

Goal-Free Evaluation of a Student-Led Literacy Mentoring Program to Support Children's Literacy in Bayan Village

Sugita Dwi Wardani, Herlina, Kholisussa'di

Prodi Pendidikan Masyarakat, Fakultas Ilmu Pendidikan dan Psikologi, Universitas Pendidikan Mandalika.

*Correspondence e-mail: kholisussakdi@undikma.ac.id

Accepted: December 2024; Revised: December 2024; Published: December 2024

.Abstract

Student-led literacy mentoring programs have the potential to expand children's learning opportunities beyond formal schooling; however, evaluations of their perceived benefits, implementation processes, experienced outcomes, and supporting conditions remain limited. This study aimed to evaluate a Literacy Mentoring Program implemented by the Student Association of the Community Education Study Program in Bayan Village, North Lombok Regency, by adapting a Goal-Free Evaluation orientation. The study employed a quantitative descriptive evaluation design. The population consisted of 40 children who participated in the program, of whom 30 were selected through purposive sampling. The primary data were collected using a 30-item Likert-scale questionnaire, supported by observation and documentation. The evaluation covered five aspects: perceived changes experienced by participants (actual effects), outcomes beyond the primary program focus (unintended effects), participants' ratings of the implementation process, perceived program benefits (program merit/worth), and contextual factors. The results showed scores of 80.22% for actual effects, 81.44% for unintended effects, 83.78% for the implementation process, 86.33% for program merit/worth, and 85.78% for contextual factors. Perceived program benefits received the highest rating, whereas perceived changes among participants received the lowest. At the item level, perceived improvement in reading fluency reached only 57.33%, substantially lower than the ratings for reading interest, self-confidence, and comprehension. These findings indicate that participants perceived the program as beneficial and supported by favorable implementation conditions. However, because the primary data were based on post-program self-reports without objective pre- and post-intervention assessments of reading performance, the findings cannot be interpreted as evidence of measurable improvement in reading ability. Future evaluations should incorporate objective reading-performance assessments and stronger evaluation designs.

Keywords: Goal-Free Evaluation; literacy mentoring; children's literacy; community education; program evaluation.

How to Cite: Wardani, S. D. ., Herlina, H., & Sa'di, K. (2024). Goal-Free Evaluation of a Student-Led Literacy Mentoring Program to Support Children's Literacy in Bayan Village. *Reflection Journal*, 4(2), 105-116. <https://doi.org/10.36312/rj.v4i2.6610>



<https://doi.org/10.36312/rj.v4i2.6610>

Copyright© 2024, Wardani et al

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) License.



INTRODUCTION

Reading ability is a fundamental skill that determines children's capacity to obtain information, understand learning materials, and develop knowledge across various fields. Reading literacy is not limited to recognizing letters, pronouncing words, or reading text mechanically, but also includes the ability to understand, interpret, evaluate, and use information obtained from written texts. Developing these abilities remains a major challenge in Indonesia. The 2022 Programme for International Student Assessment (PISA) results showed that Indonesian students achieved a reading score of 359, which was lower than the OECD average, and only about 25% of Indonesian students reached at least Level 2 in reading (OECD, 2023). This condition indicates that strengthening literacy cannot be positioned solely as the responsibility of formal schooling. Support from families, communities, social organizations, and non-formal education programs is also needed to expand children's opportunities to engage in meaningful reading experiences.

Previous studies have shown that structured and interactive reading assistance can provide an environment that supports literacy development. The meta-analysis conducted by Tang et al. (2026) found that *dialogic reading* had positive effects on various aspects of children's literacy. This approach positions children as active participants through questioning, responding, elaborating, and interacting during reading activities. Dong et al. (2024) similarly demonstrated that elements of *dialogic reading*, including prompts, feedback, and opportunities to elaborate responses, contributed to children's language development. In the Indonesian context, Suryati et al. (2017) found that dialogic reading strategies produced better outcomes in elementary school students' reading comprehension than conventional reading instruction. These findings indicate that the quality of interaction during reading activities is an important component of children's literacy experiences.

Support from mentors or peers also has the potential to create more active reading experiences. Voellinger et al. (2023) reported that a multicomponent reading intervention based on *peer-assisted learning* improved reading comprehension immediately after the intervention, although the sustainability of the effects at follow-up required further consideration. More recent evidence from a cluster randomized controlled trial showed that systematically and consistently implemented Peer Assisted Learning Strategies improved reading attainment, oral reading fluency, and comprehension (Breadmore et al., 2026). These studies emphasize the importance of structured implementation, repeated practice, social interaction, and the quality of mentoring. However, findings from experimental studies do not imply that every mentoring program will automatically generate similar improvements in reading ability because outcomes depend on program design, intensity, duration, implementation fidelity, and participant characteristics.

Literacy also includes an affective dimension that cannot be entirely separated from reading activity. Wigfield et al. (2016) explained that reading motivation is associated with children's engagement with texts and the development of reading comprehension. Children who have positive reading experiences may develop greater engagement, confidence, and willingness to interact with written texts. Therefore, community-based literacy programs are relevant not only for providing opportunities to practice reading but also for creating more personal, supportive, and enjoyable learning experiences. Outcomes such as reading interest, willingness to ask questions, independence, peer interaction, and enthusiasm for learning are therefore important to consider. Nevertheless, these social-affective outcomes should remain conceptually distinct from changes in reading ability measured through objective performance assessments.

Several programs in Indonesia have demonstrated the potential contribution of university students and communities to children's literacy development. Permatasari et al. (2022), for example, reported changes in students' reading ability after participation in literacy mentoring classes conducted through the Kampus Mengajar Program. Gunawan and Arif (2025) emphasized the importance of appropriate stimulation, attractive reading materials, and mentoring that is responsive to children's characteristics. In the context of West Nusa Tenggara, Zakiah and Herlina (2023) demonstrated that family-based literacy mentoring programs could be evaluated by considering contextual, input, process, and outcome components. Meanwhile, Herlina et al. (2025) showed that student-led literacy mentoring programs in North Lombok could combine pedagogical structures with participatory approaches to community empowerment. These studies suggest that student-led literacy mentoring has both educational and social potential, but the way such programs are evaluated remains an important methodological issue.

Previous studies have generally emphasized program implementation, changes in targeted competencies, or the extent to which predetermined program components were achieved. Such approaches provide valuable information about program implementation, but they do not always reveal how participants perceive the program, what benefits they actually experience, what conditions support implementation, and what other outcomes emerge beyond the primary program focus. In literacy program evaluation, Trianggoro and Koeswanti (2021) and Nugroho and Mawardi (2024) applied *Goal-Free Evaluation* to identify both the benefits and limitations of literacy program implementation. Nevertheless, evaluations of literacy mentoring programs initiated specifically by student organizations at the study-program level, particularly within rural community education contexts, remain relatively limited. This gap creates a need for an evaluation approach that does not merely ask whether activities are aligned with formal program objectives but also examines participants' experiences and the outcomes they perceive.

Theoretically, *Goal-Free Evaluation* was introduced by Scriven (1972) as a response to evaluation approaches that were excessively oriented toward formal program goals. Within this approach, evaluators focus on what a program actually produces and on the needs of its beneficiaries, thereby reducing the risk of *tunnel vision* created by searching only for evidence that corresponds to predetermined objectives. Scriven (1991) further emphasized that a goal-free orientation should direct attention toward a broad range of actual effects, including positive, negative, and unintended effects. In its stricter form, the approach requires independent evaluators to avoid or limit their knowledge of official program objectives in order to reduce *goal contamination*. Subsequent literature has shown that the operationalization of this principle in practice is not always uniform and requires explicit procedures before an evaluation can appropriately be characterized as genuinely goal-free (Youker et al., 2014).

In the present study, *Goal-Free Evaluation* was applied as an adapted evaluative orientation, rather than as a claim that evaluators were completely blinded to all formal program objectives. This position is important because the research instrument still contained indicators related to reading experiences and perceived program benefits. The goal-free orientation was primarily used to broaden the scope of evaluation beyond formal objective attainment toward participants' perceived outcomes, social-affective outcomes beyond the main program focus, experiences of the implementation process, perceived *merit/worth*, and contextual factors. With this limitation, the study was not designed to establish a causal relationship between the program and improvement in reading ability. Instead, it aimed to provide an evaluative description of participants' experiences and perceptions after participating in the program.

The Literacy Mentoring Program implemented by the Student Association of the Community Education Study Program in Bayan Village, North Lombok Regency, formed part of community service and empowerment activities. The program involved elementary school-aged children with varying levels of reading ability and included assisted reading activities, reading comprehension exercises, the use of reading materials, and learning interactions adjusted to children's needs. Although the program had been implemented, systematic information was still needed regarding how participants perceived changes they experienced, the quality of mentoring, the benefits of the program, outcomes beyond the main focus, and contextual conditions supporting implementation. Such evaluation is important so that decisions concerning program development are not based solely on whether activities were conducted or on general participant satisfaction.

Based on this gap, this study aimed to evaluate the Literacy Mentoring Program implemented by the Student Association of the Community Education Study Program in Bayan Village by adapting a *Goal-Free Evaluation* orientation. The evaluation focused on

five aspects: participants' perceived changes (*actual effects*), outcomes beyond the primary focus anticipated in the instrument (*unintended effects*), participants' ratings of the implementation process, perceived program benefits (*program merit/worth*), and contextual factors that supported or hindered implementation. The contribution of this study lies in providing an empirical profile of how children, as program beneficiaries, perceived a student-led literacy program within a rural community education context. The findings are expected to provide a basis for improving the mentoring design, strengthening evaluation procedures, and guiding future evaluations toward more objective and measurable assessments of reading performance.

METHOD

This study employed a quantitative descriptive evaluation design by adapting a *Goal-Free Evaluation* orientation. This approach was used to obtain a broader profile of participants' experiences with the program rather than focusing solely on the achievement of formal objectives. The evaluation addressed five aspects: participants' perceived changes (*actual effects*), outcomes beyond the primary program focus (*unintended effects*), participants' ratings of the implementation process, perceived program benefits (*program merit/worth*), and contextual factors that supported or hindered implementation. In this study, the application of *Goal-Free Evaluation* was adaptive because the evaluators were not completely blinded to the program's formal objectives. Instead, the goal-free orientation was primarily used to broaden the scope of evaluation toward participants' actual experiences and perceived outcomes. Similar approaches have been applied in evaluations of literacy programs in elementary schools (Trianggoro & Koeswanti, 2021; Nugroho & Mawardi, 2024).

The study was conducted in 2025 within the Literacy Mentoring Program implemented by the Student Association of the Community Education Study Program at SDN 2 Bayan, Bayan Village, North Lombok Regency. The population consisted of 40 children participating in the program, of whom 30 were selected as respondents through *purposive sampling*. Participation criteria included active involvement in the mentoring program, the ability to respond to the research instrument, and willingness to participate in data collection. Selecting respondents with direct experience of the program was intended to ensure that the information obtained was relevant to the evaluation objectives. This principle of aligning participant selection with research purposes is consistent with Campbell et al. (2020). Nevertheless, the use of purposive sampling in a relatively small population may introduce *selection bias*, particularly if children with lower literacy levels were less able to respond to the questionnaire. This limitation was therefore considered when interpreting the findings.

The primary instrument was a 30-item questionnaire using a five-point Likert scale ranging from strongly agree (5), agree (4), undecided (3), disagree (2), to strongly disagree (1). Each evaluation aspect was represented by six statements. The Likert scale was employed to obtain information about respondents' perceptions and levels of agreement regarding their experiences during the program (Joshi et al., 2015). The *actual effects* dimension covered perceptions of reading interest, concentration, self-confidence, comprehension, and reading fluency. *Unintended effects* addressed social-affective outcomes beyond the main program focus. Process ratings included scheduling, clarity of mentoring instructions, reading materials, learning atmosphere, and mentor assistance. *Program merit/worth* represented perceived program benefits, whereas *contextual factors* addressed conditions supporting or constraining program implementation.

The questionnaire served as the primary source of quantitative data. Observation was used as supporting evidence to examine activity implementation and participant engagement, while documentation was used to verify evidence of program

implementation. Because the process-dimension scores were derived from participants' questionnaire responses, they were interpreted as *participant-rated process* rather than independent observational data. Observation and documentation were therefore treated as contextual supporting information and were not incorporated into the questionnaire scores.

Data were analyzed using descriptive statistics to systematically summarize and describe patterns in the measurements (Mishra et al., 2019). Percentages were calculated using the formula $P = (\text{obtained score} / \text{maximum score}) \times 100\%$. Each dimension had a maximum score of 900, while the complete 30-item questionnaire had a maximum score of 4,500. Percentages were used as descriptive indices for comparing patterns across indicators and dimensions rather than as objective measures of change in reading ability. The overall score was likewise treated as an arithmetic summary rather than as a validated single construct of program effectiveness. Because no pre- and post-program reading assessments or comparison groups were included, the analysis was not intended to establish a causal relationship between participation in the program and changes in children's reading ability.

RESULTS AND DISCUSSION

The evaluation results showed that participants generally provided positive assessments of the Literacy Mentoring Program implemented by the Student Association of the Community Education Study Program in Bayan Village. However, all quantitative scores in this study should primarily be interpreted as participants' perceptions of their experiences with the program, rather than as objective measures of changes in reading ability. The main instrument was a post-program questionnaire that asked children to assess perceived changes, the implementation process, program benefits, and contextual conditions. Accordingly, the dimensions of *actual effects*, *unintended effects*, *process observation*, *program merit/worth*, and *contextual factors* were retained in accordance with the adopted evaluation framework, but their interpretation was limited to what the available data could support.

Descriptively, the highest percentage was obtained for *program merit/worth* at 86.33%, followed by *contextual factors* at 85.78%, *process observation* at 83.78%, *unintended effects* at 81.44%, and *actual effects* at 80.22%. The sum of all item scores was 3,758 out of a maximum score of 4,500, equivalent to 83.51%. Because the five dimensions represent conceptually different aspects of the program, the overall percentage of 83.51% was treated only as an arithmetic descriptive summary across dimensions, rather than as evidence that these dimensions constitute a single validated construct of "program effectiveness."

Table 1. Evaluation profile of the Literacy Mentoring Program

No.	Evaluation dimension	Obtained score	Maximum score	Percentage	Descriptive interpretation
1	<i>Actual effects</i>	722	900	80.22%	Positive perceived changes, but the lowest score across dimensions
2	<i>Unintended effects</i>	733	900	81.44%	Social-affective outcomes beyond the primary focus were rated relatively highly
3	<i>Process observation</i>	754	900	83.78%	The implementation process was positively rated by participants

No.	Evaluation dimension	Obtained score	Maximum score	Percentage	Descriptive interpretation
4	<i>Program merit/worth</i>	777	900	86.33%	Perceived benefits and value of the program received the highest rating
5	<i>Contextual factors</i>	772	900	85.78%	Conditions supporting implementation were rated relatively strongly
	Cross-dimensional summary	3,758	4,500	83.51%	Descriptive summary rather than a validated single effectiveness index

The interpretation of the results therefore focused more on patterns within individual dimensions and items than on categorical percentage labels. This was particularly important for *actual effects*, which reached 80.22%, only slightly above the 80% threshold previously used to classify results as “highly effective.” A difference of only 0.22 percentage points from a categorical threshold does not necessarily represent a substantively meaningful distinction. For this reason, the analysis emphasized the variation among indicators within each dimension.

Participants' Perceived Changes (*Actual Effects*)

The *actual effects* dimension obtained a score of 722 out of 900, equivalent to 80.22%. This dimension contained several conceptually different outcomes, including reading interest, concentration, self-confidence, reading comprehension, and reading fluency. Therefore, the aggregate score should be interpreted cautiously because it combines affective outcomes, perceived cognitive outcomes, and reading-performance-related indicators within the same dimension.

The highest scores were found for children's interest in reading books independently and their confidence in reading aloud in front of others, each scoring 131 out of 150 or 87.33%. Perceived reading comprehension was also rated relatively highly, with a score of 130 out of 150 or 86.67%. These results indicate that the program was most strongly associated, from the participants' perspective, with increased reading interest, confidence, and perceived understanding of texts.

In contrast, the item concerning reading fluency compared with the period before participation in the program obtained only 86 out of 150, or 57.33%. This percentage was substantially lower than the scores of the other indicators within the same dimension and constitutes one of the most important findings of the study. The difference suggests that children could perceive the mentoring activities as engaging and confidence-building without necessarily perceiving comparable progress in reading fluency.

Because the study did not conduct objective reading assessments before and after the program, this item cannot be interpreted as evidence of measured improvement in reading performance. The statement concerning being “more fluent than before the program” is more appropriately understood as retrospective perceived change. This result highlights the importance of distinguishing between participants' positive learning experiences and objectively verified changes in reading performance.

Outcomes Beyond the Primary Program Focus (*Unintended Effects*)

The *unintended effects* dimension obtained a score of 733 out of 900 or 81.44%. The highest ratings were found for discipline in following the program schedule and

enthusiasm for participating in other learning activities, each scoring 133 out of 150 or 88.67%. Willingness to ask questions when experiencing difficulty understanding a text, inviting peers to read together, and sharing reading activities at home each received 128 out of 150 or 85.33%. The lowest score in this dimension concerned other positive habits outside literacy activities, which reached 123 out of 150 or 82.00%.

These findings indicate that participants associated the program with social and affective outcomes extending beyond reading activities themselves. However, the term *unintended effects* requires careful qualification. Because discipline, willingness to ask questions, peer reading, and other positive behaviors had already been included as predetermined questionnaire items, the findings are better interpreted as outcomes beyond the primary program focus that had been anticipated in the instrument, rather than as entirely unforeseen effects emerging naturally from the program.

Truly unexpected outcomes would be more appropriately identified through open-ended questions, interviews, independent observations, or documentary analysis designed to capture effects that had not been specified before data collection. Thus, the findings in this dimension provide useful information about broader social-affective experiences, but they should not be interpreted as a comprehensive identification of all unintended program outcomes.

Participants' Ratings of the Implementation Process

The dimension originally labeled *process observation* obtained a score of 754 out of 900 or 83.78%. Because these numerical scores were derived from participants' questionnaire responses rather than from systematic independent observations, they are more accurately interpreted as participant-rated process.

Implementation according to the planned schedule received the highest score, at 130 out of 150 or 86.67%. Clarity of mentors' instructions and consistency of program implementation each received 129 out of 150 or 86.00%. The appropriateness of reading materials scored 124 out of 150 or 82.67%, while the comfort of the learning environment scored 123 out of 150 or 82.00%. The lowest-rated process indicator was mentors' active assistance to children experiencing reading difficulties, which obtained 119 out of 150 or 79.33%.

This pattern suggests that organizational features such as scheduling, clarity of instructions, and consistency were viewed positively by participants. Nevertheless, the comparatively lower score for individualized assistance indicates an area that may require attention, particularly because the participating children had varying levels of reading ability. Children experiencing greater difficulty may require more intensive or differentiated support than can easily be provided within a group mentoring format.

[1] Perceived Program Benefits (*Program Merit/Worth*)

The *program merit/worth* dimension obtained the highest score across all dimensions, reaching 777 out of 900 or 86.33%. Statements indicating that the program increased enthusiasm for learning and encouraged greater independence in learning to read each scored 132 out of 150 or 88.00%. The perception that the program was beneficial for supporting reading and deserved to be continued each scored 130 out of 150 or 86.67%. An enjoyable learning experience obtained 127 out of 150 or 84.67%, while perceived tangible benefits for children's development received 126 out of 150 or 84.00%.

The predominance of *program merit/worth* indicates strong participant acceptance and perceived usefulness. The children generally viewed the program as relevant, enjoyable, and beneficial to their learning experiences. Nevertheless, these ratings should not be interpreted as objective evidence that the program improved reading

performance. They provide stronger evidence regarding *acceptability*, perceived relevance, positive learning experiences, and participant-perceived value.

Profile of Contextual Factors

The *contextual factors* dimension obtained 772 out of 900 or 85.78%. Unlike outcome dimensions, this score was not interpreted as a level of effectiveness. Contextual factors represent conditions that participants perceived as supporting or influencing program implementation.

Parental support obtained the highest score, reaching 132 out of 150 or 88.00%, followed by the availability of facilities and infrastructure at 130 out of 150 or 86.67%. Availability of reading materials and suitability of implementation time each obtained 128 out of 150 or 85.33%, whereas environmental support scored 127 out of 150 or 84.67%. These findings indicate that participants generally perceived the program as operating within a relatively supportive context.

Contextual conditions should therefore be conceptually separated from program effectiveness. High levels of parental support or adequate facilities do not necessarily mean that the program objectively improves reading performance. Rather, these conditions provide a potentially favorable environment in which mentoring activities can be implemented.

Table 2. Highest- and lowest-rated items within each evaluation dimension

Dimension	Highest-rated item	Score	Lowest-rated item	Score
<i>Actual effects</i>	Independent reading interest and confidence in reading aloud	131/150 (87.33%)	Reading fluency compared with before the program	86/150 (57.33%)
<i>Unintended effects</i>	Discipline in following the schedule and enthusiasm for other learning activities	133/150 (88.67%)	Other positive habits outside literacy activities	123/150 (82.00%)
<i>Participant-rated process</i>	Activities implemented according to schedule	130/150 (86.67%)	Mentors actively assisted children experiencing reading difficulties	119/150 (79.33%)
<i>Program merit/worth</i>	Learning enthusiasm and independence in learning to read	132/150 (88.00%)	Tangible benefits for children's development	126/150 (84.00%)
<i>Contextual factors</i>	Parental support	132/150 (88.00%)	Environmental conditions supporting program implementation	127/150 (84.67%)

The findings reveal a relatively consistent pattern in which participants reported positive experiences with the Literacy Mentoring Program, although the strength of these perceptions varied across dimensions. *Program merit/worth*, contextual conditions, and implementation processes received higher ratings than *actual effects*. This pattern suggests that the program was more strongly perceived as useful, acceptable, and supported by favorable conditions than as producing equally strong perceived changes across all reading-related indicators.

The findings for *actual effects* are particularly important. Independent reading interest, self-confidence, and perceived comprehension received high ratings, whereas perceived improvement in reading fluency reached only 57.33%. This discrepancy

demonstrates heterogeneity among the outcomes captured by the questionnaire. A program can be enjoyable, increase confidence, and strengthen willingness to engage with texts without necessarily producing equivalent changes in reading fluency.

Evidence from reading intervention research helps contextualize this pattern, although the level of evidence in the present study should not be equated with experimental studies. Voellinger et al. (2023) reported improvements in reading comprehension following a strategy-based peer-assisted intervention, although the effects were less consistent at follow-up. Using a cluster randomized controlled trial, demonstrated that systematically implemented Peer Assisted Learning Strategies could improve reading attainment, oral fluency, and comprehension. These studies employed stronger designs for measuring changes in reading performance. In contrast, the present study relied primarily on post-program participant perceptions. Consequently, the percentages obtained in Bayan Village should not be directly compared with experimentally derived effect sizes or objectively measured gains. Rather, these studies indicate that improvements in reading performance typically depend on adequate intensity, structure, consistency, and implementation quality.

The interactive nature of literacy mentoring is also relevant to research on *dialogic reading*. Demonstrated through meta-analysis that dialogic reading generally contributes positively to children's literacy development. Dong et al. (2024) further showed that prompts, feedback, and opportunities to expand children's responses during shared reading can support language development. In Indonesia, Suryati et al. (2017) found that dialogic reading strategies supported elementary school students' reading comprehension. These studies provide a useful basis for strengthening the mentoring program through activities such as asking children to predict text content, explain unfamiliar vocabulary, retell stories, connect texts to everyday experiences, and engage in repeated reading. Nevertheless, these strategies should be treated as implications for program development rather than as explanations of causal mechanisms demonstrated by the present data.

The social-affective outcomes also deserve attention. Participants reported relatively high levels of discipline, willingness to ask questions, peer interaction, and enthusiasm for other learning activities. Wigfield et al. (2016) emphasized that reading motivation is associated with reading engagement and comprehension. The present findings are conceptually consistent with the view that literacy experiences extend beyond technical reading skills to include motivation, confidence, participation, and social interaction. However, affective outcomes must remain distinct from objective reading performance. For instance, showed that improvements in reading achievement do not necessarily correspond with significant improvements across all motivational or self-efficacy indicators. Social-affective outcomes should therefore be regarded as meaningful outcomes in their own right rather than substitutes for direct measures of reading ability.

Participants' ratings of the implementation process provide additional practical information. Scheduling, clarity of guidance, implementation consistency, and appropriateness of reading materials were generally viewed positively. These findings are relevant to Permatasari et al. (2022), who reported that scheduled literacy mentoring activities involving reading practice, reading habits, and story discussion could support students' learning experiences. Similarly emphasized appropriate stimulation and engaging reading materials in literacy mentoring. Within the Bayan program, individualized assistance received the lowest process rating. This result suggests that future implementation could place greater emphasis on differentiated support for children with different levels of reading ability, especially those who experience persistent reading difficulties.

The high *program merit/worth* score indicates that participants considered the program relevant, enjoyable, and valuable. This finding is consistent with Herlina et al. (2025), who showed that student-led literacy mentoring in North Lombok could combine structured learning activities with participatory community engagement. Within community education, the value of a program is not restricted to measurable academic outcomes. It may also include increased access to learning opportunities, participation, positive learning experiences, and interaction between university students and community members. Nevertheless, strong participant acceptance should remain analytically separate from evidence of objective effectiveness.

A similar distinction applies to contextual factors. Parental support was the highest-rated contextual indicator, followed by facilities, reading materials, scheduling, and environmental conditions. Zakiah and Herlina (2023) demonstrated that literacy program evaluation benefits from considering context, input, process, and implementation support. The present findings similarly suggest that literacy mentoring does not occur in isolation. Children's opportunities to engage with reading may depend on family participation, access to reading materials, school conditions, and broader community support. These contextual characteristics are better conceptualized as conditions facilitating implementation rather than as measures of program effectiveness themselves.

From a *Goal-Free Evaluation* perspective, the study provides information concerning participants' perceived outcomes, program benefits, implementation experiences, and contextual conditions. Trianggoro and Koeswanti (2021) and Nugroho and Mawardi (2024) demonstrated that *Goal-Free Evaluation* can be applied to identify both benefits and limitations of literacy programs. However, the present operationalization also has important limitations because several outcomes classified as *unintended effects* were predetermined in the questionnaire. Claims regarding genuinely unforeseen effects should therefore remain limited. Future evaluations should complement structured questionnaires with interviews, open-ended questions, systematic observations, or documentary evidence to identify outcomes that were not anticipated before data collection.

The principal contribution of this study is therefore better understood as providing an empirical profile of participants' perceptions of a student-led literacy mentoring program within a rural community education context. Participants generally rated the program highly in terms of perceived benefits, supportive contextual conditions, and implementation experiences, while perceived direct changes in reading-related outcomes were more variable. In particular, the relatively low reading-fluency score compared with other indicators demonstrates that program acceptance and perceived usefulness should not be equated with measurable gains in reading performance.

The findings must also be interpreted within the limitations of the research design. The primary evidence was based on post-program self-reports without objective pre- and post-program reading assessments. The sample included only 30 of the 40 program participants through purposive sampling, creating the possibility that children who were more capable of responding to the questionnaire were disproportionately represented compared with children with lower literacy levels. Furthermore, the study was conducted at a single school in Bayan Village, which limits the extent to which its findings can be generalized to student-led literacy programs in other communities.

Based on this profile, the program has a reasonable basis for being regarded as acceptable and beneficial from the participants' perspective, but decisions concerning its effectiveness, continuation, or expansion require additional evidence. Future evaluations should incorporate objective pre- and post-program assessments of reading fluency, accuracy, vocabulary, and comprehension, as well as individual progress monitoring and

baseline-based grouping where appropriate. Greater attention should also be paid to implementation fidelity, mentoring intensity, and differentiated assistance. Such improvements would enable future evaluations to distinguish more clearly among positive participant experiences, social-affective outcomes, implementation quality, contextual support, and objectively measured changes in children's reading performance.

CONCLUSION

This study evaluated a student-led Literacy Mentoring Program implemented by the Student Association of the Community Education Study Program in Bayan Village using an adapted Goal-Free Evaluation orientation. The findings indicate that participants generally perceived the program positively across the evaluated dimensions. *Program merit/worth* received the highest score at 86.33%, followed by contextual factors at 85.78%, participants' ratings of the implementation process at 83.78%, *unintended effects* at 81.44%, and *actual effects* at 80.22%. These results suggest that the program was particularly valued for its perceived usefulness, supportive learning environment, and implementation conditions. However, the findings also reveal variation across specific outcomes. Perceived improvement in reading fluency reached only 57.33%, considerably lower than indicators related to reading interest, self-confidence, and comprehension. This difference indicates that positive experiences and perceived benefits should not be equated with measurable improvements in reading performance. Because the primary evidence was derived from post-program self-reports and no objective pre- and post-reading assessments were conducted, the study cannot establish causal effects on children's reading ability. Accordingly, the main contribution of this evaluation lies in describing participants' perceived outcomes, program benefits, implementation experiences, and contextual conditions rather than demonstrating objective effectiveness in improving reading performance.

RECOMMENDATIONS

Based on the findings, the Literacy Mentoring Program should be refined by giving greater attention to aspects that received comparatively lower ratings, particularly reading fluency and individualized assistance for children experiencing reading difficulties. Mentoring sessions may incorporate more systematic repeated-reading activities, dialogic reading strategies, graded reading materials, and individualized feedback while maintaining the supportive and engaging learning environment already valued by participants. Parental involvement should also be strengthened because family support emerged as one of the most positively perceived contextual factors. However, decisions regarding program continuation or expansion should not rely solely on participant satisfaction or perceived benefits. Future evaluations should include objective pre- and post-program assessments of reading fluency, accuracy, vocabulary, and comprehension to determine whether measurable changes occur. Instrument quality should also be strengthened through appropriate content validation and reliability testing, while outcomes related to cognitive, affective, and reading-performance dimensions should be analyzed separately. Future studies should involve larger and more diverse samples across multiple locations and, where feasible, include comparison groups and longer implementation periods. Qualitative data from interviews, open-ended questions, observations, and documentation should also be incorporated to identify genuinely unintended outcomes and provide a more comprehensive understanding of how and why the program produces particular experiences and outcomes.

BIBLIOGRAPHY

- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing*, 25(8), 652–661. <https://doi.org/10.1177/1744987120927206>
- Dong, Y., Chow, B. W.-Y., Mo, J., Miao, X., & Zheng, H.-Y. (2024). Effects of dialogic reading elements on children's language development. *Journal of Research in Reading*, 47, 181–200. <https://doi.org/10.1111/1467-9817.12447>
- Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4), 396–403. <https://doi.org/10.9734/BJAST/2015/14975>
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67–72. <https://doi.org/10.4103/aca.ACA.157.18>
- Nugroho, C. D. A., & Mawardi, M. (2024). Evaluasi program literasi membaca sekolah dasar di SD Negeri 1 Kedungjati menggunakan model Goal Free Evaluation. *Nusantara: Jurnal Pendidikan Indonesia*, 4(3), 635–644. <https://doi.org/10.14421/njpi.2024.v4i3-2>
- OECD. (2023). PISA 2022 Results (Volume I): The state of learning and equity in education. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Permatasari, A., Yuliyanti, N., & Isnaini, Z. N. (2022). Peningkatan kemampuan membaca siswa melalui kelas pendampingan literasi Program Kampus Mengajar Angkatan 2 SD Negeri 1 Sobo. *Jurnal Loyalitas Sosial: Journal of Community Service in Humanities and Social Sciences*, 4(1), 15–26. <https://doi.org/10.32493/JLS.v4i1.p15-26>
- Suryati, N., Furaidah, & Saukah, A. (2017). The effects of dialogic reading strategy on EFL young learners' reading comprehension skills. *Jurnal Ilmu Pendidikan*, 23(2), 176–182. <https://doi.org/10.17977/jip.v23i2.10992>
- Trianggoro, I. R. W., & Koeswanti, H. D. (2021). Evaluasi program Gerakan Literasi Sekolah (Gelis) di sekolah dasar. *Jurnal Ilmiah Pendidikan Profesi Guru*, 4(3), 355–362. <https://doi.org/10.23887/jippg.v4i3.40629>
- Voellinger, V. A., Lubbe, D., & Stein, L. (2023). A quasi-experimental study of a peer-assisted strategy-based reading intervention in elementary school. *Contemporary Educational Psychology*, 73, 102180. <https://doi.org/10.1016/j.cedpsych.2023.102180>
- Wigfield, A., Gladstone, J. R., & Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives*, 10(3), 190–195. <https://doi.org/10.1111/cdep.12184>
- Zakiah, Z., & Herlina, H. (2023). Efektivitas pendidikan informal pendampingan keluarga pada program literasi di Desa Kwang Rundun Kecamatan Jerowaru Kabupaten Lombok Timur 2021. *Transformasi: Jurnal Penelitian dan Pengembangan Pendidikan Non Formal Informal*, 8(2), 126–135. <https://doi.org/10.33394/jtni.v8i2.6925>