All role-based comparisons were treated as descriptive because of the small and unequal group sizes.

Table 10

Evaluation of the Tutorial Program According to Respondent Role



Operational Area	Tutorial teachers (<i>n</i> = 23)	Administrators (<i>n</i> = 6)	Referring subject teachers (<i>n</i> = 5)	Multiple roles (<i>n</i> = 1)
Program Implementation	4.49	4.30	4.00	5
Implementation and Delivery	4.50	4.33	4.08	5
Perceived Student Progress and Outcomes	4.43	3.80	4.04	5
Tutor Support and Development	4.43	3.77	3.80	5
Program Management and Overall Satisfaction	4.45	3.77	3.80	5
Overall	4.46	3.99	3.94	5

Note. The multiple-role value represents one individual response and is reported only for transparency. The subgroup means are descriptive, and no inferential comparisons were conducted.

Descriptive Variations According to Length of Involvement

Descriptive variations were also observed according to length of involvement. Respondents involved for two to three semesters recorded the highest overall mean ($M = 4.72$), followed by those involved for one semester ($M = 4.58$). Respondents involved for less than one semester had an overall mean of 4.22, while those involved for more than three semesters recorded the lowest overall mean ($M = 3.98$).

Among respondents involved for more than three semesters, the lowest operational area means were recorded for Program Management and Overall Satisfaction ($M = 3.80$), Tutor Support and Development ($M = 3.85$), and Perceived Student Progress and Outcomes ($M = 3.90$). Respondents involved for two to three semesters recorded means ranging from 4.65 to 4.75 across the five operational areas.

Given the small subgroup sizes, particularly the group with two to three semesters of involvement ($n = 4$), these patterns should be treated as descriptive only. The data do not establish why the observed variations occurred.

Table 11

Evaluation According to Length of Involvement

Operational Area	Less than 1 semester (<i>n</i> = 13)	1 semester (<i>n</i> = 10)	2–3 semesters (<i>n</i> = 4)	More than 3 semesters (<i>n</i> = 8)
Program Implementation	4.37	4.60	4.75	4.03
Implementation and Delivery	4.31	4.58	4.65	4.33
Perceived Student Progress and Outcomes	4.20	4.54	4.70	3.90
Tutor Support and Development	4.08	4.58	4.75	3.85
Program Management and Overall Satisfaction	4.14	4.58	4.75	3.80
Overall	4.22	4.58	4.72	3.98

Note. The subgroup means are descriptive only. Some categories were small, particularly the two-to-three-semester group ($n = 4$), and should not be interpreted as stable population estimates.



Qualitative Findings

All 35 respondents entered a response in each of the three open-ended fields. Responses consisting solely of “none,” “not applicable,” “yes,” punctuation marks, or equivalent non-substantive entries were excluded from qualitative content coding but retained in the dataset. After applying this rule, 29 substantive responses remained for the question on successful program aspects, 19 for areas needing improvement, and 13 for recommendations.

Because the responses were generally brief, the qualitative findings were treated as descriptive qualitative evidence rather than as an in-depth qualitative inquiry.

Perceived Contributions of the Tutorial Program

Several respondents described perceived improvement in students’ understanding of lessons, grades, subject mastery, confidence, and engagement. They also referred to the program’s role in addressing learning gaps, particularly among students who experienced difficulty in specific subjects.

Respondents further described one-on-one or small-group instruction, modified teaching approaches, and focused guidance as beneficial. Some noted that individualized assistance allowed tutors to respond to students’ specific strengths and weaknesses.

Other responses referred to regular follow-up, review of lessons, assistance with academic tasks, monitoring of progress, and continued engagement with students. A smaller number of respondents mentioned student targeting and communication with parents as positive features.

Areas Requiring Improvement

Several respondents raised concerns about student dependence on tutors and the need to place greater emphasis on mastery, study habits, and independent learning. Some suggested that tutorial sessions should focus more clearly on remediation rather than on completing assignments.

Student eligibility also emerged as an area of concern. Responses reflected differing views on whether honor students should continue to receive tutorial services, indicating a need for clearer and consistently applied eligibility criteria.

Other responses referred to uneven workload distribution, tutor-subject mismatch, limited coverage of some subject areas, monitoring of tutorial practices, scheduling difficulties, insufficient time, and the need for more conducive tutorial spaces. Some respondents also emphasized the importance of parental follow-up and stronger coordination among stakeholders.

Recommendations for the Next Program Cycle

Respondents recommended formal orientation and training for tutors, particularly on appropriate tutorial methods, remediation, learner independence, and instructional boundaries.

They also proposed more standardized monitoring and documentation through progress checklists, written reports, and regular feedback. Other recommendations included assigning tutors according to subject expertise, distributing students more equitably, clarifying eligibility policies, improving venues and schedules, and providing more individualized support for students with foundational learning needs.

Integration of Quantitative and Qualitative Findings

The quantitative and qualitative findings generally converged in presenting the Tutorial Program as a functioning and positively evaluated academic-support initiative. High ratings for the regular conduct of sessions, instructional delivery, student selection, and personalized support were consistent with respondents’ descriptions of focused instruction, regular follow-up, and perceived improvement in student understanding.



Table 12

Joint Display of Quantitative Findings, Qualitative Evidence, and Institutional Implications

Quantitative finding	Corresponding qualitative evidence	Integrated interpretation	Institutional implication
Regular conduct of sessions received the highest item mean ($M = 4.63$, $SD = 0.55$).	Respondents referred to regular follow-up, lesson review, and continued assistance.	The program was perceived as operationally consistent in its delivery.	Retain scheduling practices while documenting actual session frequency and attendance.
Tutor training and guidance received the lowest item mean ($M = 4.00$, $SD = 0.94$).	Respondents requested formal orientation, remediation strategies, and clearer instructional boundaries.	Respondents differed in the extent to which they perceived tutor training and guidance as sufficient.	Establish a standardized pre-implementation tutor-development program.
Monitoring and feedback received a comparatively lower mean ($M = 4.20$, $SD = 0.76$).	Respondents recommended progress checklists, written reports, and closer supervision.	Respondents generally perceived monitoring as present but requested greater standardization and documentation.	Adopt common session logs and progress-monitoring tools.
Student-selection criteria received a high rating ($M = 4.51$, $SD = 0.66$).	Respondents expressed differing views regarding the participation of honor students.	Selection was generally perceived as fair, but entry, continuation, and exit rules were not uniformly understood.	Clarify eligibility criteria and distinguish remediation, reinforcement, and enrichment.
Perceived student outcomes were evaluated favorably ($M = 4.29$, $SD = 0.69$).	Respondents described better understanding, grades, confidence, and engagement, but some raised concerns about dependence on tutors.	Personnel perceived student benefits, although excessive assistance may weaken learner independence.	Combine tutorial support with gradual release of responsibility and direct student outcome monitoring.
Program Management and Overall Satisfaction received $M = 4.26$ ($SD = 0.78$).	Respondents cited scheduling, workload, subject matching, communication, and venue concerns.	Favorable management and satisfaction ratings coexisted with specific operational concerns.	Improve logistical planning, workload distribution, and stakeholder communication.

The qualitative responses also clarified the comparatively lower quantitative ratings. Tutor training and guidance was the lowest-rated item, and respondents explicitly recommended formal preparation and clearer orientation for tutors. Monitoring and feedback received one of the lower ratings in the Implementation and Delivery operational area, while respondents called for checklists, progress reports, and closer supervision.

The integrated findings also revealed issues not fully captured by the composite means. Although student-selection criteria were rated Highly Evident, respondents expressed differing views regarding the participation of honor students. Similarly, although perceived student outcomes were evaluated favorably, some respondents cautioned that tutorial practices could encourage dependence when sessions focused heavily on assignment completion.

Taken together, the findings indicate that respondents generally viewed the program positively while also identifying the need for clearer eligibility rules, more systematic tutor preparation, subject-aligned assignments, stronger monitoring, and improved logistical support.



Discussion

The findings present the Tutorial Program as an operationally established and generally favorably evaluated academic-support initiative. Regular session delivery, scheduling, and the organization of implementation received the strongest evaluations, suggesting that respondents consistently encountered the program in practice. These findings align with research emphasizing continuity, sufficient instructional exposure, curricular connection, and sustained tutor–learner engagement as important conditions of effective tutoring (Nickow et al., 2024; Robinson et al., 2024). However, the pronounced upper-end response concentration limits the distinctions that can be drawn among program areas. The favorable ratings should therefore be understood as stakeholder assessments of implementation rather than evidence that the program had achieved optimal quality.

Student selection was one of the most favorably evaluated practices, indicating broad confidence in the program's targeting procedures. Nevertheless, the qualitative responses revealed differing views about whether honor students should enter or remain in the program. This apparent inconsistency suggests that respondents may have regarded the existing procedures as generally fair while interpreting the program's purpose differently. Some viewed tutoring primarily as remediation for learners experiencing academic difficulty, whereas others appeared to recognize a continuing reinforcement function. The finding supports the development of explicit entry, continuation, and exit criteria that distinguish remediation, reinforcement, and enrichment, especially when staffing and instructional resources are limited.

Respondents perceived the program as contributing to academic performance, study habits, confidence, motivation, participation, and reduced failure or incomplete-grade incidence. These perceptions are compatible with broader evidence that structured tutoring can support student learning when it is aligned with diagnosed needs and delivered consistently (Nickow et al., 2024; Ritter et al., 2009; Slavin et al., 2011). However, the present study did not include pre-intervention and post-intervention achievement data, attendance records, student-level grades, comparison groups, or longitudinal measures. The results therefore demonstrate that personnel perceived student benefits; they do not establish the magnitude of those benefits or show that tutorial participation caused them.

Tutor Support and Development received the lowest operational-area mean, while sufficient training and guidance for tutorial teachers obtained the lowest item mean and the largest standard deviation among the 25 indicators. This pattern suggests that respondents differed more in their assessment of tutor preparation than in other program areas. The qualitative responses reinforced this concern by calling for formal orientation, clearer guidance on remediation, subject-aligned assignment, and explicit boundaries between supporting learning and completing tasks for students. These concerns are important because effective tutoring depends not only on subject knowledge but also on the ability to diagnose misconceptions, scaffold learning, provide corrective feedback, and gradually promote learner independence. Implementation research likewise identifies staff preparation, supervision, and institutional support as important determinants of program quality (Durlak & DuPre, 2008). The program would therefore benefit from a standardized tutor-development component that clarifies instructional goals, appropriate levels of assistance, and strategies for transferring responsibility progressively to the learner.

Monitoring and feedback also emerged as a comparatively weaker area. Although respondents generally perceived monitoring as present, they recommended clearer documentation, progress checklists, written reports, closer supervision, and more regular feedback. These responses suggest that existing monitoring practices may not yet be sufficiently standardized across tutors and sessions. Informal oversight can help coordinators remain aware of program activities, but systematic documentation is needed to assess whether sessions are delivered consistently, whether students are progressing, and whether adjustments are necessary. At the same time, implementation quality should not be reduced to rigid procedural compliance, since tutorial arrangements must often be adapted to staffing, schedules, learner needs, and available resources. Standardized but flexible monitoring tools could therefore support implementation consistency while still allowing tutors to respond to subject-specific and learner-specific needs (Durlak & DuPre, 2008).

Program Management and Overall Satisfaction was also rated positively, although overall satisfaction was the second lowest item in the instrument. The results suggest that respondents generally recognized the program as organized and aligned with the school's philosophy of personalized learning support. However, not all respondents expressed agreement regarding its outcomes and management. The qualitative comments identified scheduling, venue



availability, workload distribution, tutor-subject mismatch, and communication with parents as areas requiring improvement. These concerns indicate that operational support remains an important part of program effectiveness. A tutorial program may have strong instructional intentions but still be weakened by insufficient time, inappropriate spaces, unclear assignments, or uneven distribution of students among tutors.

Exploratory subgroup patterns should not be overinterpreted. Tutorial teachers reported higher descriptive means than administrators and referring teachers, while respondents with more than three semesters of involvement reported lower means than several less-experienced groups. These differences may reflect variation in role, exposure, expectations, or experience, but the small and unequal categories prevent reliable explanation. The subgroup results are therefore best regarded as questions for future interviews or larger institutional studies rather than as evidence of meaningful population differences.

The substantial upper-end response concentration across all five operational areas presents an additional methodological consideration. More than one-third of respondents attained the maximum possible score on the overall instrument, and similarly high proportions reached the maximum in individual operational areas. Such concentration reduces variability and limits the instrument's ability to distinguish between moderately strong and exceptionally strong implementation. It may also reflect response tendencies such as acquiescence or reluctance to express criticism in an internal evaluation. The open-ended responses were therefore particularly important because they revealed concerns that were not fully visible in the numerical scores. This reinforces the value of mixed-methods evaluation in institutional settings, where favorable ratings may coexist with specific calls for improvement.

The integration of numerical and written responses provides the principal contribution of the study. High ratings showed that respondents generally valued the program, while their written comments identified implementation concerns involving tutor preparation, learner dependence, monitoring, eligibility, subject matching, workload, venue, and scheduling. These findings are not contradictory. They indicate that stakeholders may support continuation of a program while also recognizing the need for procedural refinement. In institutional research, this distinction is important because favorable satisfaction scores should not prevent critical examination of implementation conditions.

The study also demonstrates how routinely collected evaluation data may be used to inform institutional improvement. The findings identify four interconnected priorities: standardized tutor preparation and instructional boundaries; clearer learner eligibility and subject matching; systematic monitoring, documentation, and communication; and future evaluation using direct student indicators and multiple stakeholder perspectives. These priorities move the evaluation beyond general satisfaction toward actionable program development (Fitzpatrick et al., 2011; Stufflebeam & Coryn, 2014).

The findings remain limited by the single-institution setting, small respondent group, upper-end response concentration, lack of documented construct validation, and brief qualitative comments. Student outcomes were assessed indirectly through personnel perceptions, and the secondary design prevented follow-up questioning. Future evaluations should incorporate student and parent perspectives, diagnostic and achievement records, attendance and completion data, repeated measurements across program cycles, and clearer documentation of tutorial dosage and participation. A larger dataset would also permit more rigorous instrument validation and more stable subgroup analysis.

Conclusion

The Tutorial Program was generally evaluated favorably by the personnel involved in its implementation and oversight. Its strongest perceived feature was the consistency of instructional delivery, particularly the regular conduct of sessions and adherence to schedules. Respondents also perceived the program as supporting learner understanding, academic progress, motivation, and participation. These judgments, however, represent stakeholder perceptions and should not be interpreted as direct evidence that the program caused improvements in student outcomes.

The quantitative and qualitative findings jointly identified tutor preparation, monitoring, eligibility rules, instructional boundaries, subject matching, workload distribution, scheduling, and venue support as priorities for improvement. Thus, the central issue was not whether the program should continue, but how its procedures could be more clearly



standardized and supported. The findings support continuation of the program together with targeted institutional reforms in tutor development, learner selection, monitoring, and logistical coordination.

As a secondary analysis of internal evaluation data, the study also demonstrates the potential value of routine institutional evidence for continuous improvement. Its conclusions remain limited by the small, single-institution sample, concentration of high ratings, absence of direct student measures, and limited depth of the open-ended responses. Future evaluation should combine stakeholder perceptions with student-level participation, achievement, attendance, and longitudinal data to determine more directly how tutorial support contributes to learner progress.

Recommendations

First, the Integrated Basic Education Department should establish a standardized tutor-preparation program before each tutorial cycle. The orientation should address remediation strategies, diagnosis of learning gaps, scaffolding, feedback, subject-specific support, and gradual development of learner independence. It should also clarify the distinction between facilitating learning and completing assignments or academic requirements for students.

Second, the department should clarify and communicate program policies governing learner entry, continued participation, and completion. These policies should distinguish remediation, reinforcement, and enrichment and should address the participation of students whose academic standing improves while they are receiving support. Tutors should be assigned according to subject expertise as far as practicable, and student loads should be distributed equitably.

Third, program monitoring should be strengthened through standardized but manageable documentation. Session logs, competency checklists, attendance records, and brief progress reports may be used to document the learning needs addressed, instructional strategies employed, student participation, and recommended follow-up. Monitoring should be used for feedback and program improvement rather than only for administrative compliance. Venue, scheduling, instructional materials, and communication among tutors, subject teachers, administrators, parents, and learners should also be reviewed as part of this monitoring system.

Fourth, future evaluations should include evidence beyond personnel perceptions. Student and parent feedback, diagnostic scores, grades, attendance, tutorial completion, failure or incomplete-grade records, and repeated measurements across program cycles should be examined where ethically and administratively feasible. Further instrument development should include independent content review, pilot testing, examination of item redundancy, and construct validation using a larger sample.

References

- Boone, H. N., Jr., & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension*, 50(2), Article 48. <https://doi.org/10.34068/joe.50.02.48>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Department of Education. (2024a, November 26). *DepEd, partners report positive educational outcomes with phone-based tutoring program*. <https://www.deped.gov.ph/2024/11/26/deped-partners-report-positive-educational-outcomes-with-phone-based-tutoring-program/>
- Department of Education. (2024b, December 21). *DepEd, DSWD renew commitment on expanded Tara, Basa! Tutoring Program*. <https://www.deped.gov.ph/2024/12/21/deped-dswd-renew-commitment-on-expanded-tara-basa-tutoring-program/>
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3-4), 327-350. <https://doi.org/10.1007/s10464-008-9165-0>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107-



115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Pearson.

Nickow, A., Oreopoulos, P., & Quan, V. (2024). The promise of tutoring for PreK–12 learning: A systematic review and meta-analysis of the experimental evidence. *American Educational Research Journal*, 61(1), 74–107. <https://doi.org/10.3102/00028312231208687>

Organisation for Economic Co-operation and Development. (2023a). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>

Organisation for Economic Co-operation and Development. (2023b). *PISA 2022 results (Volumes I and II)—Country note: Philippines*. OECD Publishing. <https://doi.org/10.1787/a0882a2d-en>

Organisation for Economic Co-operation and Development. (2023c). *Equity and inclusion in education: Finding strength through diversity*. OECD Publishing. <https://doi.org/10.1787/e9072e21-en>

Organisation for Economic Co-operation and Development. (2025). *Teacher support for student learning: Insights from PISA*. OECD Publishing. <https://doi.org/10.1787/97b3a899-en>

Republic Act No. 12028. (2024). *Academic Recovery and Accessible Learning Program Act*. <https://elibrary.judiciary.gov.ph/thebookshelf/showdocs/2/97912>

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research*, 79(1), 3–38. <https://doi.org/10.3102/0034654308325690>

Robinson, C. D., Kraft, M. A., Loeb, S., & Schueler, B. E. (2024). *Design principles for accelerating student learning with high-impact tutoring*. EdResearch for Action. <https://edresearchforaction.org/research-briefs/design-principles-for-accelerating-student-learning-with-high-impact-tutoring/>

Schreier, M. (2012). *Qualitative content analysis in practice*. SAGE.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1–26. <https://doi.org/10.1016/j.edurev.2010.07.002>

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, and applications* (2nd ed.). Jossey-Bass.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>

World Bank, United Nations Children's Fund, United Nations Educational, Scientific and Cultural Organization, & Organisation for Economic Co-operation and Development. (2022). *From learning recovery to education transformation: Insights and reflections from the fourth survey on national education responses to COVID-19 school closures*. World Bank. <https://openknowledge.worldbank.org/handle/10986/38112>