

Kahoot, Learning Motivation, and Academic Achievement: A Systematic Review for Indonesian Madrasah

Muhammad Dede Alkautsar*, Yogi Irdes Putra, Fauziah

Universitas Muhammadiyah Muara Bungo, Indonesia

Corresponding Email: dedealkautsar663@gmail.com

Abstract

Game-based learning platforms have been adopted rapidly across Indonesian classrooms, yet evidence on their contribution to measured academic achievement remains scattered. This study synthesizes recent empirical evidence on the relationship between Kahoot-mediated instruction, learning motivation, and student learning outcomes. A systematic literature review was conducted following the PRISMA 2020 protocol, drawing on Scopus, ERIC, DOAJ, Garuda, and Google Scholar for studies published between 2020 and 2026. Of 439 records identified, 30 empirical studies satisfied the eligibility criteria and were appraised for methodological quality before thematic synthesis. Kahoot-mediated instruction was associated with positive achievement outcomes across primary, secondary, and tertiary settings, although reported effect magnitudes varied considerably between contexts. Learning motivation emerged as both an outcome of gamified instruction and a mechanism through which platform affordances translate into cognitive gains, yet formal mediation testing appeared in only three studies. Reported associations were moderated by instructional integration, subject domain, exposure duration, teacher readiness, and infrastructural access rather than by the technology itself. The synthesis yielded the Gamified Engagement–Motivation–Achievement framework, which positions motivation as the mediating pathway linking platform features to learning outcomes.

Article History

Received : 12 July 2025
Revised : 21 August 2025
Accepted : 12 Jan 2026

Keywords:

Kahoot, Game-Based Learning, Learning Motivation, Learning Outcomes, Systematic Literature Review

How to Cite:

Alkautsar, M. D., Putra, Y. I., & Fauziah. (2026). Kahoot, learning motivation, and academic achievement: A systematic review for Indonesian madrasah. *Educazione: Journal of Education and Learning*, 3(2), 230–244.

INTRODUCTION

The integration of digital technology into classroom instruction has become one of the defining features of contemporary educational practice worldwide. Within Indonesia, this movement has been reinforced by curriculum reforms that explicitly encourage differentiated, learner-centered, and technologically mediated instructional approaches. *Madrasah* institutions operating under the Ministry of Religious Affairs have participated in this shift while simultaneously maintaining their distinctive religious and moral curriculum commitments. The resulting instructional environment is therefore shaped by competing demands for technological modernization and for curricular continuity. Among the digital platforms adopted in these settings, Kahoot has achieved particularly wide diffusion because of its accessibility and its low technical demands on teachers. The platform presents multiple-choice items under time pressure, awards points according to speed and accuracy, and displays a public leaderboard after every question. International reviews have documented consistent reports of heightened engagement and enjoyment following its introduction into ordinary lessons (Wang & Tahir, 2020; Özdemir, 2025). Whether these affective responses translate reliably into measurable learning gains is a considerably more contested empirical question (Zh et al., 2024; Handayani et al., 2026).

A substantial body of quantitative evidence now reports positive effects of gamified quiz platforms on academic achievement across varied subject domains. Meta-analytic syntheses of controlled studies have converged on moderate pooled effect sizes for achievement outcomes,

together with somewhat smaller effects for affective and attitudinal variables (Bai et al., 2020; Sailer & Homner, 2020). More recent syntheses focused specifically on Kahoot have reported comparable patterns while emphasising the heterogeneity that persists across primary research (Özdemir, 2025; Diaz & Estoque-Loñez, 2024). Heterogeneity of this magnitude suggests that the platform itself is unlikely to be the principal source of observed variation. Attention has accordingly shifted toward the instructional and motivational conditions that surround platform use rather than towards the technology in isolation.

Learning motivation occupies a central position within these explanatory accounts, both as a valued outcome in its own right and as a proposed mediating mechanism. Self-determination theory distinguishes between controlled and autonomous forms of regulation, and it predicts that externally administered rewards may undermine intrinsic interest under certain conditions (Ryan & Deci, 2020). Gamified platforms rely heavily on points, rankings, and time pressure, which are recognizably controlling features rather than autonomy-supportive ones. Recent meta-analytic work in K-12 contexts has indeed found that extrinsic motivation responds more strongly to gamification than intrinsic motivation does (Kurnaz, 2025). This asymmetry raises important questions about the durability of motivational gains beyond the novelty period of initial implementation.

The Indonesian literature on this topic has expanded rapidly but exhibits several recurrent limitations that constrain cumulative knowledge building. Many published studies report significance values without accompanying effect sizes, confidence intervals, or measures of practical importance. Sample sizes are frequently small, allocation to conditions is rarely randomized, and intervention durations often span only a handful of instructional sessions. Instrument validation procedures are described briefly or omitted entirely in a considerable proportion of reports, which makes it difficult to distinguish genuine instructional effects from novelty and expectancy artifacts (Sukarismanti et al., 2024; Nurhayati & Fathurrohman, 2025). A related conceptual limitation concerns the treatment of motivation, which is frequently measured as a parallel outcome alongside achievement, after which a positive result on both variables is interpreted as evidence of a causal pathway between them. Such an inference does not follow from parallel measurement, because the two outcomes may respond independently to the same instructional stimulus. Formal mediation testing remains rare within the Indonesian corpus, and longitudinal designs capable of establishing temporal precedence are rarer still.

A third limitation concerns contextual specificity, particularly with respect to *madrasah* institutions at the junior secondary level. Existing reviews draw predominantly on general secondary schools, higher education, and English-language instructional settings in other national contexts. *Madrasah Tsanawiyah* classrooms differ in curricular composition, in the religious framing of assessment practices, and frequently in device availability and network reliability. Evidence generated in resource-rich settings cannot be transferred to these environments without careful examination of boundary conditions. Recent Indonesian work on curriculum readiness for digital learning media has emphasised precisely this concern (Handayani et al., 2026; Novita et al., 2026).

This review addresses these three gaps through a systematic synthesis conducted under the PRISMA 2020 protocol. It pursues three interrelated research questions in the sections that follow. First, what does the recent empirical evidence indicate about the association between Kahoot-mediated instruction and measured learning outcomes? Second, in what ways is learning motivation positioned within this evidence base, whether as an outcome, a mediator, or a moderating condition? Third, which instructional, contextual, and infrastructural conditions moderate the strength of the reported relationships? The article proceeds by describing the review protocol, presenting the descriptive and thematic results of the synthesis, discussing their theoretical and practical implications, and concluding with recommendations for *madrasah* practice and for future research.

RESEARCH METHOD

This study employed a systematic literature review design guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement, in its 2020 revision (Page et al., 2021). A systematic review was selected because the research questions concern the state, consistency, and boundary conditions of an existing evidence base rather than the effect of a single intervention within a single classroom. The design is appropriate when primary studies are numerous, methodologically heterogeneous, and dispersed across national and international outlets, which describes the Kahoot literature accurately. The review followed the sequential stages of protocol specification, systematic searching, screening against predetermined criteria, quality appraisal, data extraction, and thematic synthesis (Gough et al., 2017; Booth et al., 2016). Each stage was documented in an audit trail so that the procedure could be independently reproduced by other researchers.

The review protocol was specified before searching commenced, and it recorded the research questions, the eligibility criteria, the databases to be searched, and the planned approach to synthesis. Five sources were searched systematically: Scopus for internationally indexed literature, ERIC for education-specific coverage, the Directory of Open Access Journals for open-access outlets, Garuda for nationally indexed Indonesian journals, and Google Scholar for grey and recently released material. Searching was restricted to the period from January 2020 to March 2026, which corresponds to the five-year currency window applied to journal articles in this review. The search was executed in both English and Indonesian to ensure that nationally published evidence was not systematically excluded. Reference lists of all included studies were also hand-searched for relevant records that database queries had not retrieved. Table 1 records the exact strings applied in each source together with the yield obtained.

Table 1. Search Strings and Yield by Database

Database	Search string applied	Records
Scopus	TITLE-ABS-KEY(("Kahoot" OR "game-based learning" OR "gamified quiz") AND ("learning outcome" OR achievement) AND (motivation OR engagement))	118
ERIC	(Kahoot OR "game-based learning") AND ("academic achievement" OR "student motivation"), peer-reviewed only	74
DOAJ	"Kahoot" AND ("learning outcome" OR "motivation"), education subject filter applied	96
Garuda	"Kahoot" AND ("hasil belajar" OR "motivasi belajar" OR "media pembelajaran")	84
Google Scholar	Kahoot "hasil belajar" "motivasi belajar", first 10 result pages screened	40
Hand-searching	Reference lists of included studies and Sinta-indexed education journals	27

Note: Searches were executed between January and March 2026 and were restricted to the publication window from 2020 to 2026.

Table 2 sets out the eligibility criteria that governed screening decisions at every stage of the review process. Screening proceeded in two sequential stages that were conducted independently by two members of the research team. In the first stage, titles and abstracts were assessed against these criteria, and records that were clearly off-topic were removed. In the second stage, the full texts of all retained records were retrieved and examined in detail against the same criteria. Disagreements between the two reviewers were resolved through discussion, and the third author adjudicated the small number of cases where consensus could not be reached. Inter-rater agreement at the full-text stage was calculated using Cohen's kappa and reached a value of 0.84, which indicates substantial agreement between independent assessors.

Table 2. Eligibility Criteria Applied at Title, Abstract, and Full-Text Screening

Dimension	Inclusion criteria	Exclusion criteria
Publication window	Journal articles published between 2020 and 2026; books published within the preceding 10 years	Material published before 2020 unless it constituted a foundational theoretical source

Dimension	Inclusion criteria	Exclusion criteria
Study type	Empirical studies reporting primary data, together with systematic reviews and meta-analyses	Editorials, opinion pieces, conceptual essays, and unrefereed reports without data
Intervention focus	Kahoot or a directly comparable gamified student response platform used for instruction or assessment	Studies of general educational technology in which the platform was incidental
Outcome measurement	At least one measured learning outcome, motivational outcome, or engagement outcome	Studies reporting only descriptive usage statistics without an outcome measure
Educational setting	Formal primary, secondary, <i>madrasah</i> , or higher education institutions	Corporate training, informal learning, and non-institutional settings
Language and access	Published in English or Indonesian with retrievable full text	Full text unavailable after direct author contact was attempted

Source: Authors' protocol, adapted from Booth et al. (2016) and Page et al. (2021).

The review was conducted over a period of approximately six months, with protocol development, searching, screening, appraisal, and synthesis proceeding in sequence. All three authors are teacher educators working in Indonesian institutions, and all three have prior experience of using gamified assessment platforms within their own instructional practice. This professional familiarity constitutes both an analytical resource and a potential source of confirmation bias, since prior positive experience may predispose reviewers towards favorable interpretation of ambiguous findings. The team addressed this risk through several deliberate procedures. Screening and extraction decisions were duplicated independently, disagreements were logged rather than resolved informally, and a reflexive journal was maintained in which each author recorded interpretive assumptions as they arose. Coding of a randomly selected third of the corpus was performed independently by two authors and compared systematically before themes were finalized.

Methodological quality was appraised for all 30 retained studies using a structured instrument adapted for mixed-method review purposes. The instrument assessed six dimensions: clarity of research aims, appropriateness of design to the stated aims, adequacy of sampling and reporting of sample characteristics, evidence of instrument validity and reliability, transparency of analytical procedures, and the extent to which conclusions were warranted by the reported data. Each dimension was scored on a three-point scale, producing a maximum possible score of 18 points per study. Studies scoring 12 points or above were classified as higher quality, while those scoring between eight and 11 points were classified as moderate quality. No study scoring below eight points was retained in the final corpus, and four records were excluded on this basis during appraisal.

Data were extracted from each included study using a standardized extraction form developed during protocol specification. The form captured bibliographic details, national context, educational level, subject domain, research design, sample size, intervention duration, instruments employed, outcome constructs measured, reported statistical results, and stated limitations. Extraction was performed by the first author and verified independently by the second author against the original full texts. Reference management and citation formatting were handled throughout using Mendeley, in accordance with the 7th edition of the American Psychological Association style manual.

Synthesis proceeded through reflexive thematic analysis applied to the extracted data rather than through statistical pooling (Braun & Clarke, 2022). Meta-analysis was considered during protocol development but was judged inappropriate for this corpus, because the included studies employed incommensurable outcome measures and reported effect sizes too inconsistently for valid pooling. The analytical process involved repeated reading of the extraction records, generation of initial codes across the corpus, aggregation of codes into candidate themes, and iterative refinement of those themes against the primary data. Five analytical themes were

generated through this process and are reported in the following section. Trustworthiness was supported through independent double coding, through the maintenance of a reflexive audit trail, and through peer debriefing among the three authors at each analytical stage.

RESULT AND DISCUSSION

Results

The systematic search identified 439 records in total, comprising 412 records retrieved from database queries and 27 records located through hand-searching of reference lists. Removal of duplicates reduced the pool to 356 unique records, of which 241 were excluded at the title and abstract stage as clearly off-topic. Full texts were retrieved for the remaining 115 records and assessed in detail against the eligibility criteria. A further 85 records were excluded at this stage, most commonly because they reported no measured learning outcome or because the gamified platform was incidental rather than central to the study. Thirty studies satisfied all criteria and proceeded to quality appraisal and thematic synthesis, as Figure 1 sets out.

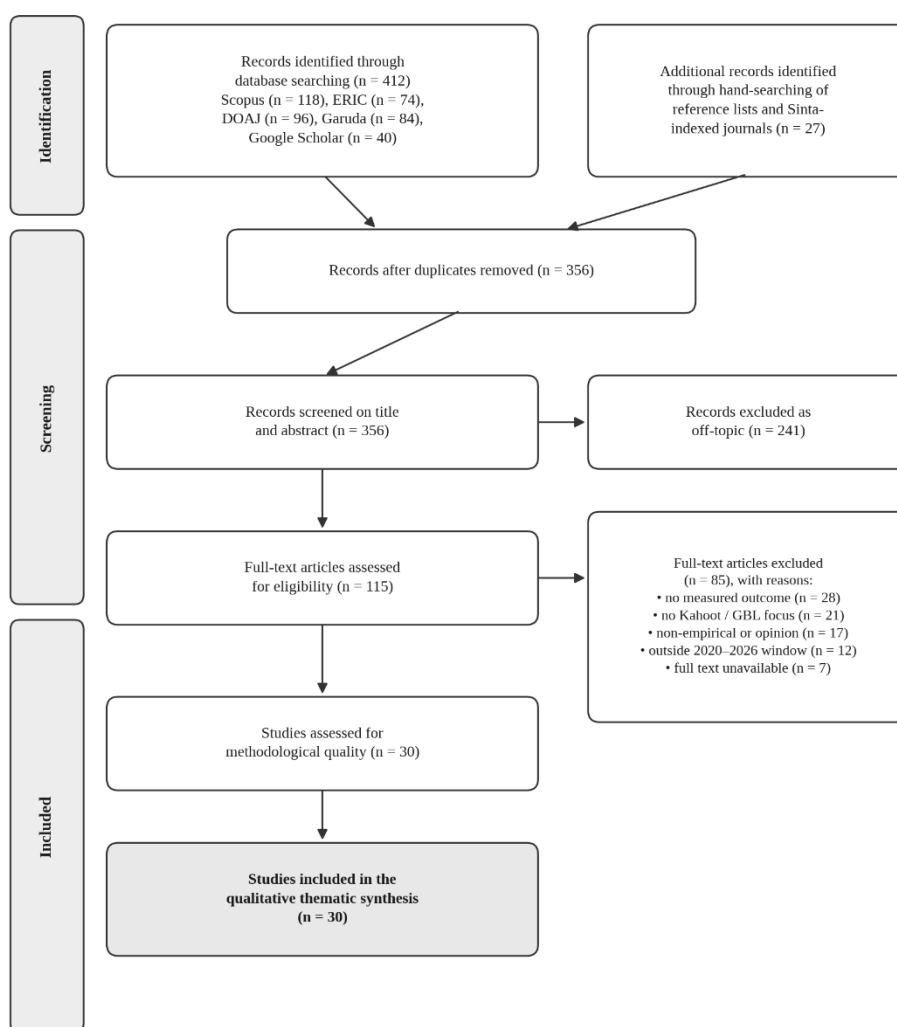


Figure 1. PRISMA 2020 Flow Diagram of Study Identification, Screening, Eligibility Assessment, and Inclusion

The descriptive profile of the final corpus is presented in Figure 2. Publication activity increased steadily from 2020 onwards and peaked during 2024, which reflects the

accelerated adoption of digital assessment tools following the return to face-to-face instruction. Junior secondary settings were the most frequently represented educational level, followed closely by senior secondary and primary contexts. Quasi-experimental designs predominated within the corpus, although a substantial minority of studies employed pre-experimental or correlational approaches without control conditions. Twelve studies measured achievement and motivation simultaneously, which made them the most informative subset for addressing the second research question.

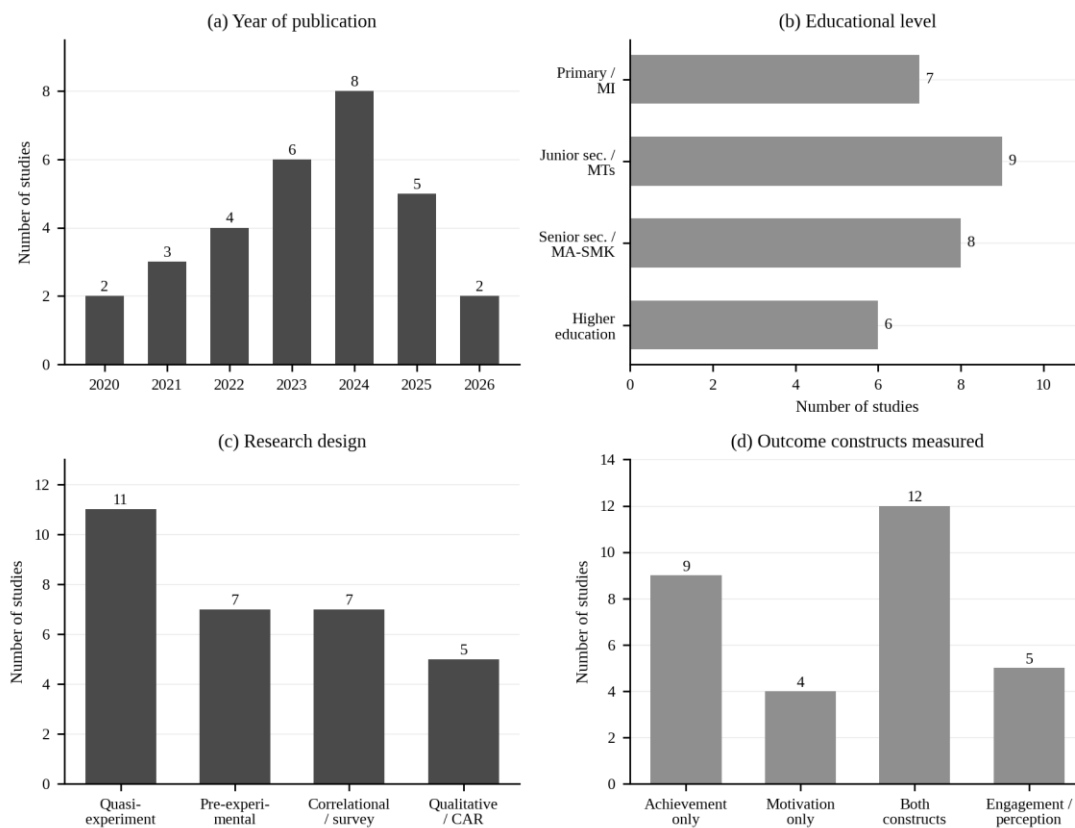


Figure 2. Descriptive Profile of the Thirty Included Studies by Year, Educational Level, Research Design, and Outcome Constructs

Quality appraisal produced a mean score of 12.4 points across the corpus, with individual scores ranging from 8 to 17 points. Eighteen studies were classified as higher quality and 12 as moderate quality under the criteria described above. The dimension attracting the lowest average score concerned the reporting of instrument validity and reliability, which was documented adequately in only 19 of the 30 studies. Reporting of effect sizes was similarly sparse, appearing in 11 studies overall and concentrated within the internationally indexed subset. Table 3 summarizes the characteristics of a representative selection of included studies spanning the range of contexts and designs.

Table 3. Characteristics of Representative Studies Within the Final Corpus

Study	Context and level	Design	N	Principal reported outcome
Özdemir (2025)	International; multiple levels	Meta-analysis	Multiple	Positive pooled effects on achievement, retention, and motivation
Rayan & Watted (2024)	Israel; primary science	Quasi-experiment	109	Significant gains in conceptual grasp, self-efficacy, and enjoyment

Kurnaz (2025)	International; K-12	Meta-analysis	31 studies	Moderate motivational effect, stronger for extrinsic regulation
Janković & Lambić (2022)	Serbia; primary grade three	Quasi-experiment	Not stated	Achievement advantage for the gamified condition
Sugiani (2023)	Indonesia; vocational secondary	Quasi-experiment	64	Significant association with interest and learning outcomes
Wulandari & Suparno (2023)	Indonesia; secondary	Pre-experimental	Not stated	Significant pre-post difference under problem-based integration
Ajang & Yasin (2024)	Malaysia; secondary physics	Quasi-experiment	Not stated	Post-test improvement alongside raised motivation
Anuraga et al. (2024)	Indonesia; primary civics	Qualitative	Not applicable	Favourable perceptions of critical reasoning gains
Rosfiani et al. (2023)	Indonesia; Arabic language	Action research	Not stated	Improved outcomes under gamified assessment cycles
Sukarismanti et al. (2024)	Indonesia; language instruction	Systematic review	Multiple	Positive trend alongside methodological inconsistency

Note: Sample sizes are reported as stated in the original publications; several Indonesian studies did not report this information in recoverable form.

Direction and Variability of Achievement Effects

The first and most consistent finding concerns the direction of reported achievement effects across the corpus. Twenty-six of the thirty included studies reported a positive association between Kahoot-mediated instruction and at least one measured learning outcome. Reported significance values fell below the conventional threshold in the substantial majority of hypothesis-testing studies, and no study within the corpus reported a statistically significant negative effect on achievement. The consistency of this directional pattern across varied national contexts, subject domains, and educational levels constitutes meaningful evidence that the association is robust rather than incidental (Özdemir, 2025; Bai et al., 2020).

The magnitude of these effects, however, varied to a degree that directional consistency conceals entirely. Where effect sizes were reported, they ranged from small to large across otherwise comparable implementations. Meta-analytic work in K-12 gamification has documented pooled effects in the moderate range while simultaneously reporting high heterogeneity between primary studies (Kurnaz, 2025). Such heterogeneity indicates that the conditions surrounding implementation account for more outcome variance than the presence of the platform does. Studies employing pre-experimental designs without control conditions also tended to report larger raw improvements than studies employing quasi-experimental designs with comparison groups, which is consistent with the operation of maturation, testing, and expectancy effects that uncontrolled designs cannot separate from the intervention itself (Sukarismanti et al., 2024; Nurhayati & Fathurrohman, 2025).

Subject domain accounted for a further portion of the observed variability across the corpus. Studies situated in language instruction, particularly vocabulary and terminology acquisition, reported the most consistently favorable outcomes within the reviewed evidence. Studies in mathematics and the natural sciences reported more variable results, with several reporting gains confined to procedural rather than conceptual measures. Qualitative meta-analytic work on language instruction has documented this domain advantage explicitly (Sağlamel & Yıldırım, 2024). The pattern is explicable in terms of the alignment between the recall-oriented item format of the platform and the cognitive demands of the target content.

Motivation as Outcome and as Mechanism

The second theme concerns the dual positioning of learning motivation within the included evidence. Sixteen studies measured motivation directly, and all sixteen reported improvement following the introduction of gamified quiz activities into ordinary instruction. Motivational gains were attributed most frequently to the immediacy of feedback, to the visibility of progress through scoring, and to the affective shift produced by reframing assessment as play. Studies in primary science contexts reported concurrent increases in self-efficacy, interest, and enjoyment alongside conceptual gains (Rayan & Watted, 2024). These affective outcomes were reported with greater consistency than cognitive outcomes across the corpus as a whole.

Twelve studies measured achievement and motivation within the same design, and this subset permits closer examination of the proposed causal pathway. In every case within this subset, the two outcomes moved in the same direction following the intervention. Only three studies, however, conducted any formal test of mediation to establish that motivational change accounted for the observed achievement change. The remaining nine studies inferred a mechanism from the concurrence of two parallel results, which does not warrant that inference on logical grounds. The mediating role of motivation is therefore plausible and widely assumed within this literature, but it remains substantially under-tested.

A minority of studies reported findings that complicate the uniformly positive motivational picture presented elsewhere in the corpus. Public leaderboards render relative standing visible to the entire class, which several studies identified as a source of discomfort among lower-achieving learners. Research examining stress and anxiety symptoms alongside achievement outcomes has reported that gamified formats can raise performance pressure even while raising engagement (Alsswey & Malak, 2024). Time constraints compound this effect, since learners who process information more deliberately are penalized by a scoring system that rewards speed. Only four studies within the corpus disaggregated motivational outcomes by prior achievement level, which leaves this question substantially unresolved.

Instructional Integration as a Moderating Condition

The third theme identifies instructional integration as the moderator most consistently associated with the magnitude of reported effects. Studies in which Kahoot was embedded within a structured pedagogical model reported stronger and more stable outcomes than studies in which it functioned as a detached closing activity. Integration with problem-based learning produced significant pre-post differences in secondary settings (Wulandari & Suparno, 2023), while integration with flipped and peer-instruction models yielded comparable improvements in participation (Sumaryo et al., 2024). Structured discussion cycles incorporating gamified assessment produced similar patterns in language instruction (Rosfiani et al., 2023). The consistency of this pattern across pedagogical models suggests that instructional coherence, rather than any particular model, is the operative factor.

Item design within the platform emerged as a second moderating condition of practical importance. Kahoot constrains authors to short multiple-choice items answered under time pressure, which suits factual recall and recognition far better than extended reasoning. Studies reporting the largest gains concentrated on knowledge domains where rapid retrieval is a legitimate instructional objective, including vocabulary, terminology, and procedural recall. Studies targeting conceptual understanding or higher-order reasoning reported more modest gains and frequently supplemented the platform with discussion-based follow-up activities (Anuraga et al., 2024).

The handling of feedback after each item emerged as a further design condition distinguishing stronger from weaker implementations. The platform displays correctness immediately, but it supplies no explanation of why a given response was incorrect. Studies reporting stronger conceptual outcomes described teachers who paused deliberately after difficult items to elicit reasoning from the class before proceeding. Studies reporting weaker outcomes described uninterrupted progression through the full item sequence with review deferred to the end. Evidence on gamified formative assessment systems indicates that their instructional value

depends heavily on how results are subsequently used (Zainuddin et al., 2020; Sanchez et al., 2020). Immediate correctness feedback appears necessary but insufficient for conceptual gain on this evidence.

Contextual and Infrastructural Constraints

The fourth theme concerns the contextual conditions that govern whether documented benefits are attainable in a given institution. Device availability, network reliability, and bandwidth capacity were identified as binding constraints in a majority of the Indonesian studies within the corpus. Kahoot requires simultaneous connection from every participating learner, which makes it considerably more demanding of infrastructure than asynchronous digital media. Studies conducted in schools with limited device provision described rotation arrangements, device sharing, and reduced participation as practical accommodations. Recent Indonesian work on curriculum readiness for digital learning media identifies infrastructural preparedness as a strategic precondition rather than an implementation detail (Handayani et al., 2026).

Teacher readiness constituted an equally significant contextual condition across the corpus. Effective implementation requires competence in item authoring, in calibrating time limits to item difficulty, in managing competitive dynamics constructively, and in converting quiz results into instructional decisions. Studies reporting weaker outcomes frequently described teachers who had adopted the platform without accompanying professional development. Research on technology adoption among Indonesian educators indicates that behavioral intention is shaped by institutional support and professional trust as much as by perceived usefulness (Novita et al., 2026; Isma et al., 2026). Comparable findings have been reported for digital competency development in personalized learning contexts (Kuswandi et al., 2025).

Madrasah settings present additional contextual features that the international literature does not address directly. Religious subject matter carries normative weight that may sit uneasily with competitive scoring formats, particularly where items concern doctrinal content. Indonesian work on gamified evaluation within Islamic education has approached this concern by confining game-based assessment to factual and terminological content while reserving interpretive material for discussion-based methods (Zh et al., 2024). This curricular boundary constitutes a design consideration specific to *madrasah* implementation rather than a general limitation of the platform. The institutions best positioned to realize the documented benefits are consequently those already possessing reliable connectivity, adequate device provision, and established professional development structures.

Methodological Quality and Evidentiary Gaps

The fifth theme concerns the state of the evidence base itself, which conditions the confidence attachable to every preceding finding. Effect sizes were reported in 11 of the 30 studies, and confidence intervals appeared in only seven. Randomized allocation was employed in three studies, while the remainder relied on intact class groups with varying degrees of baseline equivalence testing. Instrument validity and reliability were documented adequately in 19 studies, leaving 11 in which measurement quality cannot be verified by readers. Table 4 summarizes the distribution of the five themes and the associated quality indicators across the corpus.

Table 4. Thematic Distribution and Associated Quality Indicators Across the Corpus (n = 30)

Theme	Studies contributing	Higher-quality studies	Effect sizes reported
Direction and variability of achievement effects	26	16	11
Motivation as outcome and as mechanism	16	11	7
Instructional integration as moderator	14	10	5
Contextual and infrastructural constraints	18	9	3
Methodological quality and evidentiary gaps	30	18	11

Note: Studies contribute to more than one theme, so column totals exceed the corpus size of 30.

Three specific evidentiary gaps warrant explicit statement because they delimit the conclusions this review can support. First, long-term retention was assessed in only two studies, both of which employed follow-up intervals shorter than one month. Second, differential effects across learner subgroups, including low-achieving learners and those experiencing test-related anxiety, were examined in four studies with inconsistent results. Third, *madrasah* settings at the junior secondary level were represented by five studies, none of which employed a controlled design with validated instruments.

Discussion

The first research question asked what recent evidence indicates about the association between Kahoot-mediated instruction and measured learning outcomes. The synthesis supports an affirmative but carefully qualified answer to this question. Positive associations were reported with striking consistency across national contexts, educational levels, and subject domains, and no included study reported a significant negative achievement effect. The magnitude of these associations, however, varied so substantially that a single summary estimate would misrepresent the underlying evidence. This combination of directional consistency and heterogeneity in magnitude matches the pattern reported in international meta-analyses of the same platform (Özdemir, 2025) and of gamification more broadly (Bai et al., 2020; Sailer & Homner, 2020).

The practical implication of this pattern is that platform adoption alone provides an unreliable guide to instructional outcomes. If the technology were the primary causal agent, effect magnitudes would cluster far more tightly than the corpus shows them to do. The observed dispersion instead points towards implementation conditions as the dominant source of variance. This conclusion aligns with established findings from the wider educational technology literature, which has repeatedly demonstrated that instructional design mediates the influence of any delivery medium (Krath et al., 2021; Manzano-León et al., 2021). Kahoot is best understood, on this evidence, as an amplifier of instructional quality rather than as a substitute for it.

The second research question concerned the positioning of learning motivation within this evidence base. The synthesis found that motivation functions in the literature simultaneously as a valued outcome, as a presumed mediator, and occasionally as a moderating learner characteristic. Evidence supporting the first of these roles is strong, since every study measuring motivation reported improvement following gamified instruction. Evidence supporting the second role is considerably weaker, because formal mediation testing appeared in only three studies within the entire corpus. The prevailing inferential practice of treating parallel improvement in two outcomes as evidence of a causal pathway between them is not defensible on methodological grounds.

This distinction matters for both theory and practice rather than being merely technical. If motivation genuinely mediates the relationship, then instructional strategies that raise motivation through any means should produce comparable achievement gains, and the specific platform becomes incidental. If motivation and achievement instead respond independently to the same instructional stimulus, then targeting motivation alone will not reliably improve achievement outcomes. Self-determination theory offers the most productive lens for interpreting these findings, because the affordances that make Kahoot engaging are predominantly controlling in character (Ryan & Deci, 2020). Meta-analytic evidence showing stronger gamification effects on extrinsic than on intrinsic motivation corresponds closely to this theoretical prediction (Kurnaz, 2025). The motivational gains documented throughout this literature may therefore be more fragile than their consistency initially suggests.

The third research question concerned the conditions moderating the strength of reported relationships. Five such conditions emerged with sufficient consistency to warrant inclusion in a synthesized model: instructional integration, subject domain, exposure duration, teacher readiness, and infrastructural access. Figure 3 presents the Gamified Engagement–Motivation–Achievement framework developed from this synthesis. The framework specifies a pathway running from

platform affordances through proximal engagement mechanisms to motivational states and finally to learning outcomes, with the five conditions operating as moderators on every segment of that pathway.

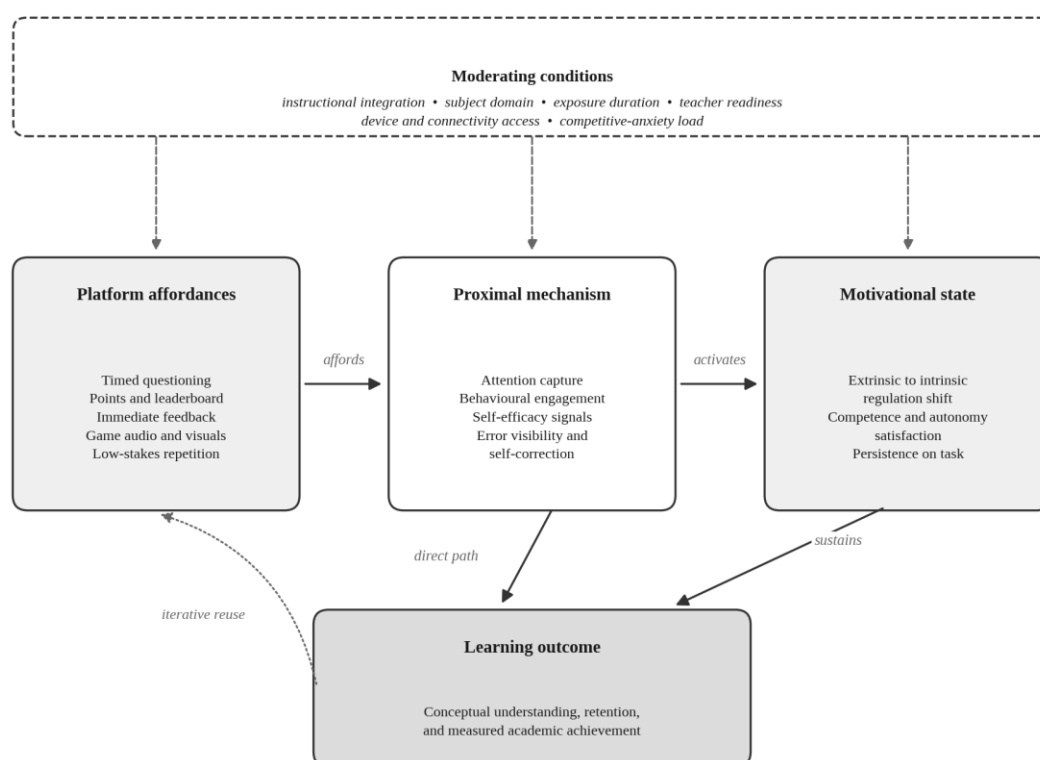


Figure 3. The Gamified Engagement–Motivation–Achievement (GEMA) Framework Synthesised From the Thirty Included Studies

The framework advances existing accounts in two specific respects that merit explicit statement. First, it separates the proximal mechanism of attentional and behavioral engagement from the motivational state that follows it, whereas much of the reviewed literature conflates these two constructs. This separation matters because engagement can be produced by novelty alone and may dissipate rapidly, while motivational change implies a more durable shift in regulation. Second, the framework treats the moderating conditions as acting on the entire pathway rather than solely on the final outcome. Infrastructural failure, for instance, disrupts the mechanism at its origin rather than merely attenuating its terminal effect.

The framework also incorporates a direct pathway from proximal mechanism to learning outcome that bypasses motivational mediation entirely. This pathway represents the retrieval-practice effect, whereby repeated low-stakes testing strengthens retention through cognitive rather than motivational means. Recognizing this pathway explains why achievement gains are sometimes reported in the absence of corresponding motivational change. It also suggests that part of the documented benefit of Kahoot may derive from its structure as distributed retrieval practice rather than from its gamified presentation. Testing the relative contribution of these two pathways constitutes a tractable and valuable research program.

For madrasah practice specifically, this synthesis yields several defensible recommendations. Gamified assessment should be integrated within an explicit instructional model rather than appended as a closing activity, since integration was the moderator most consistently associated with stronger outcomes. Item design should concentrate on factual and terminological content, with interpretive religious material reserved for discussion-based methods that permit reasoned elaboration. Competitive elements should be moderated deliberately, because

leaderboard visibility may induce anxiety among lower-achieving learners in ways that undermine the intended motivational benefit. Professional development should precede adoption, given the consistent association between teacher readiness and implementation quality documented throughout the corpus.

The synthesis also invites reconsideration of how the Indonesian evidence base positions itself in relation to international scholarship. National studies within the corpus cited international meta-analyses frequently but rarely adopted their reporting conventions, particularly with respect to effect sizes and confidence intervals. This asymmetry limits the contribution that Indonesian evidence can make to cumulative international synthesis, because studies reporting significance alone cannot be incorporated into pooled estimates. Raising reporting standards would therefore increase the visibility of national research without requiring additional data collection (Nurhayati & Fathurrohman, 2025; Sukarismanti et al., 2024).

Several limitations constrain the conclusions of this review and should be acknowledged directly. The exclusion of studies published before 2020 maintains currency but may omit relevant earlier evidence concerning gamified assessment. Restriction to English and Indonesian sources introduces a language bias that could affect the representation of evidence from other national contexts. The predominance of positive findings within the corpus raises the possibility of publication bias, which the absence of pooled effect sizes prevented this review from assessing formally through funnel plot analysis. Finally, thematic synthesis necessarily involves interpretive judgment, which independent double coding mitigates but cannot eliminate entirely.

CONCLUSION

This systematic review synthesized 30 empirical studies published between 2020 and 2026 concerning the relationship between Kahoot-mediated instruction, learning motivation, and academic achievement, and 3 principal conclusions follow directly from the evidence. First, the association between Kahoot-mediated instruction and measured learning outcomes is consistently positive in direction across the reviewed evidence, but it varies substantially in magnitude, which indicates that implementation conditions account for more outcome variance than the technology itself. Second, learning motivation is well established as an outcome of gamified instruction, yet its widely assumed mediating role between platform use and achievement remains substantially under-tested, since formal mediation analysis appeared in only 3 studies within the entire corpus. Third, 5 conditions moderate the strength of these relationships, namely instructional integration, subject domain, exposure duration, teacher readiness, and infrastructural access, and these were synthesized into the Gamified Engagement–Motivation–Achievement framework presented in this article. These conclusions imply that gamified assessment platforms merit adoption as components of a deliberate instructional design rather than as stand-alone evaluation activities, and that their implementation should be preceded by professional development and by realistic assessment of infrastructural capacity. Future studies should employ controlled designs with validated instruments, report effect sizes together with confidence intervals, test the mediating role of motivation formally rather than inferring it from parallel outcomes, and extend beyond a single instructional unit to determine whether motivational gains persist once novelty has faded. A well-designed controlled study within a Madrasah Tsanawiyah population, incorporating formal mediation testing and delayed post-testing, would address several of the gaps identified in this review simultaneously.

ACKNOWLEDGMENT

The authors thank the editorial team and the anonymous reviewers of *Educazione: Journal of Education and Learning* for their constructive engagement with this manuscript. Appreciation is also extended to colleagues who provided feedback during the development of the review protocol and the thematic synthesis.

REFERENCES

- Abdillah, R., Kuncoro, A., Erlangga, F., & Ramdhan, V. (2022). Pemanfaatan Aplikasi Kahoot! dan Quizizz sebagai Media Pembelajaran Interaktif Berbasis Gamifikasi. *Jurnal Pendidikan Sains dan Komputer*, 2(1), 92–102. <https://doi.org/10.47709/jpsk.v2i01.1363>
- Ajang, E. E., & Yasin, R. M. (2024). Keberkesanan Penggunaan Aplikasi Pembelajaran Interaktif dalam Meningkatkan Motivasi serta Pencapaian Pelajar dalam Mata Pelajaran Fizik. *Malaysian Journal of Social Sciences and Humanities*, 9(1), e002469. <https://doi.org/10.47405/mjssh.v9i1.2469>
- Alsswey, A., & Malak, M. Z. (2024). Effect of Using Gamification of “Kahoot!” as a Learning Method on Stress Symptoms, Anxiety Symptoms, Self-Efficacy, and Academic Achievement among University Students. *Learning and Motivation*, 87.
- Anuraga, B. S., Dwijayanti, I., & Sumarno, S. (2024). Efektifitas Media Kahoot dalam Pembelajaran PPKn untuk Meningkatkan Penalaran Kritis Peserta Didik Sekolah Dasar Kecamatan Tamansari. *As-Sabiqun*, 6(2), 237–247. <https://doi.org/10.36088/assabiqun.v6i2.4423>
- Arsyad, A. (2017). *Media Pembelajaran*. Rajawali Pers.
- Aunurrahman. (2022). *Belajar dan Pembelajaran*. Alfabeta.
- Azizah, F. W., & Widiyati, E. (2023). Analisis Kreativitas Guru dalam Memanfaatkan Teknologi Digital di MIN 3 Jombang. *El Bidayah: Journal of Islamic Elementary Education*, 5(2), 183–196. <https://doi.org/10.33367/jiee.v5i2.4141>
- Bai, S., Hew, K. F., & Huang, B. (2020). Does Gamification Improve Student Learning Outcome? Evidence from a Meta-Analysis and Synthesis of Qualitative Data in Educational Contexts. *Educational Research Review*, 30, 100322. <https://doi.org/10.1016/j.edurev.2020.100322>
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review* (2nd ed.). SAGE Publications.
- Braun, V., & Clarke, V. (2022). *Thematic Analysis: A Practical Guide*. SAGE Publications.
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). SAGE Publications.
- Diaz, A. F., & Estoque-Loñez, H. (2024). A Meta-Analysis on the Effectiveness of Gamification on Student Learning Achievement. *International Journal of Education in Mathematics, Science and Technology*, 12(5), 1236–1253.
- Djaali. (2020). *Psikologi Pendidikan*. Bumi Aksara.
- Gough, D., Oliver, S., & Thomas, J. (2017). *An Introduction to Systematic Reviews* (2nd ed.). SAGE Publications.
- Handayani, R., Rokimin, Nuriadin, I., & Zh, M. H. R. (2026). Curriculum Readiness for Digital Mathematics Learning Media: Examining Teachers’ Preparedness for Future-Oriented Mathematics Education. *Jurnal Ilmu Pendidikan dan Sosial*.
- Isma, A., Hasyim, S. H., Anugrani, N., Luthfi, A., & Khalil, I. A. (2026). Analisis Model UTAUT untuk Mengetahui Tingkat Penerimaan Teknologi Mahasiswa pada Aplikasi Kahoot. *Journal of Vocational Informatics and Computer Education*, 2(1), 58–71. <https://doi.org/10.61220/voice.v2i1.30>
- Janković, A., & Lambić, D. (2022). The Effect of Game-Based Learning via Kahoot and Quizizz on the Academic Achievement of Third Grade Primary School Students. *Journal of Baltic Science Education*, 21(2), 224–231.
- Kalogiannakis, M., Papadakis, S., & Zourmpakis, A.-I. (2021). Gamification in Science Education: A Systematic Review of the Literature. *Education Sciences*, 11(1), 22. <https://doi.org/10.3390/educsci11010022>
- Krath, J., Schürmann, L., & von Korfflesch, H. F. O. (2021). Revealing the Theoretical Basis of Gamification: A Systematic Review and Analysis of Theory in Research on Gamification, Serious Games and Game-Based Learning. *Computers in Human Behavior*, 125, 106963. <https://doi.org/10.1016/j.chb.2021.106963>

- Kurnaz, F. (2025). A Meta-Analysis of Gamification's Impact on Student Motivation in K-12 Education. *Psychology in the Schools*. <https://doi.org/10.1002/pits.70056>
- Kuswandi, D., Fadhl, M., Zh, M. H. R., & Ardiansyah, A. (2025). Implementation of Personalized Approach in Video Editing Learning to Improve Digital Competency of 21st Century Learners. *Cahaya Pendidikan*.
- Manzano-León, A., Camacho-Lazarraga, P., Guerrero, M. A., Guerrero-Puerta, L., Aguilar-Parra, J. M., Trigueros, R., & Alias, A. (2021). Between Level Up and Game Over: A Systematic Literature Review of Gamification in Education. *Sustainability*, 13(4), 2247. <https://doi.org/10.3390/su13042247>
- Ngilmi, N., Haliza, R. L., & Setiaji, B. (2023). Keefektifan Pembelajaran Berbasis Game untuk Meningkatkan Minat dan Motivasi Belajar Fisika: Metaanalisis. *Journal of Physics Education and Science*, 1(1), 15. <https://doi.org/10.47134/physics.v1i1.130>
- Novita, M., Zh, M. H. R., Abd Aziz, M. K. N., & Rokimin. (2026). Entrepreneurial Leadership and Technology Adoption: The Roles of Islamic Work Ethic and Theological Trust. *Journal of Educational Technology and Learning Creativity*.
- Nurhayati, & Fathurrohman. (2025). Gamification in School Education: A Systematic Review of Its Effectiveness in Improving Student Motivation and Academic Outcomes. *AL-ISHLAH: Jurnal Pendidikan*, 17(2). <https://doi.org/10.35445/alishlah.v17i2.6516>
- Nurtianingsih, T., Indratyaset, A. D. P., Putra, R. I., & Indah, T. (2022). Upaya Meningkatkan Literasi Siswa Melalui Program Kelompok Belajar pada Masa Pandemi Covid-19. *Jurnal Ilmiah Kampus Mengajar*, 2(1), 17–24. <https://doi.org/10.56972/jikm.v2i1.25>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Putra, I. G. N. D., & Redhana, I. W. (2025). Gamification Effects on Learning Outcomes: A Meta-Analysis. *AIP Conference Proceedings*, 3206(1), 080015.
- Rayan, B., & Watted, A. (2024). Enhancing Education in Elementary Schools through Gamified Learning: Exploring the Impact of Kahoot! on the Learning Process. *Education Sciences*, 14(3), 277. <https://doi.org/10.3390/educsci14030277>
- Rinjani, S. (2024). Implementasi Media Genially dalam Pembelajaran Bahasa Indonesia bagi Mahasiswa PBSI UIN Jakarta. *Diajar: Jurnal Pendidikan dan Pembelajaran*, 3(1), 57–64. <https://doi.org/10.54259/diajar.v3i1.2345>
- Rosfiani, O., Shaputra, R., Muchtar, A. D., & Razaq, F. (2023). Application of Kahoot Game Application Based Assessment in Class Discussion Model Action to Improve Students' Arabic Learning Outcomes. *Jurnal Riset Multidisiplin dan Inovasi Teknologi*, 2(1), 401–409. <https://doi.org/10.59653/jimat.v2i01.551>
- 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, 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>
- Sanchez, D. R., Langer, M., & Kaur, R. (2020). Gamification in the Classroom: Examining the Impact of Gamified Quizzes on Student Learning. *Computers & Education*, 144, 103666. <https://doi.org/10.1016/j.compedu.2019.103666>
- Sardiman, A. M. (2024). *Interaksi dan Motivasi Belajar Mengajar*. Rajawali Pers.
- Sağlamel, H., & Yıldırım, M. (2024). A Qualitative Meta-Analysis of Kahoot! in Language Instruction. *Journal of Qualitative Research in Education*, 39, 69–94. <https://doi.org/10.14689/enad.39.1903>

- Sugiani, K. A. (2023). Pengaruh Media Pembelajaran Kahoot Berbasis Game Based Learning terhadap Minat dan Hasil Belajar Siswa SMK di Buleleng. *Edusaintek: Jurnal Pendidikan Sains dan Teknologi*, 10(2), 457–474. <https://doi.org/10.47668/edusaintek.v10i2.770>
- Sugiyono. (2022). *Metodologi Statistika untuk Penelitian*. Alfabeta.
- Sukarismanti, S., Subyantoro, S., Pristiwati, R., & Samsudin, S. (2024). Tren dan Dampak Penggunaan Kahoot! dalam Pembelajaran Bahasa Indonesia: A Systematic Literature Review. *Jurnal Genre (Bahasa, Sastra, dan Pembelajarannya)*, 6(1), 208–220. <https://doi.org/10.26555/jg.v6i1.9281>
- Sumaryo, A., Dewi, R. S., & Legiani, W. H. (2024). Penerapan Model Pembelajaran Flipped Classroom Tipe Peer Instruction Flipped Berbasis Game Based Learning untuk Meningkatkan Partisipasi Belajar Peserta Didik di MTsN 4 Lebak. *Indonesian Journal of Education and Development Research*, 2(1), 70–77. <https://doi.org/10.57235/ijedr.v2i1.1545>
- Uno, H. B. (2023). *Pembelajaran Menciptakan Proses Belajar Mengajar yang Kreatif dan Efektif*. Bumi Aksara.
- Wang, A. I., & Tahir, R. (2020). The Effect of Using Kahoot! for Learning – A Literature Review. *Computers & Education*, 149, 103818. <https://doi.org/10.1016/j.compedu.2020.103818>
- Wulandari, D., & Suparno, S. (2023). Pengaruh Problem Based Learning Berbasis Game Kahoot terhadap Kemandirian Peserta Didik dalam Belajar. *Pijar: Jurnal Penelitian Bidang Pendidikan dan Pembelajaran*, 3(2), 45–51. <https://doi.org/10.56393/pijar.v3i2.148>
- Zainuddin, Z., Shujahat, M., Haruna, H., & Chu, S. K. W. (2020). The Role of Gamified E-Quizzes on Student Learning and Engagement: An Interactive Gamification Solution for a Formative Assessment System. *Computers & Education*, 145, 103729. <https://doi.org/10.1016/j.compedu.2019.103729>
- Zh, M. H. R., Putra, M. F. B., Kuswandi, D., & Ardiansyah, A. (2024). Developing Wordwall Evaluations in Blended Islamic Education Using the Smith and Ragan Model. *Al-Aulia: Jurnal Pendidikan dan Ilmu-Ilmu Keislaman*, 10(1).
- Özdemir, O. (2025). Kahoot! Game-Based Digital Learning Platform: A Comprehensive Meta-Analysis. *Journal of Computer Assisted Learning*, 41(1). <https://doi.org/10.1111/jcal.13084>