

*Playful English Teaching and Beneficiary Satisfaction in a University Community
Outreach Project in Milagro*

*Enseñanza lúdica del inglés y satisfacción de los beneficiarios en un proyecto
universitario de vinculación con la comunidad en Milagro*

*Insegnamento ludico dell'inglese e soddisfazione dei beneficiari in un progetto
universitario di servizio alla comunità a Milagro*

Edwin Alexander Pacheco Vásquez¹
Email: epachecov@unemi.edu.ec
ORCID: <https://orcid.org/0009-0006-6988-4707>

Paolo Geovanny Fabre Merchan²
Email: pfabrem@unemi.edu.ec
ORCID: <https://orcid.org/0000-0001-7457-0776>

María Noemí Alchundia Mendoza³
Email: malchundiam@unemi.edu.ec
ORCID: <https://orcid.org/0000-0002-5619-3660>

Correspondencia: Edwin Alexander Pacheco Vásquez¹ Email: epachecov@unemi.edu.ec

Artículo de investigación

* **Recibido:** 01-10-2026 * **Aceptado:** 18-10-2026 * **Publicado:** 28-10-2026

1. Universidad Estatal de Milagro, Milagro, Guayas, Ecuador.
2. Universidad Estatal de Milagro, Milagro, Guayas, Ecuador.
3. Universidad Estatal de Milagro, Milagro, Guayas, Ecuador.

Abstract

Community outreach can provide settings in which prospective teachers connect university coursework with the practical demands of teaching English. This study describes the implementation and beneficiary satisfaction of a university outreach project conducted in a parish community in Milagro, Ecuador, during 2025–2026. A retrospective descriptive design combined documentary analysis of programme records with a secondary analysis of aggregate satisfaction data. The documentary corpus included the approved project and execution contract, two individual activity logs, one individual final report, consolidated university participation records, and a subsequently developed methodological guide. The administrative records contained 38 university students across two cohorts, of whom 36 completed 96 practicum hours each. The selected documents described visual materials, songs, movement, games, guided oral exchanges and short written activities. In the satisfaction summary, each of five criteria contained 125 valid responses. Excellent ratings ranged from 82.4% for attendance to 88.0% for satisfaction of beneficiaries' needs; no Inadequate ratings were recorded. Across the five criteria, 535 of 625 item ratings were Excellent (85.6%). These results document programme delivery and a favourable appraisal of the service, rather than measured gains in English proficiency. The study identifies practical connections between playful teaching, supervised community practice and the organisation of reusable materials, while showing why satisfaction, practicum completion and language learning require separate indicators.

Keywords: English language teaching; playful learning; community outreach; beneficiary satisfaction; pre-service teacher education.

Introduction

A community English lesson requires decisions that are easy to overlook in a university classroom. Facilitators must explain a game clearly, select words that learners can understand, organise participation and provide enough support for a hesitant child to respond. When children bring different ages and levels of experience to the same setting, a single worksheet or whole-group explanation may not give everyone a reasonable opportunity to use the language. Community outreach makes these instructional decisions visible because prospective teachers have to respond to learners and to the conditions of a particular institution.

Teaching English to young learners involves more than simplifying materials originally designed for adults. Cameron (2001) provides a framework for examining how classroom activities, language use and development relate to children's learning. Pinter (2017) similarly situates instruction within children's development and considers the teaching of listening, speaking, reading and writing alongside assessment and learner agency. These perspectives support an approach in which the facilitator asks what learners will understand and do with English during an activity, rather than judging a lesson only by the number of topics covered.

Playful activities offer one way to organise such opportunities. A song may introduce recurring words; a matching game may require learners to recognise them; an exchange between partners may give those words a communicative purpose. Yet the presence of games, pictures or music does not establish that learning has occurred. The evidence needs to distinguish the resources used, the participation they supported and the language outcomes actually assessed. This distinction becomes particularly important when programme reports are expected to demonstrate both community contribution and university student training.

The project examined here, *Fortalecimiento de Comunicación Oral y Escrita en el Idioma Inglés para Niños y Niñas: Estrategias Lúdicas que Despiertan la Creatividad*, was implemented through Universidad Estatal de Milagro (UNEMI). Activities took place in the community setting of Iglesia Parroquial San Juan Bautista y Santa Madre Teresa de Calcuta in Milagro. The beneficiary institution was a parish community, rather than a school participating as the site of an ordinary English course. The project sought to create opportunities for children and adolescents to practise English through playful activities while university students undertook supervised community service.

This setting raises a practical question: what can be concluded from activity records, university participation data and a satisfaction summary when comparable pre-intervention and post-intervention language scores are unavailable? An account limited to satisfaction could overlook the instructional work behind the service. An account that treated favourable ratings as proof of language improvement would exceed the evidence. A descriptive documentary study can instead examine what was implemented and how the service was rated, while preserving the distinction between those findings and language attainment.

The objective was to describe the project's documented organisation, teaching activities and aggregate beneficiary satisfaction, and to discuss their implications for community-based English teaching and pre-service teacher preparation. Three questions guided the analysis: How was university participation organised across the two cohorts? Which instructional practices appeared in the selected records? How were the five service-evaluation criteria distributed across the available satisfaction categories? The contribution is a bounded account of one outreach experience, with an explicit separation between administrative completion, documented pedagogy and evaluation of the service.

Conceptual background

Playful activity and instructional purpose

Guided play provides a useful distinction between activity that is merely recreational and activity organised around an intended learning goal. Skene et al. (2022) reviewed guided-play interventions and found that their outcomes varied across learning domains and comparisons. Their findings do not support a universal claim that playful instruction outperforms other approaches. They do, however, justify examining how adult guidance, the activity and the learning objective are connected. In the present context, this means asking which English response a game invites and what assistance makes that response possible.

Classroom management also forms part of that connection. Carless (2002), drawing on observations in primary EFL classrooms, examined participation, mother-tongue use, noise and the role of drawing or colouring in task implementation. These issues are relevant to community teaching because an attractive activity can occupy learners without requiring much language use. A picture becomes an English-learning resource when the task asks learners to identify, ask about or describe something, and when the facilitator checks whether that response can be produced with appropriate support.

Songs warrant a similarly careful interpretation. Davis (2017) found evidence suggesting benefits for vocabulary and motivation, while noting a less developed empirical base for several other language outcomes. Songs can therefore be selected as resources for repeated exposure and participation without assuming that enjoyment demonstrates oral proficiency. The instructional question is whether the repeated language also becomes available in a short exchange, a response to an instruction or an independent attempt to name something.

Support for oral and written participation

The concept of scaffolding helps explain why guidance matters when a task initially exceeds a learner's independent performance. Wood et al. (1976) examined adult support during children's problem solving. Applied as a pedagogical lens rather than as direct evidence about this project, this perspective suggests examining the assistance supplied by a model, a gesture, a

picture or a partial sentence. Such assistance should make participation possible while leaving something for the learner to understand or produce.

For example, a facilitator may first model “It is blue,” then invite a child to choose another colour, and later ask for a response without the full model. An equivalent transition in writing might move from matching a printed word to a picture toward completing a short phrase. These are instructional possibilities, not measured learning sequences in the present dataset. Cameron (2001) treats literacy as a distinct teaching concern, which cautions against assuming that successful oral repetition automatically produces independent reading or writing.

Materials need to support these purposes. Tomlinson’s (2023) edited volume connects materials evaluation and adaptation with the development of resources for particular learners and language skills. This orientation is useful for outreach projects that rely on accessible cards, worksheets, posters and objects. Reusability is a practical advantage, but the resource still needs a clear task, comprehensible instructions and an observable response. A collection of attractive materials alone does not document either their classroom use or their effect.

Community service and evidence of learning

Service-learning research provides a framework for understanding the relationship between community contribution and teacher preparation. Chambers and Lavery (2012) describe reciprocity, real community needs and structured learning as important features, and identify professional and personal themes in prospective teachers’ reflections. Celio et al. (2011) found benefits across several student outcomes in a meta-analysis and reported associations between stronger outcomes and practices such as curricular connections, community involvement and reflection. Neither source establishes that every outreach placement produces those benefits; programme design and the evidence collected remain consequential.

For this reason, administrative completion, satisfaction and learning should be treated as different outcomes. Completion indicates that a student fulfilled a recorded requirement. Satisfaction describes an appraisal of the service. Evidence of language development requires tasks or assessments that directly examine language performance. Black and Wiliam (1998) position formative assessment around information used to adjust teaching and learning. This offers a basis for future monitoring through brief oral and written tasks, rather than using favourable service ratings as a substitute for performance evidence.

Methodology

Design and setting

The study used a retrospective descriptive documentary design with a quantitative descriptive component. It examined an existing outreach experience and an aggregate satisfaction summary; it did not introduce a new classroom intervention, construct a control

group or compare matched language-test scores. Bowen (2009) discusses the use of documents as research data and their advantages and limitations. Here, the documents were used to reconstruct selected aspects of programme implementation, with attention to the purpose and scope of each source.

The execution contract established a twelve-month project period beginning on 18 August 2025. Teaching and university participation were organised in two cohorts during 2025 and 2026. The documentary materials did not provide a fully consistent day-by-day chronology: the second-cohort activity log and individual final report differed in some dates. Accordingly, the analysis used the contractual period and the cohort years as the temporal frame, without imposing a single continuous schedule of lessons or equating the project period with uninterrupted beneficiary attendance.

The institution provided a community venue associated with the parish in Milagro. Its role was distinct from that of a formal school enrolling a fixed grade-level sample. The project addressed children and adolescents, but the available satisfaction summary did not link responses to a verified individual attendance register. Consequently, 125 valid surveys are reported as the scope of the satisfaction dataset, not as an independently verified count of unique children who attended both cohorts.

Documentary corpus and units of analysis

The corpus was purposive and bounded. It included the approved project and execution contract, two individual activity logs representing the two cohorts, one second-cohort individual final report, the consolidated university participation records and the October 2026 methodological guide. The satisfaction component used the five-criterion frequency summary provided by the project director. The selected logs and report illustrate recorded practices; they do not constitute an exhaustive set of observations from every facilitator or session.

Table 1. Sources and their role in the analysis

Source	Information used	Analytical boundary
Approved project and execution contract	Objectives, setting and contractual period	Planned commitments are not evidence of attained learning outcomes
Two activity logs and one individual final report	Recorded teaching topics, resources and reported classroom difficulties	Selected accounts do not establish how often each practice occurred across the whole project
Consolidated university participation records	Cohort, completion status and practicum hours	Administrative hours are not individual beneficiary instructional hours
Aggregate satisfaction summary	Five criteria, three response categories and 125 valid responses per criterion	No individual response vectors or breakdown by cohort were available
Methodological guide dated October 2026	Organisation of reusable activities and printable resources	A subsequent systematisation product, not an independently evaluated intervention

The analysis distinguished three units. The administrative unit was the university participant. The documentary unit was an explicit description of an activity, resource or teaching issue in the selected records. The quantitative service-evaluation unit was the response to a criterion. One survey could contribute one response to each of five criteria, so 625 item ratings do not represent 625 respondents or independent observations.

Satisfaction categories and calculations

The summary used the categories Inadequate, Adequate and Excellent. The criteria addressed attendance of teachers and university students, relationships with them, fulfilment of proposed activities, achievement of objectives and satisfaction of beneficiaries' needs. These items describe perceptions of service delivery and responsiveness. They are not a language proficiency test, nor do they directly measure the creativity mentioned in the project title.

The source identified 125 valid responses for every criterion. The analyst received category frequencies, not completed questionnaires. Information on individual administration dates, respondent type, nonresponse, assistance given during completion and the assignment of surveys to cohorts was not available. Therefore, the analysis did not estimate a response rate, describe a probability sample, calculate instrument reliability or infer differences between the two cohorts. The criterion concerning objectives was interpreted as perceived fulfilment, rather than independent verification that every language objective had been attained.

For each criterion, the three frequencies were added to verify the denominator. Category percentages were calculated as frequency divided by 125 and multiplied by 100. The percentage of Excellent ratings was checked against the supplied summary. A pooled descriptive percentage was also calculated by dividing all Excellent item ratings by all 625 item ratings. Because the response categories are ordinal and only aggregate counts were available, no mean satisfaction score, significance test or person-level composite was computed.

Documentary procedure and reporting safeguards

Documents were read for explicit references to planning, teaching resources, oral and written activity, classroom management and evaluation. Those descriptions were organised under these categories and compared across the two selected cohort logs and the individual report. The procedure was descriptive: no frequency counts of themes, inter-rater agreement or comprehensive qualitative saturation are claimed. Examples were retained where the activity was identifiable, while general statements of improvement in an individual report were not converted into quantitative language gains.

Administrative completion was checked separately from the teaching descriptions. The subsequent guide was examined to identify how the recorded topics were reorganised into a reusable resource. Its October 2026 preparation was kept separate from the earlier execution

period. The analysis thus avoided treating a later resource as proof that the same detailed activities or assessment forms had already been used during the project.

Only aggregate counts and non-identifying descriptions are presented. Children's names, addresses, photographs and individual records are excluded. The documentary source is an institutional service process, and the aggregate satisfaction data were supplied for this retrospective account. No claim is made that a new experimental research protocol, prospective ethics approval, expert instrument validation or consent procedure was documented in the materials reviewed. The authors' role in project implementation is also relevant to interpretation, particularly where the source consists of participants' own reports.

Results

University participation and completion

The participation records contained 38 university students: 25 in the first cohort and 13 in the second. Twenty-three first-cohort participants and all thirteen second-cohort participants completed the recorded requirements. Overall, 36 students completed the project participation requirement, equivalent to 94.7% of those registered. Two first-cohort records were classified as not completed. These are administrative outcomes and should not be interpreted as a measure of teaching quality or of beneficiaries' language achievement.

Table 2. University participation across the two cohorts

Indicator	Cohort 1	Cohort 2	Total
Registered university students	25	13	38
Completed participation requirements	23	13	36
Not completed	2	0	2
Recorded hours per completing student	96	96	96
Aggregate completed practicum hours	2,208	1,248	3,456

The 3,456 hours represent the sum of 96 hours for each of 36 completing university students. The activity records include preparation, induction, material development and other programme tasks as well as teaching. This total therefore cannot be read as the number of hours received by one beneficiary, the length of an individual child's course or a common instructional exposure for every survey respondent.

Teaching practices described in the records

The first-cohort log recorded greetings, numbers, colours, days of the week, emotions, animals, food and occupations, together with reinforcement sessions and material preparation. Resources included posters, flashcards, songs and worksheets. The second-cohort log recorded present simple, vocabulary, colours, animals, verbs, seasons, weather, clothing, numbers, body

parts and emotions. Its descriptions also included preparation and reinforcement activities. The individual final report added descriptions of family vocabulary, occupations and simple sentence construction; those details are attributed to that report rather than assumed to appear in every log.

Several practices linked recognition with a productive response. Visual prompts and association tasks supported the identification of words; songs and movement accompanied repetition; guided questions, short exchanges and sentence-building activities created opportunities for oral or written production. These descriptions show the instructional repertoire reported in the selected records. They do not establish a uniform teaching sequence, equal participation by all beneficiaries or independent mastery of the language practised.

The individual report identified shyness in speaking, differences in prior knowledge and distraction as teaching difficulties. It described responses involving visual support, games, group activities and closer assistance. These observations suggest that facilitators encountered heterogeneous participation needs. Because they were not collected through a standardised observation instrument or a complete set of comparable reports, their prevalence across the project cannot be estimated.

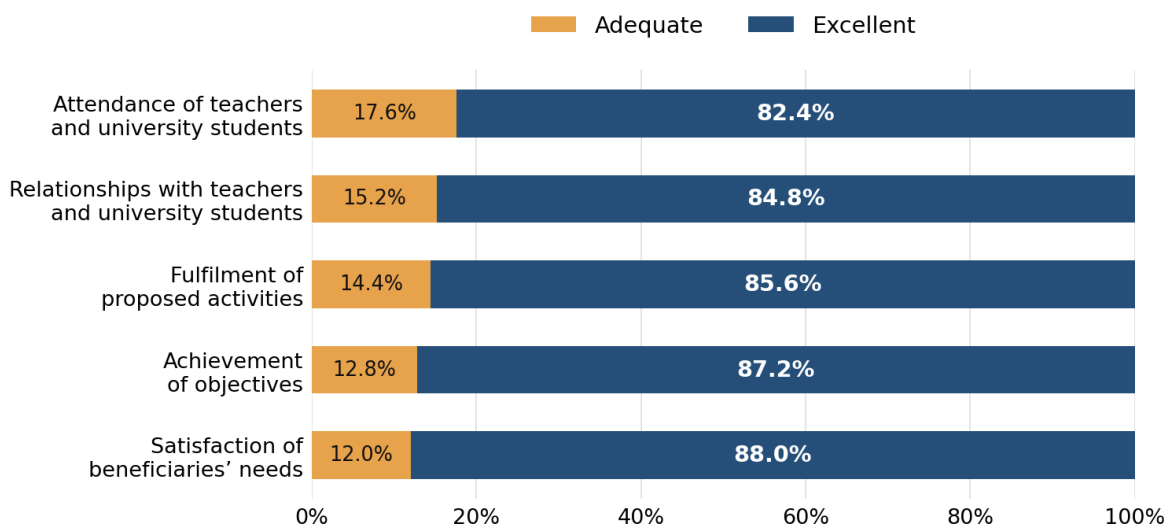
Beneficiary satisfaction

All five criteria contained 125 valid responses and no Inadequate ratings. Excellent ratings ranged from 103 to 110, corresponding to 82.4%–88.0%. Satisfaction of beneficiaries' needs received the highest proportion of Excellent ratings, followed by achievement of objectives and fulfilment of proposed activities. Attendance received the lowest proportion, although its remaining responses were all Adequate.

Table 3. Aggregate satisfaction frequencies and percentages

Criterion	Inadequate f	Adequate f	Excellent f	Valid total	Excellent %
Attendance of teachers and university students	0	22	103	125	82.4
Relationships with teachers and university students	0	19	106	125	84.8
Fulfilment of proposed activities	0	18	107	125	85.6
Achievement of objectives	0	16	109	125	87.2
Satisfaction of beneficiaries' needs	0	15	110	125	88.0
All item ratings	0	90	535	625	85.6

Note. *f* = absolute frequency. The final row pools ratings across five criteria; its denominator is 625 item responses, not 625 people. Each criterion has 125 valid responses. Source: project director's consolidated satisfaction summary.

Figure 1. Percentage distribution of Adequate and Excellent ratings by criterion.

Note. Inadequate was 0% throughout. Labels report percentages, calculated from 125 valid responses per criterion.

Across all criteria, 535 item ratings were Excellent (85.6%) and 90 were Adequate (14.4%). The highest and lowest Excellent percentages differed by 5.6 percentage points, or seven responses at the criterion level. Without linked individual questionnaires, this difference cannot be treated as a paired person-level comparison or a statistically tested ordering of service dimensions. Similarly, the pooled 85.6% does not mean that 85.6% of respondents rated every criterion Excellent.

The distribution supports a favourable appraisal of the service in the supplied summary. It does not identify why respondents selected a category, whether children and caregivers interpreted the criteria similarly, or whether the ratings changed between cohorts. No open-ended satisfaction responses were included in the summary, so explanations for the pattern cannot be derived from respondents' own comments.

Organisation of a subsequent methodological resource

The methodological guide prepared in October 2026 organised the recorded teaching themes into twelve modules with eleven printable annexes. It included facilitator guidance, suggested oral and written activities, differentiation and monitoring resources. Its structure provides a way to preserve and adapt the project's instructional ideas beyond individual logs. The guide explicitly distinguishes its newly developed activity versions from historical evidence of classroom application.

This distinction affects the interpretation of the product. The guide documents a subsequent effort to organise reusable resources. It does not show that all twelve modules were implemented as twelve historical teaching weeks, that the annexes were administered as tests, or

that the resource was evaluated or delivered during the execution period. Its practical value remains a question for future use and assessment, separate from the satisfaction results reported here.

Discussion

Instructional opportunities within the documented repertoire

The records describe teaching that combined visual support, repetition, movement and short productive tasks. This repertoire is compatible with the need to make an unfamiliar language understandable and usable in concrete activities. The useful connection is not that games automatically improve English, but that a task can give a child a reason to recognise or produce language. For instance, a number activity can move beyond repetition when learners need a partner's answer to complete a match. A colour worksheet can become more communicative when learners ask about a picture before writing a label.

Such interpretations remain pedagogical explanations rather than observed effects established by this study. Carless (2002) shows why task implementation needs attention to participation and classroom conditions, while Davis (2017) supports a cautious view of songs as potentially useful resources rather than sufficient evidence of broad language development. The project records identify resources and practices; they do not demonstrate that a particular resource caused a particular proficiency gain. This boundary preserves the value of the implementation account without turning it into an unsupported effectiveness claim.

The combination of oral and written tasks is also relevant. A learner can participate orally with substantial visual assistance yet still require support to recognise or write the same words. The documents therefore suggest a practical need for separate checks of oral response and written production. Drawing on Black and Wiliam (1998), future sessions could use a brief response to guide the next instructional decision: a learner who can identify a picture but cannot name it may need a different prompt from one who can speak the word but cannot write it. These are recommendations for monitoring, not retrospective test results.

What favourable satisfaction can and cannot establish

The largest Excellent proportion concerned satisfaction of needs, at 88.0%. This suggests that the service was favourably appraised in terms of its responsiveness, within the limits of the summary. Attendance, at 82.4%, offers a possible focus for programme review. Nevertheless, the seven-response difference between these criteria does not reveal whether a particular timetable, an absence or a communication issue explains the pattern. A follow-up instrument would need a focused comment or question to examine such explanations.

The absence of Inadequate ratings is a feature of the dataset, not evidence that every aspect of implementation was free of difficulty. The individual report described challenges involving

confidence, heterogeneous knowledge and attention. A positive appraisal of the service can coexist with those challenges. It is also possible that the categories, circumstances of administration or respondents' relationship to the programme influenced ratings, but the aggregate data do not permit those explanations to be tested. The analysis therefore reports the distribution without assigning an unverified cause to its favourable pattern.

Skene et al. (2022) also caution against treating diverse learning outcomes as interchangeable when evaluating playful approaches. Their review concerned different domains and younger populations than the full target group of this outreach project. It is relevant as a conceptual reminder to define the outcome and instructional guidance carefully, not as a benchmark that validates the present results. Here, satisfaction is the measured service outcome; changes in English proficiency and creativity remain outside what the available dataset can establish.

Teacher preparation and the continuity of community practice

The completion of 36 university students documents substantial participation in supervised community work. The logs also show that their responsibilities included planning and producing materials, not simply attending lessons. These tasks create opportunities to connect pedagogical coursework with decisions about a real group. However, a completion rate of 94.7% does not measure the quality of that connection, and a description of professional learning in one individual report cannot be generalised to all participants.

This interpretation is consistent with Chambers and Lavery (2012), who examined structured service-learning through prospective teachers' reflective accounts. Celio et al. (2011) likewise highlight the importance of design features such as reflection and curricular connections. In Milagro, a stronger evaluation of university learning would require comparable reflective tasks or supervision rubrics across cohorts. The existing records identify a context for professional learning, while a future study could examine how facilitators' explanations, adaptations and feedback change over the placement.

A related issue is continuity. When cohorts change, practical knowledge about materials, group routines and learner needs can remain dispersed across individual records. The later guide provides a possible organisational response by bringing teaching ideas into a shared resource. Tomlinson (2023) supports examining materials through evaluation and adaptation rather than equating their production with instructional success. Before wider adoption, the guide's instructions, difficulty and oral-to-written progression should be tried with learners and revised using evidence from their responses.

Limitations

The study concerns one parish community and a bounded documentary corpus. The selected activity logs and individual report were produced for programme purposes and cannot

replace systematic observation of all sessions. Some dates differed across sources, preventing a definitive reconstruction of individual beneficiary exposure. The analysis did not estimate how frequently a teaching practice occurred across the whole project or establish that every activity was implemented in the same way.

The satisfaction component relied on an aggregate summary. Original questionnaires, individual response patterns, cohort identifiers, administration details and a documented account of instrument validation were unavailable to this analysis. Response rate, scale reliability, subgroup differences and changes between cohorts could not be calculated. Because the authors participated in the project, interpretation may also reflect an insider perspective; separating recorded counts from pedagogical inference reduces, but does not remove, this limitation.

Finally, no matched language-performance dataset supported an assessment of change. The study therefore cannot estimate improvements in speaking, writing or creativity, attribute outcomes to playful instruction, or generalise the satisfaction distribution beyond the surveys summarised. Its findings remain useful as a description of implementation and service appraisal, provided those boundaries accompany any subsequent use of the results.

Conclusions and practical implications

The outreach project combined university participation with recorded opportunities to practise English through visual, playful and guided activities in a community venue. Thirty-six of 38 registered university students completed the recorded requirements. The satisfaction summary showed Excellent ratings of 82.4%–88.0% across five criteria, with 85.6% Excellent among all item ratings. These findings support a favourable descriptive appraisal of the service and document the organisation of the practicum; they do not demonstrate a measured effect on language proficiency.

The main instructional implication is to connect each resource to an observable language response. Facilitators can identify what learners should understand, say or write, demonstrate the task, and use a brief response to decide what support is needed next. Activities can remain playful while producing information that is more useful for teaching than participation or enjoyment alone. Materials shared between cohorts should include those language purposes and guidance on adaptation, rather than only a topic list.

Future programme evaluation should keep three records distinct: university practicum completion, individual beneficiary participation and language performance, and satisfaction with the service. Linking anonymous survey codes to a cohort, recording who answered the questionnaire and documenting the timing of administration would improve the interpretation of satisfaction. Comparable oral and written tasks before and after instruction would be needed for a separate study of language development. The present systematisation provides a documented starting point for those improvements without claiming outcomes that were not measured.

References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511733109>
- Carless, D. (2002). Implementing task-based learning with young learners. *ELT Journal*, 56(4), 389–396. <https://doi.org/10.1093/elt/56.4.389>
- Celio, C. I., Durlak, J., & Dymnicki, A. (2011). A meta-analysis of the impact of service-learning on students. *Journal of Experiential Education*, 34(2), 164–181. <https://doi.org/10.1177/105382591103400205>
- Chambers, D. J., & Lavery, S. (2012). Service-learning: A valuable component of pre-service teacher education. *Australian Journal of Teacher Education*, 37(4), 128–137. <https://doi.org/10.14221/ajte.2012v37n4.2>
- Davis, G. M. (2017). Songs in the young learner classroom: A critical review of evidence. *ELT Journal*, 71(4), 445–455. <https://doi.org/10.1093/elt/ccw097>
- Pinter, A. (2017). *Teaching young language learners* (2nd ed.). Oxford University Press. <https://www.oupjapan.co.jp/en/products/detail/17550>
- Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development*, 93(4), 1162–1180. <https://doi.org/10.1111/cdev.13730>
- Tomlinson, B. (Ed.). (2023). *Developing materials for language teaching* (3rd ed.). Bloomsbury Academic. <https://www.bloomsbury.com/uk/developing-materials-for-language-teaching-9781350199675/>
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>