p-ISSN 2252-7427, e-ISSN 2541-3635
https://journal.uniku.ac.id/index.php/IEFLJ/index

p-ISSN 2252-7427 e-ISSN 2541-3635
Indonesian EFL Journal
AISEE
The Association of Indonesia Scholars of English Education

FOSTERING CRITICAL THINKING AND READING COMPREHENSION THROUGH KAHOOT-BASED COLLABORATIVE LEARNING IN TEAM MODE: A QUASI-EXPERIMENTAL STUDY IN INDONESIAN EFL JUNIOR HIGH SCHOOL

Florizqul Shodiq[1], Dwita Laksmita Rachmawati[2*]
[1,2]Master's Program in English Education, Universitas Muhammadiyah Gresik, Gresik, East Java, Indonesia.

Corresponding author's email: dwitalaksmita@umg.ac.id

APA Citation: Shodiq, F., & Rachmawati, D. L. (2026). Fostering critical thinking and reading comprehension through Kahoot-based collaborative learning in team mode: A quasi-experimental study in Indonesian EFL junior high school. *Indonesian EFL Journal, 12*(3), 1057-1070. https://doi.org/10.25134/ieflj.v12i3.209

Received: 23-03-2026 Accepted: 29-06-2026 Published: 30-09-2026

Abstract: Developing reading comprehension and critical thinking is essential for English as a Foreign Language (EFL) learners, yet conventional instruction may encourage passivity and reading anxiety. This study examined the effects of Kahoot!-based collaborative team learning on factual report text comprehension, critical thinking (analysis, evaluation, and inference), and perceived reading anxiety. Using a quasi-experimental non-equivalent control group design, the study involved 67 ninth-grade students at SMP Islamic Qon, Gresik, Indonesia, assigned to experimental (n = 34) and control (n = 33) groups. The experimental group participated in gamified collaborative instruction, while the control group received textbook-based instruction using the Presentation, Practice, Production model. Data from reading tests, observational rubrics, and a post-intervention survey were analyzed using ANCOVA, t-tests, and Wilcoxon signed-rank tests. The experimental group significantly outperformed the control group in reading comprehension (M = 75.41 vs. 65.27; $F(1, 64) = 113.398$, $p < .001$, partial $\eta^2 = .639$). Observations also showed significant improvements in analysis (+1.86), evaluation (+1.88), and inference (+1.91) ($p < .001$). Survey results indicated 91.8% overall student agreement regarding collaborative learning, motivation, and perceived anxiety reduction. These findings suggest that structured team-based gamification can support reading comprehension and critical thinking while providing a more engaging and less anxiety-provoking EFL learning experience.
Keywords: *Kahoot! team mode; collaborative learning; critical thinking; reading comprehension; EFL reading anxiety; gamification*

INTRODUCTION

Reading comprehension occupies a central position in secondary language education, functioning less as one skill among several than as the mechanism through which learners access disciplinary knowledge. Contemporary accounts describe it as a hierarchical construct in which literal recall, reorganization of textual information, inferential reasoning, and evaluative judgment interact, and proficiency depends on the coordinated deployment of vocabulary knowledge, background knowledge, and strategic processing (Amumpuni et al., 2025). For English as a foreign language (EFL) learners, each of these levels is complicated by limited lexical resources, unfamiliar syntactic patterning, and restricted exposure to extended authentic text, with difficulty concentrating disproportionately at the inferential and evaluative levels. Evidence from explicit strategy instruction with secondary EFL learners indicates that these higher-level processes are teachable rather than simply maturational, but that they require sustained and deliberate instructional scaffolding (Chinpakdee & Gu, 2024).

Reading serves as a fundamental academic survival skill in the globalized era, functioning as the primary gateway for students to acquire specialized information and achieve scholastic

success (Lin et al., 2020). In English as a Foreign Language (EFL) contexts, reading is a complex cognitive activity requiring strategic decoding, lexical inferencing, and textual evaluation. The urgency of elevating reading proficiency is underscored by recent international assessments; for instance, the 2022 Program for International Student Assessment (PISA) reported that Indonesian 15-year-olds scored substantially below the OECD average in reading literacy, with over 60% failing to attain baseline Level 2 proficiency (OECD, 2023). In response, Indonesia's national curriculum framework (*Kurikulum Merdeka*) explicitly mandates the integration of Higher-Order Thinking Skills (HOTS)—specifically analysis, evaluation, and creation—into secondary language education (Ministry of Education, Culture, Research, and Technology, 2022).

Despite these curricular mandates, ninth-grade students at SMP Islamic Qon, Gresik, face persistent hurdles in mastering EFL reading comprehension, particularly when processing technical factual report texts. Preliminary observations revealed that while learners can decode surface-level vocabulary, they struggle to execute critical thinking operations, such as identifying generic structures, distinguishing main ideas from supporting details, and drawing valid inferential conclusions. This cognitive bottleneck is compounded by high levels of foreign language reading anxiety and passive learning habits engendered by traditional teacher-centered instruction (Rachmawati & Hastari, 2022).

Grounding instructional intervention in theoretical frameworks is essential to address these multi-faceted challenges. Social constructivist theory posits that cognitive development is inherently socio-cultural and mediated through collaborative interaction within the zone of proximal development (ZPD) (Vygotsky, 1978; Shabani, 2023). Through structured peer negotiation, learners internalize complex reasoning strategies that exceed their individual capacity. Concurrently, Bandura's (2001) self-efficacy theory suggests that supportive peer environments foster vicarious mastery experiences, thereby mitigating affective barriers such as reading anxiety. While Digital Game-Based Learning (DGBL) platforms like Kahoot! have gained traction for enhancing engagement, their application in EFL contexts has predominantly focused on individual competition, speed, and lower-order vocabulary recall rather than collaborative HOTS development (Wang & Tahir, 2020; Zhang & Zou, 2022).

The potential of Digital Game-Based Learning (DGBL) to address these pedagogical challenges has been increasingly highlighted in recent research. Recent studies establish that integrating gamified platforms into EFL instruction breaks student passivity and enhances reading performance (Asari, 2023; Sabani et al., 2025). Gamification fosters active participation, enhances learning outcomes, and supports meaningful language acquisition (Laksanasut, 2025). However, conventional gamified reading tasks frequently prioritize quick response times, which can inadvertently encourage superficial guessing rather than deep textual analysis (Wang & Tahir, 2020). Shifting from traditional, teacher-centered lectures to game-based reading interventions can break the cycle of student passivity in EFL classrooms (Asari, 2023; Wijaya, 2024). Recent case studies in Indonesian junior high schools indicate that DGBL significantly increases student involvement and facilitates a deeper understanding of lesson materials (Chen, 2024) . The affective benefits of gamification are also well-documented, showing that gamifying reading motivation through information and communication technology (ICT) benefits EFL students by encouraging more participation and self-reliance (Sabani et al., 2025; Wardani et al., 2024). When students are engaged in a gamified environment, their anxiety levels decrease, allowing them to process language more naturally. Similarly, positive perceptions of DGBL suggest that modern learners are psychologically ready and eager to adopt digital game elements in their formal education (Rachmawati, 2025) .

Beyond motivation, the impact of strategic interventions on academic achievement remains a critical area of focus. Recent research on student achievement in reading comprehension through HOTS indicates that specific learning strategies are required to help students analyze and evaluate information effectively (Rachmawati & Purwati, 2021). Developing HOTS requires aligning instructional tasks with revised cognitive domain taxonomies, specifically focusing on analysis, evaluation, and synthesis (Mykola P., 2026). Implementing explicit reading strategies enables students to process complex texts at deeper cognitive levels (Shin & Chon, 2023). Fostering critical thinking and HOTS in reading requires systematic scaffolding and positive instructional beliefs (Rachmawati et al., 2021). Teachers must actively employ strategies that invite students to

review, refine, and create knowledge rather than just memorizing facts (Yang, 2025) . These studies suggest that a successful reading intervention must do more than just entertain; it must provide a cognitive framework for critical analysis.

Instructional response to this pattern has increasingly emphasized active learner involvement, on the premise that comprehension difficulty is sustained rather than resolved by transmissive, teacher-fronted routines in which learners read silently and complete worksheets individually. Gamification—the incorporation of game design elements into non-game instructional contexts—has been widely proposed as one such response. Reviews of gamified instruction in K–12 settings report positive associations with engagement, motivation, and achievement, though with considerable heterogeneity across implementations and outcome measures (Dehghanzadeh et al., 2024). The evidence specific to reading is thinner than the general gamification literature suggests. A systematic review applying PRISMA procedures to studies published between 2020 and 2024 identified only fifteen eligible empirical investigations of gamified reading instruction worldwide, of which a minority were conducted at the secondary level, and reported that a majority of this literature proceeds without an explicit theoretical framework (Wang et al., 2024). Where gamified reading interventions have been evaluated experimentally, reported effects are real but context-dependent, spanning comprehension performance, cognitive engagement, and reading attitude (Ndegwa et al., 2024; Prados Sánchez et al., 2023; Qiao et al., 2025).

Within this landscape, Kahoot! has become one of the most widely adopted gamified platforms, distinguished by quiz-based interaction, real-time scoring, and immediate whole-class feedback that lends itself to formative assessment (Özdemir, 2025). Synthesis evidence supports its general instructional value while resisting overstatement. A review of ninety-three studies concluded that Kahoot! can positively affect learning performance, classroom dynamics, learner attitudes, and anxiety, while explicitly identifying studies in which it produced little or no effect and documenting recurrent learner concerns about time pressure during answer submission (Wang & Tahir, 2020). A subsequent meta-analysis reported a significant moderate positive effect on academic achievement (Özdemir, 2025). These findings establish Kahoot! as a credible instructional technology, but they aggregate across subjects and outcome types and therefore do not license inference to reading comprehension specifically.

Evidence bearing directly on Kahoot! and reading is more equivocal. Mixed-methods research with Thai EFL undergraduates reported significantly higher comprehension scores and reading engagement for learners taught through quiz-based gamified activities, including Kahoot!, than for conventionally taught peers (Matyakhan et al., 2024). Conversely, a recent quasi-experimental comparison of a Kahoot! reading activity against a worksheet task with Saudi EFL learners found no statistically significant comprehension gain in either condition, notwithstanding strongly favorable learner attitudes toward the platform (Alluhaybi, 2026). Taken together, the evidence indicates that Kahoot! reliably supports participation and affective engagement, whereas its contribution to reading comprehension appears contingent on how the activity is pedagogically structured rather than on platform use as such.

That contingency points toward the social organization of the task. Social constructivist accounts locate cognitive development in mediated interaction within the zone of proximal development, where learners accomplish with peer support what they cannot yet accomplish alone (Vygotsky, 1978). Collaborative reading models operationalize this by requiring learners to articulate interpretations, surface misunderstandings, and negotiate meaning with peers. Collaborative strategic reading has accumulated repeated quasi-experimental support in EFL settings, including Indonesian classrooms, where it has been associated with gains in comprehension alongside reported reductions in reading anxiety and increased willingness to engage difficult text (Amumpuni et al., 2025; Herda et al., 2023; Yon et al., 2024). Such findings suggest that peer negotiation may be a productive complement to gamified reading tasks, since it introduces the verbalization and justification that rapid individual quiz responses tend to bypass.

Kahoot! Team Mode is of interest precisely because it alters this social organization. Rather than eliciting an individual response, it assigns a single answer to a group working from a shared device, with a brief discussion interval preceding submission. In principle this shifts the unit of performance from the individual to the team and creates an occasion for comparison of interpretations before commitment. It should not be assumed, however, that these processes occur automatically. Evidence directly comparing gamification mechanics is cautionary: in a quasi-

experiment with secondary EFL learners, students in a competitive gamification condition significantly outperformed those in a cooperative condition on reading comprehension and related measures (Qiao et al., 2024). Collaborative structuring, in other words, is not inherently superior to competitive structuring; whether team-based mechanics support comprehension appears to depend on whether the task compels evidence-based reasoning or merely permits brief consultation before a speed-scored response.

Despite extensive research on gamification and collaborative learning, three critical gaps remain. Most Kahoot! studies in Indonesia focus on individual competition or vocabulary recall, neglecting team-mode consensus-building for reading comprehension. The mechanism of structured cognitive pauses in gamified environments remains underexplored. There is limited empirical evidence on implementing DGBL in faith-based, resource-constrained Indonesian schools. This study addresses these gaps by examining how team-based consensus in Kahoot! fosters analysis, evaluation, and inference in EFL reading.

To address the identified gaps, this study formulated three core research questions: (1) To what extent does Kahoot-based collaborative learning in team mode significantly improve EFL students' reading comprehension of factual report texts compared to conventional instruction? (2) How does the implementation of Kahoot-based collaborative learning foster students' critical thinking skills in the dimensions of analysis, evaluation, and inference? (3) What are students' perceptions of gamified collaboration in reducing reading anxiety and enhancing engagement?

The goals of this investigation are in accordance with these inquiries into the topic of. (1) To examine the effectiveness of Kahoot-based collaborative learning in improving students' reading comprehension scores. (2) To analyze the development of students' critical thinking dimensions through gamified collaborative tasks. (3) To describe the students' engagement and affective responses toward the use of gamified collaboration.

This study offers three novel contributions to the field: (1) Pedagogical Novelty: Introducing a structured “cognitive pause" within Kahoot! team mode, which transforms rapid gamified mechanics into an evidence-based negotiation process requiring physical text verification. (2) Contextual Novelty: Providing empirical data from an Indonesian Islamic junior high school (SMP Islamic Qon), demonstrating how collaborative DGBL can be successfully adapted to under-resourced, single-device-per-team classroom settings.

This research holds theoretical, practical, and policy significance. *Theoretically*, it extends social constructivist principles to digital gamified environments, demonstrating that peer scaffolding within ZPD can be operationalized through structured game mechanics. *Practically*, it offers EFL educators a field-tested, low-cost pedagogical protocol for bridging technical text complexity and HOTS without requiring high-end technology infrastructure. *Policy-wise*, it provides evidence for curriculum developers and school leaders on operationalizing *Kurikulum Merdeka* literacy and critical thinking standards in resource-constrained environments.

METHOD

This study employed a quasi-experimental non-equivalent control group design involving pre-test, intervention, and post-test phases across two intact classes (Creswell & Creswell, 2018). Random assignment of individual students was administratively unfeasible and ethically disruptive to the established institutional cohort structure at SMP Islamic Qon. To account for potential baseline non-equivalence between intact classes, one-way analysis of covariance (ANCOVA) was specified as the primary analytical model, using pre-test reading scores as the baseline covariate.

p-ISSN 2252-7427, e-ISSN 2541-3635
https://journal.uniku.ac.id/index.php/IEFLJ/index

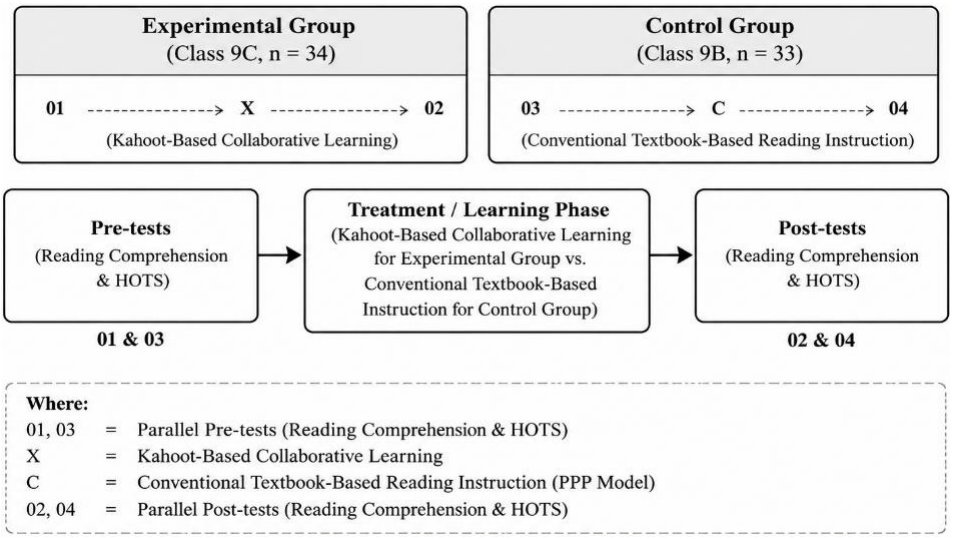


Figure 1. *Research design flow chart*

A purposive sampling technique was utilized to select two intact ninth-grade classes at SMP Islamic Qon, GKB Manyar, Gresik, East Java, Indonesia. The school operates under the Al-Qona'ah Islamic Boarding School Foundation. The sample comprised 67 students (*N = 67*): Class 9C (*n = 34*; 18 male, 16 female) served as the Experimental Group, and Class 9B (*n = 33*; 19 male, 14 female) served as the Control Group. Participants ranged in age from 14 to 15 years (*M = 14.3, SD = 0.48*). All participants had completed two years of national standardized secondary English instruction, establishing an intermediate-low baseline proficiency. Participants belonged to middle-to-lower socioeconomic backgrounds. The study protocol was reviewed and approved by the Institutional Review Board and School Administration of SMP Islamic Qon. Formal written informed consent was obtained from all participating students and their legal parents/guardians prior to data collection. Participants were informed that participation was entirely voluntary, data would be kept confidential and anonymized, and withdrawal from the study at any point would carry no academic or personal penalty.

Table 1. *Participant characteristics*

Variable	Control Group	Experimental Group
N	33	34
Grade	IX	IX
Age	14 years old	14 years old
Gender	19 male and 14 female	18 male and 16 female
School	SMP Islamic Qon	SMP Islamic Qon
Pre-test Mean	58.82	59.97
Instruction	Conventional Instruction	Kahoot!-based gamified collaborative instruction

The data collection utilized quantitative instruments to ensure robust triangulation. Data were collected using three validated quantitative instruments to ensure thorough methodological triangulation: parallel reading comprehension tests, a structured critical thinking observation rubric, and a post-intervention student perception survey. Factual report text comprehension was evaluated using parallel 25-item multiple-choice pre- and post-tests targeting literal understanding and higher-order thinking skills (HOTS). While content validity was established by two senior EFL university lecturers (*CVI = 0.92*), and pilot-testing on a non-sample cohort (*n = 30*) demonstrated internal consistency (pre-test *KR-20 = .84*; post-test *KR-20 = .86*).

Formative critical thinking performance during experimental treatment sessions was observed using a structured rubric adapted from Facione's (1990) Delphi consensus framework. The rubric operationalized three core cognitive dimensions into observable classroom behaviors scored on a 4-point scale (1.0 = Beginning, 2.0 = Developing, 3.0 = Proficient, 4.0 = Advanced): (a) Analysis (identifying generic text structures, classifying main ideas, and locating factual details); (b)

Evaluation (assessing source credibility, text logic, and statement consistency); and (c) Inference (synthesizing textual evidence to draw valid conclusions). Content and face validity were established through expert review by senior ELT lecturers. To ensure scoring consistency, two non-participant observers completed a calibration session prior to data collection, benchmarking ratings against standardized exemplars. During live classroom observations in Treatment Meetings 2 and 3, both observers independently rated student teams. Inter-rater reliability evaluated on a dual-observed sample yielded high agreement (92% absolute agreement; Cohen's κ = 0.84), with rating discrepancies resolved through immediate post-session consensus discussion.

Finally, student affective and cognitive responses were measured in the experimental group following the four-meeting intervention via a 12-item questionnaire rated on a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree") across four dimensions—materials understanding, intrinsic motivation, perceived reduction in reading anxiety, and collaborative value—yielding internal consistency with an overall Cronbach's α = .89.

The study was conducted across a two-meeting study schedule involving an experimental group (n = 34) and a control group (n = 33). In Meeting 1, both groups completed a 60-minute standardized, closed-book parallel pre-test to establish baseline reading performance. Across Meetings 2 and 3 (two 80-minute instructional sessions), the experimental group received gamified collaborative instruction focusing on factual report text structure, technical vocabulary, evaluation, and implicit inference. Students worked in seven heterogeneous, mixed-ability teams of four to five members (comprising one high-, two medium-, and one to two low-achieving students based on pre-test quartiles) using Kahoot! team mode.

In contrast, the control group received conventional textbook-based reading instruction following the Presentation, Practice, Production (PPP) framework across an identical timeframe (two 80-minute instructional meetings). To control for instructor variability, both experimental and control classes were taught by the same certified EFL instructor using identical factual report reading passages (e.g., natural phenomena and technological devices). In the control condition, instruction followed standard transmission routines: the teacher presented target vocabulary and text structures, students completed silent individual reading, and written comprehension exercises were completed independently on paper worksheets. Whole-class teacher-led answer verification followed each task. Meeting 4 involved a 60-minute post-test for both groups, followed by the perception questionnaire administered to the experimental group. Instructional fidelity across both conditions was monitored using a 10-item observation checklist evaluated by an independent observer, confirming 95% protocol adherence.

Quantitative data were analyzed using IBM SPSS Statistics, beginning with univariate normality checks via the Shapiro-Wilk test ($p > .05$) and baseline equivalence assessment using an independent-samples t-test on pre-test scores. To evaluate the primary treatment effect, a one-way ANCOVA was conducted with post-test scores as the dependent variable, instructional group as the fixed factor, and pre-test scores as the baseline covariate, following homogeneity of regression slopes verification through covariate-by-treatment interaction analysis. Additionally, observational rubric scores were quantitatively examined through descriptive statistics, while perception survey items were summarized using item means, standard deviations, and percentage agreement.

RESULTS AND DISCUSSION

To determine whether baseline reading comprehension differed significantly prior to instruction, an independent-samples *t*-test was performed on pre-test scores.

Table 2*: Group statistic*

Group	N	Pre-test M	SD	Std. Error M
Control	33	58.82	4.23	.737
Experimental	34	59.97	3.65	.627

Table 3*: Baseline equivalence based on pre-test scores*

Pre-Test	Levene's Test (F)	Levene's Test (sig)	t	df	p	Cohen's d
Equal variances assumed	1.26	0.266	-1.194	65	.237	-.292

Equal variances not assumed	-1.191	63.06	.238	

The negative sign reported for the *t*-statistic (*t = -1.194*) reflects the directional computational formula in SPSS when entering group codes. As the pre-test mean difference was only 1.15 points and $p = .237 > .05$, baseline reading comprehension capacity was statistically equivalent across both classes prior to treatment. The effect size (*d = -.292*) represents a small, non-significant baseline variance, while the negative sign reflects the direction in SPSS when entering group codes. Homogeneity of variance was examined using Levene's test prior to the independent-samples t-test analysis. The results showed no statistically significant variance differences between the control and experiment groups, $F = 1.26$, $p = .266$. Therefore, the assumption of equal variances was considered satisfied.

Table 4*: Test of normality (Shapiro-Wilk)*

Group	Test	N	Statistic	*p*	Skewness	Kurtosis
Control	Pre test	33	.966	.380	.356	-.148
Experimental	Pre test	34	.956	.190	-.517	1.084

Because all *p*-values exceeded the $\alpha = .05$ threshold, skewness/kurtosis values remained within the standard $\pm$ 1.0 range, and visual inspection of the diagnostic plots confirmed the absence of substantial deviations from normality, linear relationships, and influential outliers, the assumption of normality was satisfied for parametric testing.

Table 5*: Descriptive statistics of report text reading comprehension*

Group	N	Pre-test M (SD)	Post Test M (SD)	Gain M	Post Min-Max
Control	33	58.82 (4.23)	65.27 (4.86)	+ 6.45	60-77
Experimental	34	59.97 (3.65)	75.41 (5.24)	+ 15.44	67-88

At baseline, both groups exhibited comparable reading comprehension levels, with a pre-test mean difference of only 1.15 points (experimental *M = 59.97, SD = 3.65*; control *M = 58.82, SD = 4.23*). Following the intervention, the experimental group descriptively outperformed the control group on the post-test (*M = 75.41, SD = 5.24* vs. *M = 65.27, SD = 4.86*). Specifically, the experimental group achieved a mean score gain of +15.44 points from pre-test to post-test, whereas the control group achieved a mean score gain of +6.45 points that represents the pre-to-post mean score differences on the test scale rather than percentage gains. This resulted in an 8.99-point net gain difference in favor of the experimental condition. Furthermore, individual post-test scores in the experimental group ranged from 67 to 88, outperforming the control group's range of 60 to 77 as well as both groups' baseline averages.

The homogeneity of regression slopes assumption was assessed by examining the interaction between treatment group and pre-test scores. The interaction was not statistically significant, $F(1, 63) = 0.214$, $p = .645$, indicating that the relationship between pre-test and post-test scores did not significantly differ between the control and experimental groups. Thus, the assumption of homogeneity of regression slopes required for standard one-way ANCOVA was fully satisfied.

To determine whether the post-test mean difference was statistically significant when adjusting for pre-test baseline variations, a one-way ANCOVA was executed.

Table 6*: ANCOVA results with post-test as dependent variable*

Source	Sum Squares	df	M Square	F	p	η^2_p
Pre-test (covariate)	905.291	1	905.291	76.488	< .001	.544
Instructional	1342.148	1	1342.148	113.398	< .001	.639
Error	757.490	64	11.836			

Controlling for baseline pre-test performance, the ANCOVA confirmed a statistically significant main effect for instructional condition, $F(1, 64) = 113.398$, $p < .001$, $\eta^2_p = .639$. The instructional condition was associated with a large effect on adjusted post-test reading comprehension scores ($d \approx 2.66$).

Table 7*: Paired -samples t-test on both group between pre-test and post-test scores*

Group	Mean	Std. Deviation	t	df	Sig. (2-tailed)	Cohen's *d*
Control	+6.45	2.48	−14.975	32	< .001	-2.61
Experimental	+15.44	4.14	−21.772	33	< .001	-3.73

To further evaluate the robustness of the treatment advantage across analytical specifications, an independent-samples t-test was conducted on unadjusted pre-to-post mean gain scores (ΔM = M_post − M_pre). Negative *t*-statistics occur because baseline pre-test scores were subtracted from post-test scores. The results confirmed that the experimental group achieved significantly greater mean score improvements than the control group (+15.44, SD = 4.14 vs. +6.45, SD = 2.48). This sensitivity analysis establishes that the superior reading comprehension performance of the experimental group remains highly robust independent of baseline covariate adjustment. Both groups demonstrated significant within-group gains ($p < .001$); however, the experimental group exhibited a substantially higher effect size (d = -3.73, extremely large) than the control group (d = -2.61, extremely large).

To evaluate unadjusted post-test reading comprehension performance, an independent-samples *t*-test was conducted on post-test scores across both groups.

Table 8*: Independent-samples t-test on post-test scores*

Variance	Levene's F	Sig.	t	df	Sig. (2-tailed)	Std. Error Difference	cohen's *d*
Equal Variance Assumed	0.159	0.691	−8.203	65	< .001	1.23595	-2.01
Equal Variance not Assumed	—	—	−8.213	64.872	< .001	1.23455	

Table 8 The analysis revealed a statistically significant post-test performance difference favoring the experimental group, $t(65) = -8.203$, $p < .001$. The magnitude of the post-test mean difference was extremely large, with a standardized effect size of Cohen's $d = -2.01$. This result demonstrates a substantial unadjusted post-test performance gap between students instructed through Kahoot!-based collaborative learning in team mode and those in the conventional instruction condition.

Table 9. *Development of critical thinking dimensions during treatment meetings*

Critical Thinking Dimension	Session 1 Mean (SD)	Session 2 Mean (SD)	Gain	*z*	*p*
Analysis (Identifying generic structure & classifications)	1.76 (0.43)	3.62 (0.49)	1.86	-5.01	< .001
Evaluation (Assessing factual accuracy & text logic)	1.59 (0.50)	3.47 (0.51)	1.88	-5.04	< .001
Inference (Drawing valid, evidence-based deductions)	1.44 (0.50)	3.35 (0.48)	1.91	-5.08	< .001

Data in Table 9 demonstrate progressive growth across all three dimensions. Inference showed the largest absolute rating increase (+1.91 points, moving from 1.44 in session 1 to 3.35 in session 2), followed by evaluation (+1.88 points) and analysis (+1.86 points). By session 2, mean performance across all three critical thinking dimensions approached the upper bounds of the rating rubric (3.0 = Proficient, 4.0 = Advanced). To assess students' perceptions, a post-intervention questionnaire was administered to the experimental group, measuring four distinct dimensions on a 5-point Likert scale.

Table 10. *Student perceptions of Kahoot-based collaborative learning*

Questionnaire Item / Indicator	Mean (*SD*)	% Agreement
Materials Understanding: "Using the collaborative Kahoot quizzes helped me understand the specialized vocabulary and structure of report texts."	4.51 (0.51)	90.2%
Intrinsic Motivation: "Working with my teammates to solve the Kahoot challenges made me excited and motivated to read the texts."	4.68 (0.47)	93.6%
Perceived reduction in reading anxiety: "I felt less anxious reading difficult report texts because I could discuss them with my team before answering."	4.44 (0.56)	88.8%
Collaborative Value: "Negotiating and reaching a consensus with my classmates helped me learn how to analyze text evidence better."	4.71 (0.46)	94.2%
Overall Average Perceptions (α = *.89*)	4.59 (0.41)	91.8%

As displayed in Table 10, overall student agreement reached 91.8% (*M = 4.59*). The highest ratings were observed for collaborative value (*M = 4.71, 94.2%*) and intrinsic motivation (*M = 4.68, 93.6%*), with perceived reduction in reading anxiety scoring *M = 4.44* (*88.8%*) and materials understanding scoring *M = 4.51* (90.2%), confirming that gamified team-based learning fosters a highly supportive, engaging, and low-anxiety reading environment.

The quantitative results indicate that Kahoot!-based collaborative learning in team mode was associated with higher factual report text reading comprehension scores among EFL learners. After controlling for baseline pre-test performance, the ANCOVA revealed a statistically significant difference between the instructional groups, $F(1, 64) = 113.398$, $p < .001$, partial $\eta^2 = .639$. supported by a substantial adjusted post-test effect size ($d \approx 2.01$). Testing of ANCOVA assumptions confirmed that the homogeneity of regression slopes was fully satisfied, *F(1, 63) = 0.214, p = .645*, verifying that covariate adjustment was statistically sound across both groups.

The observed post-test mean difference was also substantial, with the experimental group (*M = 75.41, SD = 5.24*) outperforming the control group (*M = 65.27, SD = 4.86*). The corresponding independent post-test Cohen's *d* was approximately *2.01*, further indicating a large standardized difference between the groups. The results also confirmed that the experimental group achieved significantly greater mean score improvements than the control group (*+15.44 vs. +6.45; t(65) = -8.203, p < .001*; *d=2.01*). Together, the ANCOVA, independent post-test *t*-test, and gain-score analysis establish the robust nature of the intervention effect across multiple statistical approaches.

In terms of higher-order cognitive processing, observational rubric data collected across the treatment meetings demonstrated progressive within-group score increases in the experimental group across analysis (*1.76 ⇒ 3.62*), evaluation (*1.59 ⇒ 3.47*), and inference (*1.44 ⇒ 3.35*). Because these observational metrics were gathered exclusively within the experimental condition during instruction, they provide descriptive evidence of developmental progression during gamified team negotiation rather than a controlled, between-group estimate of treatment efficacy.

These empirical findings align directly with social constructivism and Vygotsky's (1978) zone of proximal development. Throughout the intervention, heterogeneously grouped students actively negotiated text meaning, defended their interpretations, and resolved conceptual misunderstandings prior to submitting final answers. Within these collaborative peer structures, higher-ability students provided dynamic linguistic and cognitive scaffolding to lower-ability team members (Shabani, 2023).

This structural requirement directly supported the observed growth across critical thinking dimensions (Table 9), with inferential reasoning exhibiting the largest absolute increase (*+1.91*). In factual report texts, drawing valid inferences demands synthesizing scattered details and identifying implicit relationships. The team consensus mechanism forced students to externalize their thought processes, effectively transforming implicit cognitive operations into explicit, evidence-based peer discussions (Lo & Hew, 2023; Sun & Zhang, 2023).

From an affective perspective, the significant reduction in self-reported reading anxiety (*88.8%* agreement, *M = 4.44*) strongly supports Bandura's (2001) self-efficacy theory. Because EFL reading anxiety frequently stems from fear of negative evaluation when encountering complex texts (Tsai & Lee, 2023), shifting the evaluative unit from the individual to the team established a supportive, low-stakes learning environment. Shared academic risk lowered learners' affective filters, encouraging active engagement with challenging report texts (Rachmawati & Fadhilawati, 2024).

Regarding alternative explanations and potential confounding variables, the progressive session-to-session gains observed in critical thinking rubrics (Table 9) suggest that cognitive development was driven by structured team negotiation rather than temporary novelty effects associated with introducing digital tools. Additionally, standardized lesson plans and adherence checklists were implemented to promote instructional consistency. Potential teacher enthusiasm and novelty effects cannot be entirely ruled out in a short-term intervention. These measures should be viewed as context-specific procedural safeguards whose sustained effectiveness requires longitudinal verification.

This study evaluated the impact of Kahoot-based collaborative learning in team mode on ninth-grade EFL students' report text reading comprehension, critical thinking skills, and reading anxiety. Controlling for baseline pre-test scores, the experimental group achieved significantly higher post-test reading comprehension scores compared to the control group, demonstrating a substantial treatment effect. Furthermore, observational data confirmed significant progressive growth across all key critical thinking dimensions, specifically analysis, evaluation, and inferential reasoning. Student perception surveys further indicated strong overall agreement regarding the collaborative value of the framework, its ability to enhance intrinsic motivation, and its role in mitigating reading anxiety.

Second, the active intervention was brief (spanning two 80-minute instructional sessions across four total study meetings) and lacked a delayed post-test, precluding assessment of long-term skill retention. While standardized lesson plans, fidelity checklists, and structured classroom observations were used as procedural safeguards, this short duration means that participant reactivity and initial novelty effects cannot be entirely ruled out in this specific context. Consequently, these findings represent a short-term, context-specific application whose sustained impact requires longitudinal verification.

Third, critical thinking was observed only in the experimental group, with no control-group baseline, which limits comparative and causal interpretation.

Fourth, reading anxiety was measured only after the intervention through a self-report survey rather than a pre/post scale, so no change in anxiety can be inferred. Self-report data are also susceptible to social desirability bias.

Fifth, the sample consisted of 67 students from a single faith-based junior high school in regional East Java, which limits generalizability. Finally, the multilingual code-switching observations were informal and not systematically coded.

CONCLUSION

This study evaluated the impact of Kahoot-based collaborative learning in team mode on ninth-grade EFL students' report text reading comprehension, critical thinking skills, and reading anxiety. Controlling for baseline pre-test scores, the experimental group achieved significantly higher post-test reading comprehension scores compared to the control group, demonstrating a substantial treatment effect. Furthermore, observational data confirmed significant progressive growth across all key critical thinking dimensions, specifically analysis, evaluation, and inferential reasoning. Student perception surveys further indicated strong overall agreement regarding the collaborative value of the framework, its ability to enhance intrinsic motivation, and its role in mitigating reading anxiety.

Beyond empirical gains, this research offers meaningful theoretical contributions by extending social constructivist theory to digital game-based learning environments. It demonstrates that digital team consensus, when mediated by structured cognitive pauses, effectively operationalizes peer scaffolding within Vygotsky's zone of proximal development. By reorienting standard gamification mechanics away from individual speed and rapid guessing toward collaborative, evidence-based inquiry, the study establishes a conceptual framework for transforming digital formative quizzes into platforms for deep critical thinking development.

Practically, these findings provide actionable insights for EFL educators seeking to optimize digital game-based learning in reading instruction. Teachers are encouraged to move beyond utilizing platforms like Kahoot! for individual competition or vocabulary recall. Transitioning to team mode and enforcing a structured cognitive pause provides a field-tested pedagogical protocol that compels learners to consult physical text evidence, negotiate interpretations, and co-construct meaning before submitting final answers, thereby strengthening both reading comprehension and collaborative skills.

From an institutional and contextual perspective, the study demonstrates a scalable, value-aligned model for under-resourced and faith-based educational settings, such as Indonesian

Islamic junior high schools. It establishes that digital formative assessment does not require costly individual student devices; rather, meaningful peer engagement and textual analysis can be achieved using a single projected teacher screen, shared physical worksheets, and one digital device per group. Curriculum designers and school administrators in similar educational contexts can leverage this low-cost framework to fulfill national literacy and higher-order thinking skills (HOTS) mandates without exceeding existing infrastructure limitations

Finally, several future research directions emerge to build upon these findings. Future studies should adopt longitudinal experimental designs featuring delayed post-tests to determine the long-term retention of comprehension and critical thinking gains. Replicating the intervention across diverse secondary school settings, including public and vocational institutions, will further establish generalizability. Additionally, prospective research should consider mediation analysis to examine the complex interactions between perceived reduction in reading anxiety, peer negotiation language, and critical thinking development during gamified collaboration.

REFERENCES

Alluhaybi, M. (2026). When fun meets focus: Comparing Kahoot and worksheets in Saudi EFL reading instruction. *Frontiers in Education*, *11*, Article 1795606. https://doi.org/10.3389/feduc.2026.1795606

Amumpuni, R. S., Hartono, R., Rukmini, D., & Haryanti, R. P. (2025). The impact of collaborative strategic reading on EFL students with different critical thinking levels. *Social Sciences & Humanities Open*, *12*, Article 101904. https://doi.org/10.1016/j.ssaho.2025.101904

Asari, S. (2023). Vocabulary development of EFL young learner after long peer interaction with video games. *JEES (Journal of English Educators Society)*, *8*(1), 45–56. https://doi.org/10.21070/jees.v8i1.1708

Asari, S., & Ma'ruf, N. (2023). Digital authentic resource and activity needs of English major students for autonomous learning over Self-Access Center (SAC). *Edulitics Journal*, *8*(2), 112–125.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bin-Hady, W. R. A., Ali, J. K. M., & Al-humari, M. A. (2024). The effect of digital interactive tools on EFL students' social and emotional learning. *Journal of Research in Innovative Teaching & Learning*, *17*(2), 243–255. https://doi.org/10.1108/JRIT-02-2024-0036

Chen, Y. C. (2024). Perceptions of digital game-facilitated language learning: A study on interactive storytelling in EFL contexts. *JALT CALL Journal*, *20*(3), 215–238. https://doi.org/10.29140/jaltcall.v20n3.2089

Chen, Y. C., & Wang, L. (2024). Collaborative digital reading in secondary EFL classrooms: Peer scaffolding and reading strategy adoption. *Computer Assisted Language Learning*, *37*(4), 789–813. https://doi.org/10.1080/09588221.2023.2189012

Chinpakdee, M., & Gu, P. Y. (2024). The impact of explicit strategy instruction on EFL secondary school learners' reading. *Language Teaching Research*, *28*(1), 296–319. https://doi.org/10.1177/1362168821994157

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dehghanzadeh, H., Farrokhnia, M., Dehghanzadeh, H., Taghipour, K., & Noroozi, O. (2024). Using gamification to support learning in K-12 education: A systematic literature review. *British Journal of Educational Technology*, *55*(1), 34–70. https://doi.org/10.1111/bjet.13335

Herda, R. K., Margana, M., & Putro, N. H. P. S. (2023). Role of collaborative strategic reading in post-pandemic EFL reading class: A quasi-experimental study. *Issues in Language Studies*, *12*(2), 211–226. https://doi.org/10.33736/ils.5491.2023

Laksanasut, S. (2025). Gamification in ESL/EFL education: Transforming language learning and teaching through play. *TESOL and Technology Studies*, *6*(1), 16–29. https://doi.org/10.48185/tts.v6i1.1562

Lin, C. C., Lin, V., Liu, G. Z., Kou, X., Kulikova, A., & Lin, W. (2020). Mobile-assisted reading development: A review from the Activity Theory perspective. *Computer Assisted Language Learning*, *33*(8), 833–864. https://doi.org/10.1080/09588221.2019.1594919

Lo, C. K., & Hew, K. F. (2023). Fostering higher-order thinking skills through gamified interactive learning: A meta-analysis. *Educational Research Review*, *38*, Article 100500. https://doi.org/10.1016/j.edurev.2022.100500

Maruti, E. S., Afdholy, N., Hidayat, N., Arifani, Y., & Rachmawati, D. L. (2025). Divergent, convergent, or both: Manipulating language use among secondary students in a multilingual setting. *International Journal of Communication & Linguistic Studies*, *23*(2), 39–54. https://doi.org/10.18848/2327-7882/CGP/v23i02/39-54

Matyakhan, T., Chaowanakritsanakul, T., & Santos, J. A. L. (2024). Implementing gamification to enhance reading engagement and reading comprehension of Thai EFL university students. *LEARN Journal: Language Education and Acquisition Research Network*, *17*(1), 212–239.

Ministry of Education, Culture, Research, and Technology. (2022). *Kurikulum Merdeka: Pengembangan kerangka kurikulum dan Capaian Pembelajaran Bahasa Inggris*. Kemendikbudristek RI.

Mykola Petrenko, Melnyk, S., Kovalenko, I., Kovács, I., Tóth, A., Horváth, M., & Szabó, J. (2026). A critical review of Bloom's taxonomy in contemporary educational practices. *International Journal of Education and Social Science Studies*, *2*(1), 9–17. https://doi.org/10.60153/ijesss.v2i1.263

Ndegwa, A., Gutiérrez-Colón, M., & Manegre, M. (2024). Impact of a gamified application on reading comprehension and attitude of Swahili among young learners in Kenya. *Interactive Learning Environments*, *32*(6), 3047–3059. https://doi.org/10.1080/10494820.2023.2165507

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Özdemir, O. (2025). Kahoot! Game-based digital learning platform: A comprehensive meta-analysis. *Journal of Computer Assisted Learning*, *41*(1), Article e13084. https://doi.org/10.1111/jcal.13084

Prados Sánchez, G., Cózar-Gutiérrez, R., Del Olmo-Muñoz, J., & González-Calero, J. A. (2023). Impact of a gamified platform in the promotion of reading comprehension and attitudes towards reading in primary education. *Computer Assisted Language Learning*, *36*(4), 669–693. https://doi.org/10.1080/09588221.2021.1939388

Qiao, S., Chu, S. K. W., & Yeung, S. S. (2025). Understanding how gamification of English morphological analysis in a blended learning environment influences students' engagement and reading comprehension. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2023.2230273

Qiao, S., Yeung, S. S., Shen, X. A., Leung, J. K. L., Ng, D. T. K., & Chu, S. K. W. (2024). How competitive, cooperative, and collaborative gamification impacts student learning and engagement. *Language Learning & Technology*, *28*(1), 1–19. https://doi.org/10.64152/10125/73546

Qiao, S., Yeung, S. S., Zainuddin, Z., Ng, D. T. K., & Chu, S. K. W. (2023). Examining the effects of mixed and non-digital gamification on students' learning performance, cognitive engagement and course satisfaction. *British Journal of Educational Technology*, *54*(1), 394–413. https://doi.org/10.1111/bjet.13249

Rachmawati, D. L. (2024). Meningkatkan keterampilan membaca siswa SD dengan metode extensive reading dan media digital. *IJECS: Indonesian Journal of Empowerment and Community Services*, *5*(2), 150–165. https://doi.org/10.32585/ijecs.v5i2.5612

Rachmawati, D. L. (2025). Digital-based assessment in secondary English classrooms: A pathway to effective evaluation. *J-Shelves of Indragiri*, *7*(1), 205–218.

Rachmawati, D. L., & Fadhilawati, D. (2024). Exploring the challenges faced by EFL students in academic text analysis and critical reading. *JETLEE: Journal of English Language Teaching, Linguistics, and Literature*, *4*(2), 190–206. https://doi.org/10.47766/jetlee.v4i2.2962

Rachmawati, D. L., & Hastari, S. (2022). Formative assessment as an innovative strategies to develop ESP students' writing skills. *Voices of English Language Education Society*, *6*(1), 78–90. https://doi.org/10.29408/veles.v6i1.5174

Rachmawati, D. L., Hastari, S., & Oktafiah, Y. (2022). Strategi dialogic reading untuk meningkatkan interaksi guru dan siswa pada kelas Bahasa Inggris. *JMM (Jurnal Masyarakat Mandiri)*, *6*(3), 2332–2343. https://doi.org/10.31764/jmm.v6i3.8513

Rachmawati, D. L., Ma'ruf, N., & Fadhilawati, D. (2025). Digital formative assessment in EFL classrooms: Redefining learner affect at Islamic institutions. *Register Journal*, *18*(2), 341–368. https://doi.org/10.18326/rgt.v18i2.341-368

Rachmawati, D. L., & Purwati, O. (2021). EFL teachers' attitudes and competence in developing HOTS-based formative assessment. *JEES (Journal of English Educators Society)*, *6*(2), 180–189. https://doi.org/10.21070/jees.v6i2.1060

Rachmawati, D. L., Purwati, O., Anam, S., & Setiawan, S. (2021). Between perception and practice: Encouraging EFL teachers to implement HOTS in secondary classrooms. *TESOL International Journal*, *16*(4), 112–128.

Reynaldo, C., Rachmawati, D. L., & Arifani, Y. (2021). Kahoot! in team mode: Enhancing collaborative engagement in secondary EFL reading. *Indonesian Journal of Applied Linguistics*, *11*(2), 341–352. https://doi.org/10.17509/ijal.v11i2.38901

Sabani, E., Sabilah, F., & Setyaningrum, R. W. (2025). Integrating Mobile-Assisted Language Learning (MALL) in teaching secondary EFL: A study on Indonesian teachers. *Englie: English Learning Innovation*, *6*(2), 88–102. https://doi.org/10.22219/englie.v6i2.32100

Shabani, K. (2023). Revisiting Zone of Proximal Development in digital language learning

environments. *Language Teaching Research*, *27*(4), 890–912. https://doi.org/10.1177/13621688211054321

Shin, D., & Chon, Y. V. (2023). Second language learners' strategic reading behaviors in technology-enhanced environments. *Language Learning & Technology*, *27*(1), 1–25.

Sun, T., & Zhang, L. J. (2023). Promoting critical thinking through collaborative learning in EFL contexts: A systematic review. *Educational Psychology Review*, *35*(2), Article 45. https://doi.org/10.1007/s10648-023-09756-3

Tsai, Y. L., & Lee, C. Y. (2023). Alleviating foreign language reading anxiety through digital peer collaboration. *System*, *115*, Article 103050. https://doi.org/10.1016/j.system.2023.103050

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wahyudi, M. R., Santosa, H., & Indrayani, L. (2025). The influence of using mobile digital tools on EFL students' language learning engagement. *International Journal of Language and Literature*, *9*(1), 45–58. https://doi.org/10.23887/ijll.v9i1.98091

Wang, A. I., & Tahir, R. (2020). The effect of using Kahoot! for learning—A literature review. *Computers & Education*, *149*, Article 103818. https://doi.org/10.1016/j.compedu.2020.103818

Wang, Z., Harun, J., & Yuan, Y. (2024). Enhancing reading instruction through gamification: A systematic review of theoretical models, implementation strategies, and measurable outcomes (2020–2024). *Journal of Information Technology Education: Research*, *23*, Article 28. https://doi.org/10.28945/5394

Wardani, H. K., Mazidah, E. N., Hidayah, B., Alimin, A., & Sa'adah, Z. (2024). Exploring opportunities and barriers of digital tools on EFL academic performance. *Language and Education Journal*, *9*(2), 112–126.

Wijaya, K. F. (2024). The benefits of applying Kahoot media in modern EFL teaching-learning contexts. *Erudita: Journal of English Language Teaching*, *4*(2), 165–178. https://doi.org/10.28918/erudita.v4i2.8896

Yang, Y. (2025). Technology-supported secondary EFL vocabulary acquisition: A systematic review. *Education and Information Technologies*, *30*(13), 17995–18029. https://doi.org/10.1007/s10639-025-13417-8

Yuan, Y., Harun, J., Wong, B. Y., & Li, L. (2025). Mobile-assisted collaborative language learning: Effects on secondary EFL student engagement. *SAGE Open*, *15*(2), 1–16. https://doi.org/10.1177/21582440251329687

Yon, A. E., Halimah, & Nuraida, I. (2024). The effectiveness of collaborative strategic reading (CSR) in improving students' reading comprehension. *Journal of Language Teaching and Research*, *15*(4), 1292–1302. https://doi.org/10.17507/jltr.1504.27

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*, Article 103729. https://doi.org/10.1016/j.compedu.2019.103729

Zhang, R., & Zou, D. (2022). Self-directed language learning with technology: A systematic review. *Computer Assisted Language Learning*, *35*(8), 2050–2088. https://doi.org/10.1080/09588221.2020.1847120