Journal of Multidisciplinary Novel Journeys & Explorations

ISSN: 3091-1990

Vol. 4 Núm. 1 (2026)

Methodological preference by instructional purpose among teacher led explanation, non competitive digital activities and competitive gamification in English learning

Preferencia metodológica según propósito didáctico entre explicación docente, actividades digitales no competitivas y gamificación competitiva en el aprendizaje del inglés

Autores:
Rodrigo Alexander Rincón Zambrano
Universidad Estatal del Sur de Manabí
Jipijapa – Ecuador
rodrigo.rincon@unesum.edu.ec
https://orcid.org/0000-0002-2713-5111

Jahiry del Rosario Molineros Ronquillo
Universidad Estatal del Sur de Manabí
Jipijapa – Ecuador
jahiry.molineros@unesum.edu.ec
https://orcid.org/0009-0009-7013-5801

Yerson José Muñoz-Chavarría
Universidad Estatal del Sur de Manabí
Jipijapa – Ecuador
yerson.munoz@unesum.edu.ec
https://orcid.org/0009-0006-6310-3500

Vanessa Valeria Hidalgo-Tumbaco
Universidad Estatal del Sur de Manabí
Jipijapa – Ecuador
vanessa.hidalgo@unesum.edu.ec
https://orcid.org/0009-0005-2011-9900

Autor de correspondencia: *Rodrigo Alexander Rincón Zambrano*, rodrigo.rincon@unesum.edu.ec

Enviado: 13-05-2026 **Revisado:** 02-07-2026
Aceptado: 29-08-2026 **Publicado:** 30-08-2026

Doi: https://doi.org/10.63688/yfwqry56

Journal of Multidisciplinary Novel Journeys & Explorations

ISSN: 3091-1990

Vol. 4 Núm. 1 (2026)

Cómo citar este artículo:

Rincón Zambrano, R. A., Molineros Ronquillo, J. del R., Muñoz Chavarría, Y. J., & Hidalgo Tumbaco, V. V. (2026). Preferencia metodológica según propósito didáctico entre explicación docente, actividades digitales no competitivas y gamificación competitiva en el aprendizaje del inglés. *Journal of Multidisciplinary Novel Journeys & Explorations*, *4*(1), 1-24. https://doi.org/10.63688/yfwqry56

© 2026; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (https://creativecommons.org/licenses/by/4.0) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea correctamente citada.

ABSTRACT

Introduction: research on gamification in English language teaching has documented that students value digital tools positively, but has explored less which function they assign to each modality within a single instructional sequence. The aim was to describe the methodological preference among teacher-led explanation, non-competitive digital activities and competitive gamification, according to instructional purpose.

Methodology: a within-subject, cross-sectional, descriptive-comparative study with a complementary qualitative component, in a university A1-level English course (N = 20, population census) after eight weeks of methodological alternation. Comparative Likert items, three forced-choice questions and open-ended questions were combined, and the Friedman test with permutation significance, the exact Wilcoxon test and Wilson intervals were applied.

Results: agreement items did not discriminate between modalities ($\chi^2 = 3.60$; $p = .311$; Kendall's W = 0.09), with 95.3 % of responses in the two highest scale categories. Forced choice did discriminate: the combination was the majority option across the three purposes (60 %, 55 % and 55 %) and the second option reordered, from competitive gamification for motivation (25 %) to non-competitive digital activities for review (20 %). Fifty-five per cent changed their choice at least once.

Conclusions: methodological preference operated as a purpose-dependent decision rather than a stable disposition towards a tool. Agreement scales affected by ceiling effects may conceal differentiations that forced choice reveals at no additional cost. The study was conducted within research and outreach projects of the Universidad Estatal del Sur de Manabí.

Keywords: foreign language instruction; educational technology; gamification; teaching methods; higher education.

RESUMEN

Introducción: la investigación sobre gamificación en la enseñanza del inglés ha documentado que los estudiantes valoran positivamente las herramientas digitales, pero ha explorado menos qué función atribuyen a cada modalidad dentro de una misma secuencia didáctica. El objetivo fue describir la preferencia metodológica entre explicación docente, actividades digitales no competitivas y gamificación competitiva, según el propósito didáctico.

Metodología: estudio intrasujeto, transversal y descriptivo-comparativo, con componente cualitativo, en un curso universitario de inglés A1 (N = 20, censo poblacional) tras ocho semanas de alternancia metodológica. Se combinaron ítems Likert de valoración comparada, tres preguntas de elección forzada y preguntas abiertas, y se aplicaron Friedman por permutación, Wilcoxon exacta e intervalos de Wilson.

Resultados: los ítems de acuerdo no discriminaron entre modalidades ($\chi^2 = 3.60$; $p = .311$; W de Kendall = 0.09), con el 95.3 % de las respuestas en las dos categorías superiores. La elección forzada sí discriminó: la combinación fue mayoritaria en los tres propósitos (60 %, 55 % y 55 %) y la segunda opción se reordenó, de la gamificación competitiva para motivar (25 %) a las actividades no competitivas para repasar (20 %). El 55 % modificó su elección al menos una vez.

Conclusiones: la preferencia metodológica operó como una decisión dependiente del propósito y no como una disposición estable. Las escalas de acuerdo con efecto techo pueden ocultar diferenciaciones que la elección forzada sí revela. El estudio se desarrolló en el marco de proyectos de investigación y vinculación de la Universidad Estatal del Sur de Manabí.

Palabras clave: enseñanza de lenguas extranjeras; tecnología educativa; gamificación; método de enseñanza; educación superior.

1. INTRODUCTION

The incorporation of digital tools into the teaching of English as a foreign language has ceased to be a matter of availability and has become a matter of judgement. Latin American university classrooms today combine teacher-led explanation supported by slides, autonomous practice platforms, and gamified quizzes with real-time scoring and leaderboards, so that the problem faced by teaching staff no longer consists of deciding whether to use technology, but which technology to use, at which point of the instructional sequence and for what purpose. Recent literature agrees that the value of these resources does not lie in the tool itself but in its pedagogical articulation: digital technologies expand opportunities for interaction and practice when they are integrated with clear objectives, and contribute little when they are used as isolated activities (Bui, 2022; Moorhouse & Yan, 2023; Sim & Ismail, 2023). In the regional context this warning acquires an additional nuance, since introducing technology without a critical reading of the setting may deepen inequalities rather than reduce them (UNESCO, 2023; Torres Duque & Argudo Garzón, 2024).

Within this landscape, gamification, understood as the use of game design elements in non-game contexts (Deterding et al., 2011), occupies a prominent place. The available meta-analyses report favourable effects on intrinsic motivation and on perceptions of autonomy and relatedness, although with a considerably smaller impact on perceived competence and with notable variability depending on the design of the activity (Li et al., 2024; Sailer & Homner, 2020). In English as a foreign language instruction, systematic reviews describe a similar picture: benefits for participation and attitudes, conditioned by the type of task, by the balance between play and content and, especially, by the way the competitive component is administered (Lo, 2024; Zhang & Hasim, 2022). Empirical work along these lines shows that both digital and non-digital gamification may affect linguistic knowledge and the perception of flow depending on the design of the activity (Foroutan Far & Taghizadeh, 2024), that game-based learning may increase learners' enjoyment and intrinsic motivation (Zhou, 2024), and that gamified settings foster concentration and language enjoyment in university English courses (Xiao, 2025). Studies focused on specific platforms point in the same direction and recommend using them as part of a blended approach combining teacher-

Doi: https://doi.org/10.63688/yfwqry56

led explanation and appropriate time management, rather than as a substitute for direct instruction (Permana et al., 2023; Quyen & Nguyen, 2025).

A considerable part of this literature, however, has been built upon a binary question: whether or not students value gamification. The answer, repeated across dozens of perception studies, is affirmative and stable, so that the question is quickly exhausted and leaves unresolved another one that is more useful for instructional decision-making. When the same group has worked with several modalities within a single course, what function does it assign to each of them? A student may hold that gamified competition provides motivation, that the teacher's explanation enables comprehension, and that a digital activity without time pressure is more useful for review, without any of these statements contradicting the others. That distribution of functions remains invisible when each methodology is evaluated separately by means of an agreement scale, because the format allows the respondent to agree with all of them at once.

The problem also has a seldom discussed methodological component. Perception studies on educational technology tend to produce strongly skewed distributions, with most responses concentrated in the upper values of the scale. This skewness is not exclusive to the field: recent psychometric research has shown that idiosyncratic response biases systematically contaminate self-reported measures and that standard analytic procedures do not control for them (Grimmond et al., 2025). When this occurs, the means of all items are compressed into a narrow band and the differences between them cease to be informative, even though they continue to be reported and interpreted as if they were. An instrument combining agreement scales with forced-choice questions makes it possible to verify directly whether the apparent homogeneity of the ratings holds when students are required to establish priorities, and turns a common limitation of this type of study into an object of analysis. The debate on the superiority of the forced-choice format is far from settled, since there is evidence that it provides information equivalent to that of agreement scales in the measurement of personality traits (Bäckström & Björklund, 2024) and also evidence that it improves discrimination in survey design (Combrinck, 2024), an additional reason to contrast both formats within the same sample.

The rationale for the present work rests on these two gaps. First, describing the function that a specific group assigns to each modality offers teachers actionable information

that the mere confirmation of a positive rating does not provide, because it guides the decision as to where each resource should be placed within the classroom sequence. Second, contrasting two response formats within the same sample provides evidence on the discriminative capacity of the instruments routinely used in this line of research. Both contributions are relevant in beginner-level courses, where learning vocabulary and grammatical structures requires repetition, guided practice and frequent retrieval of information.

The general objective of the study was to describe the methodological preference of a group of university English students among teacher-led explanation with slides, non-competitive digital activities and competitive gamification, according to the instructional purpose of each choice. The specific objectives were to compare the rating of the three modalities by means of agreement items and to determine whether the instrument discriminates among them; to analyse the distribution of preference when a forced choice is requested for three different purposes, namely motivation, participation and language review; and to examine the individual consistency of that preference across the three purposes.

The variables considered were, on the one hand, the declared rating of each modality, measured on a five-point ordinal scale; on the other, methodological preference, measured as a nominal variable with four categories across three choice situations; and, as a derived grouping variable, the consistency of preference, defined as whether or not the chosen modality coincided across the three purposes. Age and gender were also included as descriptive characterisation variables. The study was conducted as a within-subject descriptive-comparative investigation, cross-sectional in design and with a complementary qualitative component, without manipulation of variables or measurement of learning outcomes, so that its findings describe perceptions and declared decisions within a delimited context.

This study forms part of the research project "Fortalecimiento del nivel didáctico en la formación docente en las instituciones públicas del cantón Jipijapa" and of the outreach project "Tareas dirigidas y apoyo psicopedagógico para fortalecer el aprendizaje de los estudiantes en la Educación Básica Pública de Jipijapa", both attached to the Universidad Estatal del Sur de Manabí, within whose framework the teaching experience described here was designed and implemented.

2. METHODOLOGY

The research was conducted under a quantitative approach of descriptive-comparative scope, with a complementary qualitative component of subordinate status, which corresponds to a convergent mixed design with quantitative dominance (Creswell & Creswell, 2023). The design is cross-sectional, because the information was collected at a single point in time, and within-subject, because the same participants rated and chose among the three modalities, so that each student acts as their own control and the variability attributable to between-group differences is eliminated. There was no manipulation of variables, random assignment or comparison group, and no language achievement tests were administered; consequently, no hypotheses were formulated and the statistical tests serve an exploratory function and a control function regarding the discriminative capacity of the instrument, rather than a confirmatory one.

The population consisted of the students enrolled in the A1-level English course of the Common European Framework of Reference in which the experience took place, corresponding to the second level of the institutional programme. A population census was used rather than a probabilistic sample: all 20 enrolled students completed the questionnaire, which places the response rate at 100 % and eliminates non-response bias, although not the limitation derived from the small size of the group. The inclusion criteria were being enrolled in the course, having participated in the three modalities of work during the period of the experience, and expressly agreeing to participate; no exclusions or incomplete responses were recorded, so the 20 cases were analysed without imputation of missing data.

The study was carried out at an Ecuadorian public university, in a face-to-face general English course belonging to the institution's foreign language training provision. The classroom was equipped with a projector and an internet connection, and students accessed the digital platforms from their own mobile devices, a condition that is relevant for interpreting the results, since device accessibility subsequently emerged as one of the declared reasons for preference. Over eight weeks the teacher alternated three modalities of work in a planned manner: teacher-led explanation with slides and guided exercises, used for the presentation of lexical and grammatical content; non-competitive digital activities, without scoring or time limits, developed with resources such as Educaplay, interactive

crosswords and video quizzes, intended for individual practice; and competitive gamified activities through Blooket and Quizizz, with scoring, a timer and a visible leaderboard, used for the review and retrieval of previously covered content. The three modalities were used recurrently throughout the period, so that all participants had experienced each of them on multiple occasions before responding.

The instrument was a structured self-administered questionnaire, hosted in an online form and developed ad hoc for this experience on the basis of the recurrent dimensions in the literature on gamification and digital tools in English language teaching. It comprised a consent section, general characterisation data, twenty five-point Likert scale items (1 = strongly disagree; 5 = strongly agree), three forced-choice questions and three open-ended questions. The present article analyses exclusively the components linked to the comparison between modalities: the five scale items that comparatively rate the three modalities and their combination; the three forced-choice questions, which offered as options teacher-led explanation, non-competitive digital activities, competitive gamification and the combination of the three, referring respectively to which methodology is most motivating, with which one students participate most, and which is most useful for reviewing vocabulary and grammar; and the open-ended question asking which of the three methodologies helps most in learning English and requesting a justification. The two remaining open-ended questions, concerning the most valued aspects of competitive platforms and suggestions for improvement, are not analysed in this work.

Verification of the instrument was approached through descriptive indicators of internal consistency, with the caution imposed by the size of the group. The complete set of twenty items yielded a Cronbach's alpha of .965. This value is not interpreted as evidence of quality: a coefficient above .90 indicates redundancy among items rather than psychometric soundness and, with twenty cases, its estimation is markedly unstable, besides being compatible with the presence of a single response factor associated with acquiescence. The block of five items analysed in this work obtained an alpha of .730, and the corrected item-total correlations ranged from .277 to .686; the lowest value corresponded to the item referring to non-competitive digital activities, which was also the one with the greatest dispersion. These coefficients are reported for transparency and not as validation, since the block is not composed of parallel indicators of a single construct but of ratings of

objects that differ from one another, a condition under which alpha is not the appropriate statistic. The instrument was not submitted to content validation by expert judgement or to a prior pilot test, and the size of the group makes any analysis of internal structure inadvisable; both limitations are explicitly declared.

The procedure was developed in four phases. In the first, the alternation of the three modalities across the eight weeks was planned and the questionnaire was designed. In the second, once the period of the experience had concluded, authorisation was requested from the programme coordination and the group was informed about the objective of the study, the intended use of the data, the voluntary nature of participation, its independence from grading and the right not to respond. In the third phase the questionnaire was administered online; the form recorded the institutional e-mail address for the sole purpose of preventing duplicate responses. In the fourth phase the database was downloaded, dissociated by replacing identifiers with numerical codes, and cleaned, verifying the range of admissible values for each variable and the absence of incomplete cases before analysis.

Data analysis was performed with Python 3, using the pandas, numpy and scipy libraries. For the scale items, the mean, standard deviation, median, frequency distribution and percentage of agreement were calculated, the latter defined as the sum of responses 4 and 5. In order to test whether the instrument discriminates between modalities, the Friedman test was applied to the three items that directly rate each of them, accompanied by Kendall's W as a measure of effect size, and pairwise comparisons were carried out by means of the Wilcoxon signed-rank test. The Friedman statistic is reported with the tie correction, a procedure that is appropriate given the high proportion of identical values within each block and consistent with the implementation used by standard statistical packages. Given the size of the group, the significance of the Friedman test was obtained by within-subject permutation of ranks with 20,000 replications, and the pairwise comparisons were resolved by exact enumeration of all sign combinations, thus avoiding the asymptotic approximation, which is not reliable with twenty cases. It should be noted that the tie correction factor remains constant under within-subject permutation, so that the p value obtained through this procedure is identical whether or not the correction is applied; the correction affects the magnitude of the statistic and of the effect size, not the significance. For the forced-choice questions, frequencies, percentages and 95 % confidence intervals were calculated using the Wilson

method, which is appropriate for proportions with small samples, and a table of individual profiles was constructed that makes it possible to observe whether each participant maintained or modified their choice across the three purposes. A significance level of .05 was adopted, with the express warning that the statistical power of this design is very low and that the absence of significance is not equivalent to a demonstration of equivalence. The open-ended question was analysed through thematic categorisation in two successive readings, one to classify the modality indicated and another to identify the reasons adduced, with non-exclusive categories; coding was carried out by a single coder, so no inter-rater agreement index is available.

As regards ethical considerations, the study was conducted in accordance with the principles of voluntary participation, informed consent and confidentiality applicable to research with human participants. Informed consent was obtained through an initial question in the questionnaire in which each student had to expressly accept participation before accessing the remaining items; only those cases that granted such consent were included in the analysis, and the group was informed beforehand about the objective of the study, the intended use of the data, the voluntary nature of participation, its independence from grading and the right not to respond or to interrupt completion at any moment.

The treatment of the information was confidential and not anonymous at the moment of collection: the form recorded the institutional e-mail account for the sole technical purpose of preventing duplicate responses, and at no point were responses associated with an identified participant consulted or examined. That field was replaced by numerical codes before processing, so that the entire analysis was performed on an anonymised file. The data were stored in a password-protected digital file, hosted in a restricted-access repository to which only the research team signing the article had access; the records will be kept for five years from the date of publication and will be permanently deleted at the end of that period. Results are reported only in aggregate form and the open-ended responses are presented without any element that would allow individual identification.

Administration of the instrument was authorised by the programme coordination of the institution. Given that the study involved no intervention on the participants, collected no sensitive data or health information, included no minors, and was limited to recording perceptions and declared preferences regarding ordinary teaching activities by

means of a questionnaire with confidential treatment and anonymised analysis, it did not require assessment by a research ethics committee for human subjects in accordance with the institutional regulations in force; nevertheless, the principles of voluntariness, confidentiality, non-maleficence and respect for participants were safeguarded at all times. It is expressly declared that the principal investigator was the teacher of the course, a condition that introduces a power asymmetry liable to favour socially desirable responses; in order to mitigate it, administration took place outside any assessment context, students were informed that their responses would not affect their grades, and at no point were responses linked to academic records. This condition is taken up again in the discussion because of its relevance for interpreting the distribution of the data.

3. RESULTS

The results are organised into four blocks. The first describes the overall distribution of responses on the agreement scale, a finding that conditions the interpretation of everything that follows. The second presents the comparative rating of the three modalities and the statistical contrast of their differences. The third sets out the preference declared in the three forced-choice situations and the analysis of individual profiles. The fourth summarises the reasons adduced in the open-ended question. The characterisation of participants is presented beforehand in Table 1.

Table 1.

Characterisation of participants (N = 20)

Variable	Category	f	%
Age	18–23 years	12	60.0
	28–31 years	4	20.0
	32–35 years	1	5.0
	35–38 years	3	15.0
Gender	Female	13	65.0
	Male	6	30.0
	Prefers not to answer	1	5.0
Course level	A1 (CEFR)	20	100.0

Note. the age ranges reproduce the options offered by the form, which overlapped between the 32–35 and 35–38 brackets and omitted the 24–27 interval; the data are reported exactly as collected and this limitation is addressed below. Own elaboration based on the results of the questionnaire administered.

Doi: https://doi.org/10.63688/yfwqry56

The first result does not concern the content of the responses but their distribution. Of the 400 responses recorded on the Likert scale (20 items by 20 participants), 381, that is, 95.3 %, fell within the two highest categories. Option 2 was never selected and option 1 was used only once in the entire dataset, as shown in Table 2. This concentration compresses the means of the twenty items into a band of barely 0.35 points, between 4.40 and 4.75, with a standard deviation between means of 0.087; consequently, any ranking of items based on differences of one or two hundredths lacks support and is not reported.

Table 2.

Overall distribution of responses on the agreement scale (400 responses)

Response option	f	%	Cumulative %
5 = Strongly agree	252	63.0	63.0
4 = Agree	129	32.3	95.3
3 = Neither agree nor disagree	18	4.5	99.8
2 = Disagree	0	0.0	99.8
1 = Strongly disagree	1	0.3	100.0

Note. calculation based on the 20 items of the instrument answered by the 20 participants. Own elaboration based on the results of the questionnaire administered.

Table 3 presents the five items that comparatively rate the modalities. All three receive high and practically indistinguishable ratings. Competitive gamification and teacher-led explanation obtain the highest values and reach the same number of maximum-agreement responses. Non-competitive digital activities show the lowest mean and, above all, the greatest dispersion of the block, with a standard deviation of 0.94: this is the only item of the instrument that received a strong disagreement response.

Table 3.

Comparative rating of the modalities of classroom work (N = 20)

Item	M	SD	Mdn	Agreement (4–5)	Maximum (5)
1. Teacher-led explanation with slides helps comprehension	4.70	0.57	5	19 (95 %)	15 (75 %)
2. Non-competitive digital activities help practice	4.45	0.94	5	19 (95 %)	12 (60 %)

Doi: https://doi.org/10.63688/yfwqry56

Item	M	SD	Mdn	Agreement (4–5)	Maximum (5)
3. Activities with Blooket or Quizizz help reinforce what has been learned	4.75	0.44	5	20 (100 %)	15 (75 %)
4. Comprehension improves when the three modalities are combined	4.55	0.60	5	19 (95 %)	12 (60 %)
5. Digital activities generate motivation	4.55	0.60	5	19 (95 %)	12 (60 %)

Note. scale from 1 to 5. Item wording abbreviated. Agreement corresponds to the sum of responses 4 and 5. Own elaboration based on the results of the questionnaire administered.

The Friedman test applied to the three items that rate each modality confirms that these differences are not statistically distinguishable: $\chi^2(2) = 3.60$; $p = .311$, with significance obtained by permutation over 20,000 replications, and a very small effect size (Kendall's W = 0.09). The mean ranks were practically identical: 2.08 for teacher-led explanation, 2.08 for competitive gamification and 1.85 for non-competitive activities. The pairwise comparisons point in the same direction and reveal the extent to which the scale lacks discriminative capacity in this group: between competitive gamification and teacher-led explanation only 2 of the 20 participants assigned different scores ($p = 1.000$); between gamification and non-competitive activities, 3 of 20 ($p = .250$); and between teacher-led explanation and non-competitive activities, 5 of 20 ($p = .500$). In other words, between 75 % and 90 % of the students gave exactly the same score to modalities that the literature describes as pedagogically different. This result should not be interpreted as evidence that the three modalities are equivalent for students, but as confirmation that an agreement scale applied to this group captures no differentiation among them.

The picture changes when students are required to choose. Table 4 presents the distribution of the three forced-choice questions with their confidence intervals, and Figure 1 represents it graphically. Two findings emerge from them.

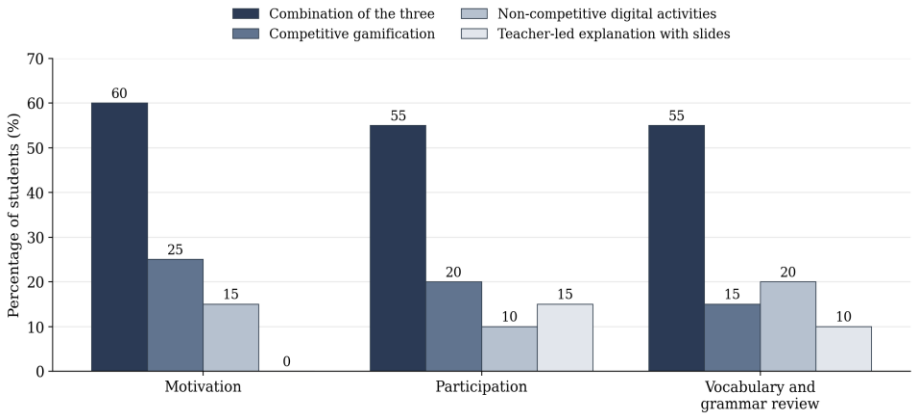
Table 4. Methodological preference by instructional purpose and 95 % confidence intervals (N = 20)

Modality chosen	Most motivating: f (%) [95 % CI]	Greatest participation: f (%) [95 % CI]	Most useful for review: f (%) [95 % CI]
Combination of the three	12 (60.0) [38.7–78.1]	11 (55.0) [34.2–74.2]	11 (55.0) [34.2–74.2]
Competitive gamification	5 (25.0) [11.2–46.9]	4 (20.0) [8.1–41.6]	3 (15.0) [5.2–36.0]
Non-competitive digital activities	3 (15.0) [5.2–36.0]	2 (10.0) [2.8–30.1]	4 (20.0) [8.1–41.6]
Teacher-led explanation with slides	0 (0.0) [0.0–16.1]	3 (15.0) [5.2–36.0]	2 (10.0) [2.8–30.1]
Total	20 (100)	20 (100)	20 (100)

Note. 95 % confidence intervals calculated using the Wilson method. Response labels were normalised in order to group equivalent wordings across questions. Own elaboration based on the results of the questionnaire administered.

Figure 1.

Methodological preference by instructional purpose



Note. percentage of students who chose each modality in each of the three forced-choice situations (N = 20). Own elaboration based on the results of the questionnaire administered.

The first finding is the primacy of the combination: across the three purposes, more than half of the participants did not choose a specific modality but rather their integration. The result is consistent with item 4 of Table 3 and suggests that the group perceives the instructional sequence as a unit and not as a set of interchangeable resources. It should be noted, however, that the confidence intervals are wide, as corresponds to the size of the group, so that the exact magnitude of that majority must be read with caution.

The second finding is more specific and constitutes the central contribution of this work: the second option reorders according to purpose. When students are asked about motivation, the modality following the combination is competitive gamification, with 25 %, and

teacher-led explanation receives no choices at all. When they are asked about participation, gamification retains second place with 20 %, but teacher-led explanation reappears with 15 %, above non-competitive activities. And when they are asked about usefulness for reviewing vocabulary and grammar, the order is reversed: non-competitive digital activities move into second place with 20 %, surpassing competitive gamification, which drops to 15 %. The modality that the group associates with motivation is not, therefore, the same one it associates with language consolidation.

The aggregate distributions could be compatible with two different scenarios: that each student has a stable preference and that the differences between columns are due to a few cases, or that a considerable proportion of the group changes option according to purpose. The analysis of individual profiles, presented in Table 5, makes it possible to distinguish between them. Nine participants (45.0 %; 95 % CI [25.8–65.8]) chose the same modality in the three questions: eight of them always opted for the combination and one always for competitive gamification. The remaining eleven (55.0 %; 95 % CI [34.2–74.2]) modified their choice at least once according to the purpose posed.

Table 5.

Individual choice profiles across the three purposes

Most motivating	Greatest participation	Most useful for review	f
Combination	Combination	Combination	8
Gamification	Gamification	Gamification	1
Combination	Combination	Gamification	1
Combination	Gamification	Combination	1
Combination	Gamification	Non-competitive	1
Combination	Non-competitive	Combination	1
Gamification	Combination	Combination	1
Gamification	Teacher-led explanation	Gamification	1
Gamification	Teacher-led explanation	Non-competitive	1
Gamification	Gamification	Teacher-led explanation	1
Non-competitive	Combination	Non-competitive	1
Non-competitive	Teacher-led explanation	Non-competitive	1
Non-competitive	Non-competitive	Teacher-led explanation	1

Note. in bold, the profiles of constant preference (9 participants). The remaining eleven rows correspond to participants who modified their choice at least once depending on the purpose. Own elaboration based on the results of the questionnaire administered.

The group with constant preference is composed almost entirely of those who always choose the combination, which admits two readings: a pedagogical conviction according

to which all modalities are necessary, or else a low-differentiation response equivalent to the ceiling effect observed on the scale. Among those who vary, the most frequent pattern consists of reserving competitive gamification for the functions of motivation or participation and shifting towards teacher-led explanation or non-competitive activities when the purpose is language review. No participant followed the reverse trajectory.

The open-ended question asked students to indicate which methodology helps most in learning English and to justify their choice. The twenty responses were distributed as follows: seven pointed to competitive gamification; six, to the combination of the three or the use of several methodologies; two, to non-competitive digital activities; one, to teacher-led explanation; one answered generically "digital" without specifying; and three produced answers that did not allow a modality to be identified. It is significant that this distribution does not reproduce that of the equivalent closed question: when the combination appears explicitly as a selectable option it is chosen by between 55 % and 60 % of the group, whereas when it must be formulated spontaneously only 30 % mention it and competitive gamification moves into first place. The justification categories are presented in Table 6.

Table 6.

Categories of justification in the open-ended question

Category	f	Content
Comprehension and perceived learning	11	Justifications centred on learning, comprehending or understanding the subject better, without reference to the playful dimension.
Variety and multimodality	3	References to the possibility of seeing, reading and practising simultaneously and to not depending on a single tool.
Dynamism and entertainment	3	References to the activity being more fun, dynamic or entertaining.
Attention and participation	2	Mentions of staying attentive, listening and participating better during class.

Doi: https://doi.org/10.63688/yfwqry56

Category	f	Content
Error recognition without pressure	1	Explicit preference for the non-competitive modality because it allows errors to be identified without the stress associated with competing.
Device accessibility	1	Choice determined by the possibility of participating from a mobile phone.

Note. non-exclusive categories: a single response could express more than one reason. Own elaboration based on the results of the questionnaire administered.

The most consistent finding is that the dominant justification is not playful but cognitive: eleven of the twenty responses explain the choice by appealing to comprehension or learning, and only three to entertainment. Even among those who chose competitive gamification, the most frequent reason was that it helps them remember, understand the topic or test what they have learned. One response illustrates precisely the logic of functional differentiation described in the preceding sections by justifying the preference for the non-competitive modality on the grounds that it allows students to see their own errors without the stress associated with competing.

4. DISCUSSION

The results support three claims, with differing degrees of robustness. The first is methodological in nature. In this group the agreement scale did not discriminate between pedagogically distinct modalities, whereas the same students did distribute themselves in a differentiated manner when they were required to choose. The contrast has a practical implication for research on the perception of educational technology: studies relying exclusively on agreement scales run the risk of concluding that students value everything positively, which is a trivial conclusion, when what actually happens is that the response format does not allow priorities to be expressed. Including forced-choice questions, ranking items or constant-sum formats constitutes a low-cost design decision with a considerable informational return, consistent with the evidence attributing greater discriminative capacity to the ipsative format (Combrinck, 2024). It should be noted, however, that this superiority is not universal: Bäckström and Björklund (2024) show that, in the measurement of

personality traits, the forced-choice format provides information equivalent to that of agreement scales, including information relating to social desirability. What the present study documents, therefore, is not a general advantage of the format but its specific usefulness in a context of pronounced ceiling effect.

The ceiling effect observed is not a sampling accident but a phenomenon with identifiable causes that deserve to be made explicit. The twenty items are worded in a positive direction, with no reverse items to interrupt automatic responding; the instrument was administered by the teacher to their own students, with the institutional e-mail address recorded, which activates social desirability in a context of power asymmetry; and consent was resolved through a dichotomous item. Any of these factors would suffice to explain part of the skewness observed, and recent psychometric literature warns that idiosyncratic response biases are not corrected by standard analytic procedures (Grimmond et al., 2025). Acknowledging this does not weaken the forced-choice results, which are less vulnerable to this bias precisely because they do not admit generalised agreement.

The second claim concerns the primacy of the combination. Across the three purposes the majority of the group chose the integration of the three modalities over any of them separately. The result is consistent with the line of research maintaining that the value of digital tools depends on their pedagogical articulation and not on their intrinsic properties (Bui, 2022; Moorhouse & Yan, 2023; Sim & Ismail, 2023), as well as with studies on specific platforms that recommend their use within a blended approach combining teacher-led explanation and appropriate time management (Permana et al., 2023; Quyen & Nguyen, 2025). The strength of the finding should nevertheless be qualified. The comparison between the closed question and the open-ended question showed that the preference for the combination is halved when the option is not available in the form, so that it is plausible that part of that 55–60 % reflects the convenience of an inclusive alternative rather than an elaborated pedagogical position. The confidence intervals obtained, wide in every case, reinforce this caution.

The third claim is the most specific and the one that provides new information. Methodological preference did not behave in this group as a stable trait of the student but as a purpose-dependent decision: 55 % modified their choice at least once across the three questions and the aggregate reordering followed a coherent direction. Competitive

gamification occupied second place for motivation and for participation, but yielded that position to non-competitive digital activities when the purpose was reviewing vocabulary and grammar, without any participant making the reverse shift.

This functional differentiation can be interpreted in the light of self-determination theory (Ryan & Deci, 2020). Competitive elements, such as points, the timer and the leaderboard, generate activation and a sense of challenge, conditions that favour motivation and immediate participation; recent evidence in university English classrooms shows that gamified settings improve concentration and language enjoyment (Xiao, 2025). Yet those same elements introduce a time pressure that competes with consolidation processes, which require time to review, correct and try again. The only open-ended response that explicitly opted for the non-competitive modality did so precisely in these terms. The available meta-analyses point in the same direction when they report consistent effects of gamification on intrinsic motivation and minimal effects on perceived competence (Li et al., 2024), and reviews of gamification in English language teaching warn that poorly regulated competition may limit the benefits of the activity (Lo, 2024; Zhang & Hasim, 2022; Jaramillo-Mediavilla et al., 2024).

For teaching practice the implication is more concrete than the usual recommendation to combine methodologies. If students assign different functions to different resources, the relevant decision does not consist of determining how much gamification to incorporate but at which point of the sequence to place it. The data from this group suggest using gamified competition at moments of activation, such as the opening of a session, the retrieval of previous content or a dynamic closing, and using digital activities without time pressure in the consolidation phases, where students need to detect and correct errors. Teacher-led explanation, which received no choices as the most motivating modality but reappeared with 15 % when participation was asked about and with 10 % when review was asked about, retains a place that no digital tool occupied in this group, which is consistent with warnings about the irreplaceable role of teacher mediation in technology integration (Torres Duque & Argudo Garzón, 2024; UNESCO, 2023).

Finally, it is worth noting that the spontaneous justifications were predominantly cognitive rather than playful. Eleven of twenty responses explained the choice in terms of understanding or learning better and only three in terms of enjoyment. This finding qualifies the frequent representation of gamification as an essentially motivational resource:

at least in the reasons these students declare, the platform is valued for what it allows them to verify and remember rather than for what it allows them to enjoy.

The results must be read with caution for six reasons that should be declared without mitigation. First, the size of the group: twenty participants from a single course, so that a difference of five percentage points corresponds to a single student and no result is generalisable beyond the context described. Second, the absence of validity evidence for the instrument, developed ad hoc, without expert judgement, without a pilot test and without analysis of internal structure, which the size of the group makes inadvisable. Third, the ceiling effect and social desirability derived from administration by the teacher and from the recording of the institutional e-mail address. Fourth, construction flaws in the instrument, in particular the overlap and omission of brackets in the age ranges and the non-identical wording of the options across the three forced-choice questions, which required label normalisation. Fifth, qualitative coding by a single coder, without an inter-rater agreement index. And sixth, the self-reported nature of the data, without measurement of actual learning, which precludes deriving from these results any claim regarding language achievement.

5. CONCLUSION

The agreement scale items did not discriminate between the three modalities analysed. Some 95.3 % of the responses were concentrated in the two highest categories, the Friedman test detected no differences between teacher-led explanation, non-competitive digital activities and competitive gamification ($\chi^2(2) = 3.60$; $p = .311$), and the effect size was very small ($W = 0.09$), so that any reading based on the ranking of the means of these items would lack foundation.

The forced-choice questions, by contrast, did produce an interpretable pattern. The combination of the three modalities was the majority option across the three purposes, but the second option reordered according to purpose, with competitive gamification for motivating and for participating and non-competitive digital activities for reviewing vocabulary and grammar, and only 45 % of the participants maintained the same choice across the three questions. In this group, therefore, methodological preference operated as a purpose-dependent decision and not as a general disposition towards a tool.

Doi: https://doi.org/10.63688/yfwqry56

For future lines of study it is recommended to replicate the design with larger samples and across several courses, incorporating reverse items and administration by a third party external to the teaching in order to mitigate the ceiling effect; to contrast declared preference with objective measures such as time on task, number of attempts or scores on vocabulary tests administered before and after; and to submit to experimental testing the functional differentiation hypothesis that this work only suggests, verifying whether the sequence of application, with competition in activation and pressure-free practice in consolidation, produces measurable differences in lexical retention.

The finding transcends the case of English language teaching and points to a broader discussion on the assessment of student perception in educational technology. If agreement formats tend to generate uniformly favourable responses, a substantial part of the accumulated evidence on the acceptance of digital tools may be describing a property of the instrument rather than a property of the phenomenon. This issue affects equally other areas of teaching innovation in which positive appraisal has become an expected outcome, and suggests the advisability of systematically incorporating formats that require priorities to be established.

In prospective terms, if the functional differentiation observed is confirmed in larger samples, the design of technology-mediated foreign language classes could move towards instructional architectures that assign each resource a specific position within the sequence, instead of accumulating resources for their appeal. The foreseeable consequence would not be greater use of gamification, but a more selective use, in which competition is reserved for the functions that students recognise in it and activities without time pressure occupy the phases in which errors need to be identified and corrected.

BIBLIOGRAPHIC REFERENCES

Bui, T. H. (2022). English teachers' integration of digital technologies in the classroom. International Journal of Educational Research Open, 3, Article 100204. https://doi.org/10.1016/j.ijedro.2022.100204

Bäckström, M., & Björklund, F. (2024). Why forced-choice and Likert items provide the same information on personality, including social desirability. Educational and

Psychological Measurement, 84(3), 549–576. https://doi.org/10.1177/00131644231178721

Combrinck, C. (2024). Not liking the Likert? A Rasch analysis of forced-choice format and usefulness in survey design. SAGE Open, 14(4), Article 21582440241295501. https://doi.org/10.1177/21582440241295501

Creswell, J. W., & Creswell, J. D. (2023). Research design: Qualitative, quantitative, and mixed methods approaches (6th ed.). SAGE.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, 9–15. https://doi.org/10.1145/2181037.2181040

Foroutan Far, F., & Taghizadeh, M. (2024). Comparing the effects of digital and non-digital gamification on EFL learners' collocation knowledge, perceptions, and sense of flow. Computer Assisted Language Learning, 37(7), 2083–2115. https://doi.org/10.1080/09588221.2022.2146724

Grimmond, J., Brown, S. D., & Hawkins, G. E. (2025). A solution to the pervasive problem of response bias in self-reports. Proceedings of the National Academy of Sciences, 122(3), Article e2412807122. https://doi.org/10.1073/pnas.2412807122

Jaramillo-Mediavilla, L., Basantes-Andrade, A. V., Cabezas-González, M., & Casillas-Martín, S. (2024). Impact of gamification on motivation and academic performance: A systematic review. Education Sciences, 14(6), Article 639. https://doi.org/10.3390/educsci14060639

Li, L., Hew, K. F., & Du, J. (2024). Gamification enhances student intrinsic motivation, perceptions of autonomy and relatedness, but minimal impact on competency: A meta-analysis and systematic review. Educational Technology Research and Development, 72(2), 765–796. https://doi.org/10.1007/s11423-023-10337-7

Lo, N. (2024). Enhancing EFL/ESL instruction through gamification: A comprehensive review of empirical evidence. Frontiers in Education, 9, Article 1395155. https://doi.org/10.3389/feduc.2024.1395155

Moorhouse, B. L., & Yan, L. (2023). Use of digital tools by English language schoolteachers. Education Sciences, 13(3), Article 226. https://doi.org/10.3390/educsci13030226

Permana, P., Permatawati, I., & Khoerudin, E. (2023). Foreign language learning gamification using Quizizz: A systematic review based on students' perception. Eralingua: Jurnal Pendidikan Bahasa Asing dan Sastra, 7(2), 233–249. https://doi.org/10.26858/eralingua.v7i2.23969

Quyen, C. T., & Nguyen, H. C. (2025). Perceptions and dominant factors of Blooket as a game-based learning tool among Vietnamese non-English majored students. Dinamika Ilmu, 25(1), 123–144. https://doi.org/10.21093/di.v25i1.10268

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. Contemporary Educational Psychology, 61, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. Educational Psychology Review, 32(1), 77–112. https://doi.org/10.1007/s10648-019-09498-w

Sim, J. S. E., & Ismail, H. H. (2023). Using digital tools in teaching and learning English: Delving into English language teachers' perspectives. Creative Education, 14(10), 2021–2036. https://doi.org/10.4236/ce.2023.1410129

Torres Duque, J. A., & Argudo Garzón, A. L. (2024). Exploración del impacto transformador de la integración de las TIC en el aula de inglés como lengua extranjera. Religación. Revista de Ciencias Sociales y Humanidades, 9(40), Article e2401172. https://doi.org/10.46652/rgn.v9i40.1172

UNESCO. (2023). Global education monitoring report 2023: Technology in education. A tool on whose terms? https://www.unesco.org/gem-report/en/technology

Xiao, Q. (2025). Effects of gamification on EFL learning: A quasi-experimental study of reading proficiency and language enjoyment among Chinese undergraduates. Frontiers in Psychology, 16, Article 1448916. https://doi.org/10.3389/fpsyg.2025.1448916

Zhang, S., & Hasim, Z. (2022). Gamification in EFL/ESL instruction: A systematic review of empirical research. Frontiers in Psychology, 13, Article 1030790. https://doi.org/10.3389/fpsyg.2022.1030790

Zhou, S. (2024). Gamifying language education: The impact of digital game-based learning on Chinese EFL learners. Humanities and Social Sciences Communications, 11(1), Article 1518. https://doi.org/10.1057/s41599-024-04073-3

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest in this study and that the processes established by this journal have been ethically followed. They further certify that this work has not been published, in part or in full, in any other journal. It is declared that the principal investigator was the teacher of the course in which the experience took place; the measures adopted to mitigate this condition are described in the section on ethical considerations.

FINANCING

This research was carried out within the framework of the research project "Fortalecimiento del nivel didáctico en la formación docente en las instituciones públicas del cantón Jipijapa" and of the outreach project "Tareas dirigidas y apoyo psicopedagógico para fortalecer el aprendizaje de los estudiantes en la Educación Básica Pública de Jipijapa", both of the Universidad Estatal del Sur de Manabí. It received no external funding from public, commercial or not-for-profit agencies.

AUTHORSHIP CONTRIBUTION

Conceptualization: Rodrigo Alexander Rincón Zambrano, Jahiry del Rosario Molineros Ronquillo.
Methodology: Rodrigo Alexander Rincón Zambrano, Yerson José Muñoz-Chavarría.
Research: Rodrigo Alexander Rincón Zambrano, Yerson José Muñoz-Chavarría.
Writing – original draft: Jahiry del Rosario Molineros Ronquillo, Vanessa Valeria Hidalgo-Tumbaco.
Writing – revision and editing: Rodrigo Alexander Rincón Zambrano, Jahiry del Rosario Molineros Ronquillo, Yerson José Muñoz-Chavarría, Vanessa Valeria Hidalgo-Tumbaco.