Available online at www at www.jeltl.org
doi: http://dx.doi.org/10.21462/jeltl.v11i1.1951

Journal of English Language Teaching and Linguistics, 11 (1), 2026

Jigsaw Cooperative Learning and Critical Thinking Performance in Pre-Service Teacher Trainees at ESEF Kenitra, Morocco

Ayoub Elbaioui[1], Latifa Hafdi Idrissi[2]
[1-2]*The Higher School for Education and Training (ESEF), Ibn Tofail University, Morocco*
e-mail: ayoub.elbaioui@uit.ac.ma

Received 02 February 2026 | Received in revised form 24 March 2026 | Accepted 26 March 2026

APA Citation:
Elbaioui, A. & Idrissi, L. H. (2026). Jigsaw Cooperative Learning and Critical Thinking Performance in Pre-Service Teacher Trainees at ESEF Kenitra, Morocco. *Journal of English Language Teaching and Linguistics, 11(1), 2026,* 113-129. doi: http://dx.doi.org/10.21462/jeltl.v11i1.1951

Abstract

This study examined whether the jigsaw cooperative learning technique improves critical-thinking performance among pre-service teacher trainees at the École Supérieure de l'Éducation et de la Formation (ESEF) in Kenitra. Using a posttest-only randomized control-group design, 120 trainees were initially assigned to either a jigsaw-based instructional condition or a lecture–discussion condition; 113 participants completed the study (60 experimental; 53 control). Critical thinking was measured immediately after an 11-week instructional period using a researcher-developed 34-item, scenario-based multiple-choice post-test (score range: 0–34) administered via Google Forms. The instrument demonstrated satisfactory internal consistency (Cronbach's α = .877). As score distributions deviated from normality, between-group differences were examined using a Mann–Whitney U test. Findings indicated that trainees in the jigsaw condition achieved higher post-test performance than those in the lecture–discussion condition, suggesting that structured peer interdependence and peer teaching can support stronger critical-thinking performance in pre-service teacher education. These findings imply that teacher-education curricula should move beyond transmission-oriented instruction by embedding structured collaborative routines that make reasoning visible and accountable. It is recommended that teacher-education programs incorporate jigsaw-based activities more systematically in methodology courses and micro-teaching sessions to foster critical thinking in context.

Keywords: cooperative learning, critical thinking, jigsaw, pre-service teacher education, reasoning skills

1. Introduction

Critical thinking is widely regarded as a central goal of higher education because it involves higher-order processes such as analysis, inference, evaluation, interpretation, and explanation, together with the ability to assess evidence and reach warranted conclusions (Facione, 1990; Salido et al., 2025; Halpern, 2014). In teacher education, these capacities are more than desirable academic outcomes. They are part of the professional judgement that pre-service teacher trainees are expected to develop as they interpret classroom situations, justify instructional decisions, adapt lessons to learners' needs, and anticipate the likely consequences of different pedagogical choices. This need has become even more pressing in a context marked by an abundance of information of uneven quality. Recent scholarship has therefore continued to emphasize that critical thinking in teacher education should be taught deliberately rather than assumed to emerge automatically through coursework (Huang & Sang, 2023; Marangio et al., 2024; Önal, 2025).

An instructional method that seems particularly suited to this purpose is cooperative learning, particularly the jigsaw method which is defined as each member of the group is made responsible for learning one part of the content and then teaching that part to the other group members (Aronson, 1978; Vives et al., 2024). The benefit of this structure is the types of thinking it makes necessary. Learners must think about the organization of their ideas, explain different concepts, rationalize interpretations, provide answers to questions, and integrate different views. All of these requirements are closely related to critical thinking. Recent studies in teacher education demonstrate that cooperative learning has positive outcomes on pre-service teachers' content knowledge, teaching self-efficacy, motivation, academic engagement, and even wider professional development (Chan et al., 2021; Fernández-Río et al., 2022; Chan et al., 2024; Salvador-Garcia & Sanahuja Ribés, 2025). In a more direct sense, Silva et al. (2023) found that critical and creative thinking, as well as critical thinking dispositions, improved in the jigsaw group, and cooperative learning methods showed this improvement. In this sense, jigsaw is more than just a method of putting students in groups. Jigsaw is a more structured method that makes participation more accountable, more dialogic, and more cognitively rigorous (Sekano et al., 2026; Liu et al., 2024).

This is particularly relevant in language teacher education. Although the present intervention focused on critical-thinking performance, it was implemented in a language-teaching methodology course in which trainees are expected not only to acquire pedagogical knowledge, but also to justify their instructional choices with evidence and principled reasoning. In language education, recent review work suggests that critical thinking is more likely to develop when it is explicitly embedded in pedagogy rather than treated as an incidental outcome of language practice (Yin et al., 2023). In a related EFL context, Banaruee et al. (2025) reported that jigsaw-based listening tasks improved learners' listening performance while also promoting strong peer interaction. Such findings are important because future teachers need to experience the kinds of interaction-rich pedagogies they may later be expected to design and implement in their own classrooms.

At the same time, the place of critical thinking in Moroccan higher education remains uneven. Ouahani & Hiba (2023) note that research on critical thinking in the Moroccan context is still relatively limited, while Beniche (2025) shows that university professors continue to face substantial challenges in trying to foster critical thinking in practice. This matters for teacher education in particular. Pre-service teachers are expected to become reflective professionals, yet they may still be trained in environments where traditional practices remain influential and where sustained opportunities for higher-order reasoning are not systematically built into everyday instruction.

When taken as a whole, the recent body of work shows both promise and a clear need for more evidence. The research on cooperative learning in the teacher education context has most notably documented the benefits of knowledge development, self-efficacy, self-motivation,

engagement, and professional development (Chan et al., 2021; Fernández-Río et al., 2022; Chan et al., 2024; Salvador-Garcia & Sanahuja Ribés, 2025). Conversely, research on the development of critical thinking skills in pre-service teachers has focused on broader pedagogical frameworks, including technology-facilitated peer assessment, digital storytelling, and reflective teaching (Barahona et al., 2023; Isaacs et al., 2024; Marangio et al., 2024). What has been most absent is practice-based empirical research on the extent to which jigsaw-based instruction can foster pre-service teachers' critical thinking in a Moroccan teacher education context. This is the gap that the present study aims to address.

Accordingly, this study examines whether jigsaw-based instruction improves critical-thinking performance among pre-service teacher trainees at ESEF Kenitra, Morocco, compared with a lecture–discussion condition. Using a posttest-only randomized control-group design and a researcher-developed scenario-based critical-thinking post-test, the study investigates whether trainees exposed to jigsaw achieve higher post-intervention critical-thinking scores than those taught through more traditional instruction. By offering performance-based evidence from a Moroccan teacher-education context, the study contributes to current discussions on how teacher-training programs can more deliberately foster critical thinking through structured and accountable collaboration. Accordingly, the study addresses the following research question:

"Does jigsaw cooperative learning significantly improve the critical-thinking performance of pre-service teacher trainees, compared with lecture–discussion instruction?"

2. Literature Review

Across teacher education, critical thinking (CT) is typically framed as a cluster of higher-order cognitive processes—analysis, inference, evaluation, interpretation, and explanation—applied to judging evidence and claims (Facione, 1990; Ennis, 1991; Paul & Elder, 2014; Facione, 2015). For pre-service teachers, CT supports routine professional judgment: diagnosing learning needs, selecting or adapting methods, evaluating classroom evidence, and anticipating pedagogical consequences—capacities cultivated through reflective practice that systematically questions assumptions and justifies decisions (Schön, 1983, 1987; Brookfield, 2012; Lipman, 2003). Programmatically, this moves curricula from "hoping" for CT to designing for it—embedding tasks that require explanation, comparison, counter-argument, and evidence use, alongside assessments aligned to explicit CT outcomes (Biggs & Tang, 2011; Halpern, 2014; Paul, 1993).

The intentional development process requires proof of learning, and evidence is needed for CT's assessment through instruments that focus on reasoning instead of recall and are defendable through psychometric studies. (Facione, 1990; Evans, 2020; Liu et al., 2016; Gelerstein et al., 2016). Locally designed assessments can closely align with the specific CT constructs (e.g., argument evaluation, causal reasoning), so long as the items are written clearly, the cognitive demand is of the right level, and there is validity/reliability evidence (Facione, 1990; Biggs & Tang, 2011; DeVellis, 2016). Within that logic of design for assessment, collaborative learning and especially Jigsaw, is posited to enhance measurable CT through cognitive strength mechanisms, namely elaborated explanations, peer teaching, perspective alignment, and immediate feedback among interdependently structured groups (Aronson, 1978; Johnson & Johnson, 2009; Johnson et al., 2007; Webb, 2009). Learners who are prepared to teach their subtopic and interweave their peers' contributions perform the very operations that are the targets of CT assessments, including justifying, weighing, and reconciling, and synthesizing (Prince, 2004; Johnson & Johnson, 2009).

Empirically, active/cooperative designs often outperform lecture on objective learning outcomes across disciplines, with meta-analytic syntheses reporting moderate average effects (Prince, 2004; Freeman et al., 2014; Slavin, 1995, 2014). Studies that isolate the Jigsaw structure likewise show advantages on content learning and higher-order performance relative to traditional

instruction in pre–post or post-only comparisons (Karacop & Doymus, 2013; Karacop & Diken, 2017; Sağsöz et al., 2017). In higher education—including pre-service teacher education—Jigsaw-based cooperative learning has been associated with improved critical-thinking dispositions and thinking performance (Silva et al., 2023), with proposed mechanisms emphasizing positive interdependence and individual accountability, the expert/home-group cycle that requires explanation and conceptual reorganization, and peer dialogue that elicits justification, evaluation, and synthesis (Johnson et al., 2007; Johnson & Johnson, 2009; Webb, 2009).

Despite promising results, the context and the population under consideration are significant. Reviews state the inadequate representation of the pre-service cohorts, variable fidelity of implementation, and the scarcity of studies conducted in non-Anglophone or Global South contexts, thus raising questions regarding the transportability of these studies across different institutional cultures and the varying demands of the curricula (Lorencová et al., 2019; Mrabti et al., 2023; Qasserras & Qasserras, 2023). Furthermore, there is a methodological gap between studies pertaining to the perception of critical thinking (CT) and studies that are evidence-based regarding the achievement of CT; the latter is crucial in supporting the assertion that the cooperative structures influence cognitive change (i.e., attitudes) (Freeman et al., 2014; Prince, 2004). The aforementioned gaps are what drive the current article's post-test experimental comparison of Jigsaw and lecture with CT instrument that was developed by the author, that is, the instrument is aligned to the accepted constructs, and the instrument is reported in a manner that reflects the utilization of appropriate psychometric and statistical techniques (Facione, 1990; Evans, 2020; Fritz et al., 2012).

3. Research Methods

3.1 Research Design

This study utilized a posttest-only randomized control group design to measure the effect of the jigsaw cooperative learning technique on the critical thinking abilities of pre-service teacher trainees. Before the instructional period commenced, participants from the Semester 4 participant list were randomized at the individual level to the jigsaw experimental group or the control group (lecture-discussion) using stratified permuted-block randomization (strata = administrative groups; 1:1 allocation). Following randomization, trainees were split into two classes based on the group role: the experimental group received jigsaw instruction in one classroom during the assigned time, while the control group received lecture-discussion instruction in another classroom at a different time. In order to minimize contextual variation, one professor (the researcher) taught both groups, taught the same course, and provided the same number of instructional hours. To avoid the possibility of testing and sensitization effects due to the inclusion of a pretest, a posttest-only design was utilized (Campbell & Stanley, 1966; Shadish et al., 2002). The outcome measure was evaluated immediately after the intervention using a standardized scoring key for the critical thinking post-test.

3.2 Participants and Setting

The sample consisted of 113 pre-service teacher trainees enrolled in their semester 4 of the teacher education program at the École Supérieure de l'Éducation et de la Formation (ESEF), Kenitra. At the commencement of the semester, the administration handed the researcher the official Semester 4 roster (N = 361), which provided details on trainees' institutional IDs and their administrative group affiliations. From this roster, 120 trainees were randomized for the study, and study conditions were assigned via individual-level stratified permuted-block randomization (strata = administrative groups; 1:1). Of those randomized, 113 participants (60 in the jigsaw condition; 53 in the lecture–discussion condition) were included in the analysis. Seven trainees assigned to the

control group withdrew from the study due to external factors that hindered their continued attendance. All participants were enrolled in the same Language Teaching Methodology course during the study period, and the study was conducted at the ÉSEF within the standard teaching framework. Analyses were conducted on participants who were post-test completers, and the potential impact of the shared instructional setting was factored in when drawing conclusions.

3.3 Randomization and Allocation

The Semester 4 roster provided by the administration was used for randomization, containing 361 students along with their institutional IDs and lists of their administrative group affiliations. The first step was to use the random procedure in Microsoft Excel (RAND function and sorting) to randomly select 120 trainees. The next step involved assigning participants to experimental and control conditions using stratified permuted block randomization, where the admin group was used as the stratification variable. Each stratum contained permuted blocks of size 4 or 6 to keep the balance random and assigned in a 1:1 ratio (60 experimental; 60 control) for randomization. The allocation list gave the instructors the ability to assign trainees to the designated classes for each condition before instruction started.

3.4 Intervention

The experimental group had an active learning technique, the Jigsaw technique of cooperative learning (Aronson, 1978), integrated into the Semester 4 Language Teaching Methodology course, which was implemented over 11 consecutive weeks. After selecting and randomizing the trainees, they were split into two groups, functioning as a measure to ensure the integrity of the treatment and limit the risk of contamination between groups. Each group used different classrooms and times. Prior to the initial complete jigsaw session, the professor provided a specific orientation to the experimental group. He explained the principle of positive interdependence and the importance of individual accountability, and described the repetitive cycle of activities: home group, expert group, and back to home group. He also explained functional roles, such as timekeeper, note-taker, facilitator, and presenter. The instructor defined collaboration roles and established clear collaboration norms, including equal participation, respectful turn-taking, and the use of evidence to support explanations. In Week 1, trainees were instructed to complete a short cycle exercise to avoid the loss of instructional time as a result of procedural confusion.

Each jigsaw session had a similar overarching instructional approach from week to week and remained closely aligned with the course content and the related course learning outcomes. In order to ensure responsibility and cohesion over time, the trainees worked in home groups of approximately four members. Trainers began each session by explaining the topic of the lesson, the objectives of the unit, and key points. He then classified the lesson content into four subtopics, each with equal scope and cognitive complexity. For example, in a unit on behaviorism, the subtopics would include classical conditioning, operant conditioning, reinforcement schedules, and applications in the classroom. Each home group member was assigned one subtopic and one reading package, which was prepared by the instructor. The reading materials had analysis and evaluation guiding questions, which were not designed to trigger simple recall, but to encourage trainees to analyze and evaluate the material, whereby they were expected to determine the central definition of the questions, and the claims vs. the evidence provided in support of the claim, provide classroom examples, and address potential misconceptions or rival interpretations.

As part of the expert-group phase, trainees left their home groups for a short period to join other group members who had the same subtopic assignment. Here, they worked through the content of the material, cleared up meanings, and negotiated different interpretations. In each group, they also prepared brief summaries of the material to be used in the later peer teaching segment. This phase took approximately 30-40 minutes. To enhance individual accountability, each

trainee was required to submit a short written preparation, usually a bullet-point summary and an illustrative example, before returning to the home group. The instructor aimed to observe each group and pose some leading and guiding questions. When he noticed some conceptual inaccuracies, gaps in applying the concepts to the real-world scenario, or areas of weakness, he provided some targeted support. However, he kept the majority of his support to a minimum. He did not do the work for the trainees, but did support and clarify, so the trainees kept most of the work with the responsibility of meaning-making.

In the home-group teaching phase, each "expert" taught their subtopic to the home-group members, who received, asked follow-up questions, and took notes. In order to encourage deeper processing and to avoid surface processing, groups were given short integration prompts after each mini lesson (for example, stating the strongest evidence in support of a given claim, applying the concept to a real-life classroom decision, or stating a limitation/counterexample that would explain). These prompts required trainees to relate subtopic ideas, explain and justify their reasoning, and evaluate the strength of the explanation, all of which are behaviors that are associated with higher-order or critical thinking processes. As an additional synthesis task, trainees collaboratively produced a diagram summarizing the unit content (Appendix A). In this diagram activity, participants were able to show the relationships among concepts, such as cause and effect, contrasts, and hierarchies, and organize and prioritize information. In class, diagrams were reviewed, and the instructor provided informal feedback pertaining to issues of clarity, accuracy, and completeness.

Trainees' teaching at the Jigsaw began to demonstrate improvement in implementing Jigsaw teaching peers each week due to the teaching model's fidelity. This was due to the fact that there was a consistent lesson rhythm, equal allocation of time to each phase of the jigsaw (overview, expert preparation, home group teaching, wrap up), and the same readings, teaching prompts, and integration questions provided to trainees. Each weekly lesson rhythm went towards closing the lesson and provided a consolidation of the key ideas, addressed teaching misconceptions that emerged from the peer teaching, and framed the unit content to the teaching goals of the course. Credited to the trainees' teaching, the gaps during the transition to teaching and the imbalanced participation were reduced, leaving more time to engage in discussions and provide more substantive explanations.

The control group was taught the same course units for the same 11-week duration by employing a lecture-discussion model with the use of regular course materials. The instructor commenced each session with an introduction to the topic, followed by direct instruction of the key concepts and subtopics. The instructor guided the whole-class discussion by posing questions and responding to the class's short clarifying questions. The majority of the students were passive participants by listening, taking notes, and responding to the instructor's questions during the plenary discussion. Participation and engagement were low, and activities of the session, where students were asked to work individually or in pairs, were conducted to reinforce comprehension, with the aim of limiting constructive peer-to-peer learning. A consistent weekly schedule and instructional hours were employed to control for task engagement and scope of the curriculum, with the primary difference being the addition of the collaborative jigsaw activity for the experimental condition compared to teacher-led exposition and discussion for the control condition.

3.5 Instruments

To assess post-intervention critical thinking, we used a researcher-developed critical thinking post-test as the primary instrument. All participants completed this assessment immediately after the 11-week instructional period. The test consisted of 34 scenario-based multiple-choice items designed to evaluate core facets of critical thinking relevant to teacher education, including analysis,

inference, evaluation, inductive reasoning, and deductive reasoning. Each item presented a short statement or classroom-related situation followed by response options requiring participants to interpret information, draw a warranted conclusion, or judge the strength of an explanation or argument. In this way, the assessment aimed to reflect the kinds of reasoning demands pre-service teachers routinely face when interpreting evidence, weighing alternatives, and making pedagogical decisions.

In developing the assessment, we followed best-practice recommendations commonly adopted when standardized critical-thinking measures are impractical. For example, Mapeala & Siew (2015) argue that limited availability of suitable instruments may necessitate developing a context-appropriate test that is "cost-effective and easy to administer." Streamed from the established frameworks for critical-thinking assessment (e.g., Facione, 1990; Paul & Elder, 2006), items were purposely constructed to prioritize clarity, focus, and integration of the subject matter distributed across the varied clusters of thinking and reasoning. The questions balanced the cognitive load of the participants while encouraging them to engage in deep reading, reflective thinking, and reasoning supported by evidence.

To support content validity, the initial item pool underwent expert review by Dr. Latifa Hafdi Idrissi, an associate professor of ELT methodology at ESEF, who evaluated item clarity, alignment with the targeted critical-thinking skills, and suitability for the trainees' level. Her feedback guided iterative revisions, including simplifying wording, removing ambiguous phrasing, and ensuring that each scenario required warranted inference and evaluation of claims. The revised version was then piloted with a small group of trainees (n = 5) to confirm comprehensibility and time feasibility. Pilot participants were not included in the main study sample. Based on pilot feedback, items that appeared difficult to interpret or contained wording likely to hinder understanding were refined or removed prior to final administration (DeVellis, 2016).

In order to maintain content validity, the initial pool of items was submitted to an expert review by Dr. Latifa Hafdi Idrissi, Associate Professor of ELT methodology at ESEF, who evaluated the clarity of the items, the alignment of the items to the critical thinking skills targeted, and the appropriateness of the items to the trainees' level. Her suggestions were the basis of several revisions, such as simplification of language, the removal of vague language, and the modification of scenarios to ensure that they demanded justification and evaluation of claims. The revised version was also piloted with 5 pre-service trainees to ascertain comprehensibility and time on task. The pilot participants were not part of the main study sample. As pilot participants suggested, items that were perceived as incomplete or too vague were revised or omitted before the final administration (DeVellis, 2016).

Items were scored dichotomously (1 = correct; 0 = incorrect), producing total scores from 0 to 34, with higher scores indicating stronger critical-thinking performance. Administration was conducted online via Google Forms under supervised classroom conditions. Participants were given 30 minutes, could navigate back and forth between items before final submission, and were allowed scratch paper for notes. To reduce the likelihood of answer-sharing among closely seated students, the form was configured to randomize the order of items across test-takers. Responses were automatically recorded and exported for analysis. Internal consistency reliability for the 34-item test was examined using KR-20 (reported as Cronbach's α, which is equivalent for dichotomously scored items) in the present sample.

3.6 Data Analysis Techniques

Following data collection, responses were exported from Google Forms and prepared for analysis in IBM SPSS Statistics (Version 27). Each participant's critical-thinking post-test was scored using a predefined answer key, with items coded dichotomously (1 = correct, 0 = incorrect) and

summed to produce a total score ranging from 0 to 34. Descriptive statistics were computed separately for the experimental and control groups to provide an initial picture of performance patterns on the post-test, including means and standard deviations, as well as medians and interquartile ranges. To support the quality of the measurement used in this study, internal consistency was examined using Cronbach's alpha, which is equivalent to KR-20 when items are scored dichotomously (Cronbach, 1951; Kuder & Richardson, 1937). In addition, item-level checks were conducted by inspecting item difficulty (proportion correct) and corrected item–total correlations, allowing us to identify items that were unusually easy/difficult or weakly related to the overall score.

Before comparing groups, we examined whether the distribution of total scores met the assumptions required for parametric testing. Normality was assessed using Shapiro–Wilk tests alongside visual inspection of Q–Q plots (Shapiro & Wilk, 1965). As these checks indicated departures from normality, we adopted a nonparametric approach. Group differences in post-test critical-thinking performance were therefore tested using the Mann–Whitney U test, which is appropriate for comparing independent groups when distributional assumptions are not satisfied (Mann & Whitney, 1947). All tests were conducted using a two-tailed significance level of .05. To accompany significance testing with an estimate of practical importance, effect size was calculated as r using the standardized z value from the Mann–Whitney procedure ($|z|/\sqrt{N}$), consistent with recommendations for nonparametric effect-size reporting (Fritz et al., 2012).

3.7 Ethical Considerations

The ethical considerations involved in this study have been in accordance with the ethical codes of conduct for research in education. An explanation of the study's purpose and the study's procedures was given to the trainees, and they were assured the right to remain anonymous and were informed of the right to decline participation or withdraw from the study at any time without any negative consequences. The trainees gave the researchers consent. The researcher was able to sample and randomize using the official Semester 4 list of enrolments given by the administration. For the purpose of organizational research, only institutional IDs and group IDs were used, and the names of students were not included in the dataset for analysis. No individual trainees were identified as a result of the anonymization of the data, and the results were expressed in total or as a whole. The data were also secured and only accessed for the purposes of research.

The researchers intended to design the study to be respectful, low-risk, and appropriate for an educational environment. Due to the random assignment, the trainees were split into two classes for teaching purposes based on the conditions. This division was purely for the conduct of the study and was in no way administrative. The researchers attempted to mitigate unfairness and discomfort by offering trainees the opportunity to complete the critical thinking post-test, which was administered in a controlled setting via Google Forms. Throughout the study, the researchers maintained the dignity and privacy of the participants with regard to the test and the study.

4. Results

4.1 Data Screening and Preliminary Checks

A total of 113 pre-service teacher trainees provided complete post-test data and were included in the analyses (control n = 53; experimental n = 60). The critical-thinking post-test was scored using a predefined answer key, with items coded dichotomously (0 = incorrect, 1 = correct) and summed to yield total scores ranging from 0 to 34. All observed values fell within this expected range. Score distributions were inspected for unusual patterns and extreme values using boxplots and

standardized scores. No extreme outliers requiring removal were identified; therefore, all scores were retained as plausible indicators of post-test performance.

4.2 Instrument Quality

Internal consistency for the 34-item critical-thinking post-test was satisfactory in the present sample (N = 113), with Cronbach's α = .877 (k = 34), which is equivalent to KR-20 for dichotomously scored items (Cronbach, 1951; Kuder & Richardson, 1937). Item difficulty indices (proportion correct) indicated a broad spread of difficulty across the instrument (range = .18–.88), suggesting that the test included both challenging and more accessible items. Corrected item–total correlations were generally acceptable; however, a small number of items showed weak discrimination (CITC < .20), indicating that these items were less strongly aligned with the overall score. Deleting the weakest item produced only a negligible improvement in reliability; therefore, all items were retained for the primary analyses and flagged for potential refinement in future administrations.

Table 1: Internal Consistency and Item Functioning of the 34-Item Critical Thinking Post-Test

Measure	Result
Number of items (k)	34
Internal consistency (Cronbach's alpha / KR-20 equivalent)	.877
Item difficulty range (p = proportion correct)	.18–.88
Corrected item–total correlation (CITC) range	.024–.674
Items flagged as very difficult (p < .20)	Q25 (p = .18)
Items flagged as very easy (p > .85)	Q29 (p = .88)
Items flagged as weak discrimination (CITC < .20)	Q20 (CITC = .133), Q25 (CITC = .024), Q29 (CITC = .199)

Note. p values represent the proportion of correct responses for each dichotomously scored item (0/1). CITC = corrected item–total correlation. For the weakest items, Cronbach's alpha if item deleted was .879 (Q20), .880 (Q25), and .877 (Q29), indicating negligible improvement in reliability; therefore, all items were retained for the main analyses and flagged for potential refinement in future administrations.

4.3 Descriptive Statistics by Group

Descriptive statistics indicated higher post-test performance in the experimental group than in the control group (Table 2). The experimental group obtained a higher mean score (M = 23.30, SD = 7.32) and median (Mdn = 25, IQR = 13) than the control group (M = 18.45, SD = 5.33; Mdn = 17, IQR = 6). Observed scores ranged from 10 to 32 in the control group and from 11 to 34 in the experimental group. Group score distributions are displayed in Figure 1.

Table 2: *Descriptive Statistics for Critical Thinking Post-Test Scores by Group*

Groups	n	M	SD	Mdn	IQR	Min	Max
Control (lecture-based)	53	18.45	5.33	17	6	10	32
Experimental (jigsaw)	60	23.30	7.32	25	13	11	34

Note. M = mean; SD = standard deviation; Mdn = median; IQR = interquartile range. Observed post-test scores ranged from 10 to 34, with higher scores indicating stronger critical-thinking performance.

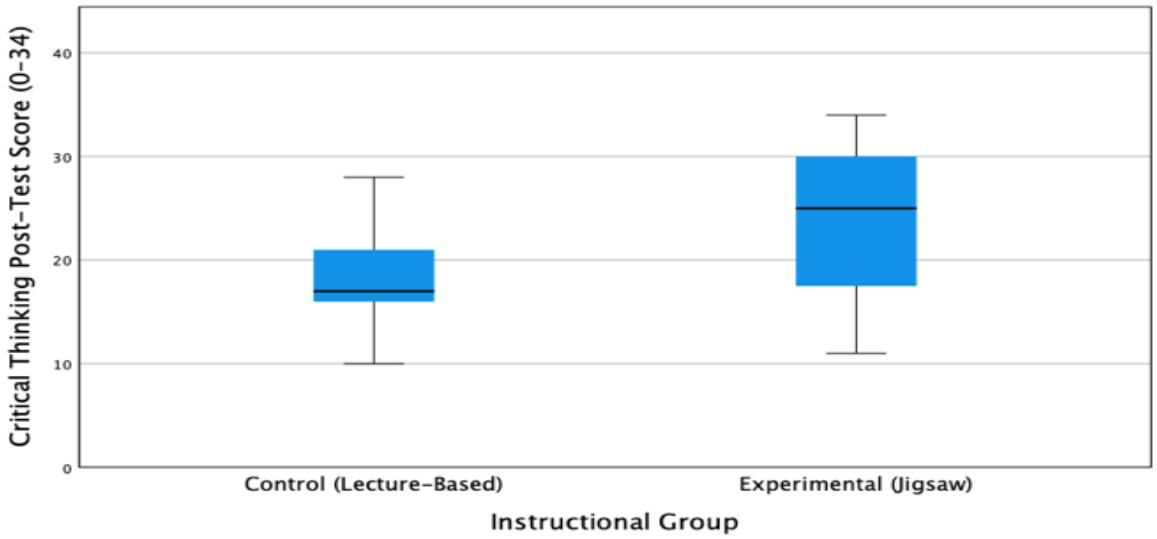

Figure 1: *Boxplot of Critical Thinking Post-Test Scores by Group*

Note. The box represents the interquartile range (IQR), the center line indicates the median, whiskers extend to the most extreme non-outlier values, and points (if shown) indicate outliers.

4.4 Group Difference in Post-Test Performance

Shapiro–Wilk tests indicated statistically significant departures from normality in both groups (control: $W = .926$, $p = .003$; experimental: $W = .926$, $p = .001$). In light of these distributional characteristics, an independent-groups nonparametric procedure was used. A Mann–Whitney U test (Mann & Whitney, 1947) indicated a statistically significant difference in critical-thinking post-test performance between conditions. The experimental group showed higher ranks (mean rank = 67.37) than the control group (mean rank = 45.26), $U = 968.00$, $z = -3.59$, $p < .001$. The effect size was moderate ($r = .34$), indicating a practically meaningful difference in performance between groups (Cohen, 1988; Fritz et al., 2012).

5. Discussion

The present findings suggest that jigsaw-based cooperative learning can meaningfully support pre-service teacher trainees' critical-thinking performance when compared with lecture–discussion instruction. On the post-intervention assessment, the experimental group achieved higher total scores than the control group, and this difference was statistically reliable using a Mann–Whitney U test. Importantly, the observed effect was not only statistically significant but also practically relevant (i.e., a moderate effect size), indicating that the advantage associated with jigsaw is large enough to matter in authentic teacher-education settings rather than reflecting a purely "technical" difference (Fritz et al., 2012). This pattern is consistent with recent research showing that cooperative learning can improve important outcomes in pre-service teacher education. For instance, Chan et al. (2021) identified improvements in teaching self-efficacy and content knowledge among EFL pre-service teachers, whereas Chan et al. (2024) indicated positive effects on learning motivation and academic engagement. In agreement with this perspective, Fernández-Rio et al. (2022) examined the findings of several research studies and identified academic, interpersonal, transversal, and professional outcomes of cooperative learning in teacher education. Silva et al. (2023) noted that the jigsaw and group investigation techniques, as opposed to the lecture-based instruction, were better at fostering critical and creative thinking. The present result aligns with recent literature on the topic. The

existing evidence, in a scenario where critical thinking is often declared as an educational goal rather than systematically taught or assessed, supports the view that collaborative tasks that are designed to achieve interdependence and individual accountability can foster conditions in which trainees practice higher-order reasoning in a transferable and sustained manner.

From an ELT perspective, the observed advantage for the jigsaw condition is pedagogically meaningful because critical thinking in language-teacher education is expressed through how trainees interrogate methodological claims, evaluate classroom evidence, and justify task and assessment choices—not merely through decontextualized reasoning ability (Borg, 2003; Farrell, 2015). This interpretation is also consistent with recent teacher-education scholarship, which increasingly frames critical thinking as context-sensitive professional reasoning that should be cultivated explicitly in pre-service programs rather than assumed to emerge indirectly through exposure to course content (Huang & Sang, 2023; Önal, 2025). In a language-teaching methodology course, well-structured jigsaw routines can operationalize communicative and task-based principles by creating sustained peer interaction in which trainees must clarify meaning, respond to questions, and produce coherent explanations—interactional practices central to communicative language teaching (Long, 1996; Swain, 1995). When designed with positive interdependence and individual accountability, such collaboration is more likely to elicit elaborated explanation and academically focused dialogue (Johnson & Johnson, 2009; Webb, 2009), and EFL research similarly reports that jigsaw listening tasks can support language development while promoting social interaction (Banaruee et al., 2025). In a related pre-service EFL context, Chan et al. (2024) likewise found that cooperative learning enhanced trainees' motivation and academic engagement, suggesting that structured peer work may create the motivational conditions under which critical and pedagogically productive participation becomes more likely. Because the present study assessed critical-thinking performance rather than linguistic proficiency, future work should pair critical-thinking measures with ELT-relevant outcomes (e.g., speaking interaction quality or listening comprehension) to test whether these cooperative mechanisms translate into parallel language gains (Long, 1996; Swain, 1995).

A plausible explanation for this advantage lies in the cognitive work that jigsaw requires learners to perform. Because each trainee is responsible for mastering a portion of content and then teaching it to peers, the task naturally prompts explanation, justification, and the organization of ideas into coherent messages—processes that align closely with higher-order reasoning. Research on "learning by teaching" suggests that preparing to teach and actually teaching can enhance understanding and transfer because it pushes learners toward active processing rather than passive reception (Fiorella & Mayer, 2013). The present result is also compatible with recent empirical work showing that cooperative-learning designs such as jigsaw can promote critical and creative thinking more effectively than lecture-based instruction when learners are required to explain, justify, and synthesize ideas collaboratively (Silva et al., 2023). From a cooperative-learning perspective, the jigsaw structure is also consistent with social interdependence theory: when learners' success becomes linked, they are more likely to invest effort, take responsibility, and engage in purposeful interaction (Johnson & Johnson, 2009). Recent review evidence in cooperative language learning likewise suggests that the quality of interaction within structured collaboration is central to its effectiveness, including its potential to support critical thinking, communication, and reflective engagement (Liu et al., 2024). In this sense, the observed group difference is not simply a matter of "working in groups," but of working in groups under an instructional design that makes thinking visible, shared, and consequential (Aronson, 1978; Johnson & Johnson, 2009).

The findings also fit well within broader cooperative-learning scholarship, showing that well-structured collaboration tends to outperform competitive or individualistic structures on learning

outcomes, particularly when tasks require elaboration and meaningful peer interaction rather than simple division of labor (Gillies, 2014; Johnson & Johnson, 2009; Slavin, 2014). Recent work in teacher education points in the same direction. Fernández-Río et al. (2022) found that cooperative-learning interventions support future teachers' global development across academic, interpersonal, and professional domains, while Salvador-Garcia & Sanahuja Ribés (2025) reported that pre-service teachers view cooperative learning as valuable for the development of collaboration, communication, and other 21st-century skills. At the same time, recent studies that focus more directly on critical thinking in pre-service teacher education have often examined other structured pedagogies, such as technology-scaffolded peer assessment or digital storytelling (Barahona et al., 2023; Isaacs et al., 2024). In this respect, the present study adds something more specific: it provides performance-based evidence for jigsaw as a structured cooperative design in Moroccan pre-service teacher education rather than for a broader cluster of student-centered strategies. Beyond cooperative learning specifically, syntheses of active learning research similarly conclude that student-centered approaches that promote discussion, explanation, and application often yield stronger objective outcomes than lecture-dominant instruction (Freeman et al., 2014; Prince, 2004). Taken together, these perspectives reinforce an important point for teacher education: when trainees repeatedly participate in dialogue that requires evidence-based claims, evaluation of alternatives, and explanation to others, critical thinking becomes less of an abstract ideal and more of a routine practice embedded in everyday learning.

Even though this paper focuses on post-test outcomes, the implementation experience helps elucidate how jigsaw is likely to yield cognitive gains. The research on cooperative learning indicates that the effectiveness of the jigsaw technique is dependent on how it is designed which means having specific operational guidelines, sequencing activities, and determining teacher roles that focus on and regulate equitable involvement and constructive discussions, rather than assuming that group work is automatically beneficial (Baloche & Brody, 2017; Baines et al., 2017; Gillies, 2014). In concrete terms, the current findings advocate for the necessity of incorporating jigsaw (and similarly designed collaborative routines) into the preparation of preservice teachers not only as content to be learned but as pedagogy to be practiced. This puts several demands on teacher education institutions.

Accordingly, the findings suggest that teacher education programs should make deliberate curricular and pedagogical adjustments. First, the jigsaw technique should be embedded in all of the content courses so that participants engage in structured cooperative learning, rather than just having it explained to them. Second, teacher educators should ensure that the jigsaw classroom technique is linked to micro-teaching and lesson planning so that participants can acquire the skills of discussion leadership, participation control, and oversight of explanations. Third, it is imperative that course content and assessments are designed and described to include elements that demonstrate considered engagement through peer assessments, group achievements, and decisions made through instructional reasoning. These suggestions corroborate the recent findings that focus on technology-assisted peer assessments and digital storytelling, which can help develop pre-service teachers' critical thinking, as they have to externalize, revise, and justify their thinking (Barahona et al., 2023; Isaacs et al., 2024). Additionally, pre-service teachers view cooperative learning as a worthwhile incentive for communication, collaboration, and future professional practice (Salvador-Garcia & Sanahuja Ribés, 2025). Design priorities for implementation include the delegation of defined roles, the imposition of time limits, group output that is visible (e.g., organized notes and diagrams), and mechanisms to ensure accountability for each trainee's contribution. These are the elements of design that deepen the rationale for cooperative learning to foster high-order thinking.

Several limitations should guide interpretation. First, the absence of a pretest limits claims about growth because baseline critical-thinking performance cannot be directly demonstrated in a post-test-only design. Second, differential attrition—where withdrawals occurred in the control condition—may have introduced bias if those who did not complete the study differed systematically from completers. Third, although the custom instrument demonstrated strong internal consistency, any locally developed test benefits from continued validation work, including replication across cohorts, item refinement, and additional evidence of construct validity (Taber, 2018). Finally, implementation fidelity was not treated as a measured variable; future work could quantify fidelity (e.g., observation checklists) and examine whether stronger implementation predicts larger learning gains (Gillies, 2014; Slavin, 2014; Johnson & Johnson, 2009).

Future research could extend the present contribution by adopting pre–post or longitudinal designs to capture trajectories and durability of gains, comparing jigsaw with other cooperative structures to identify which design features are most strongly associated with critical-thinking development, and examining potential mediators (e.g., quality of peer explanation, distribution of talk, accountability practices) to clarify the mechanisms linking structured cooperation to reasoning outcomes (Fiorella & Mayer, 2013; Johnson & Johnson, 2009; Slavin, 2014). It would also be valuable to compare jigsaw with other recent critical-thinking pedagogies used in teacher education, including peer-assessment and digital-storytelling designs, in order to identify which forms of structured participation are most effective for different learning outcomes (Barahona et al., 2023; Isaacs et al., 2024). In teacher-education contexts specifically, it would be valuable to explore whether sustained exposure to jigsaw not only improves trainees' critical-thinking performance but also shapes their readiness to enact learner-centered pedagogy during practicum and early-career teaching (Baloche & Brody, 2017).

6. Conclusion

The authors contend that the study demonstrates that jigsaw cooperative learning fosters greater critical thinking among pre-service teacher trainees than a lecture–discussion approach. In this posttest-only control-group design, the jigsaw group outperformed the comparison group on the 34-item scenario-based critical thinking post-test, which represented a statistically significant and practically meaningful difference. This study adds to the body of research that supports the assertion that structured peer interdependence creates an environment that facilitates the development of higher-order reasoning.

The work of the researchers shows that the instrument for the evaluation of critical thinking skills was relevant in this context. The researchers were able to determine this due to the test's evaluation of internal consistency and the evaluation of the items that showed a good distribution in construction difficulty and discrimination, as well as a low number of items that were evaluated for improvement in future evaluations. This creates confidence in the stated outcome and shows that local assessments designed with a good level of detail provide valid assessments of learning when there is no Greater test available.

For teacher training, the consequences are useful as well as theoretical. If the goal of the program is to develop future teachers who are critical and reflective, then structures that facilitate cooperative learning, such as the jigsaw technique, should be incorporated, not as one of the topics to be discussed, but as one of the central teaching strategies that the trainees should participate in. In particular, this should be the case in courses on the methodology of teaching languages. These structures will allow future teachers to practice the teaching of the explanation of ideas, the evaluation of the options, and the justification of the decisions taken in a teaching sense, which is the reflection on the strategies to promote learning based on evidence. Future studies should report

similar findings across different cohorts and different institutions. These studies should use pre-post or longitudinal studies to measure growth and determine how long the growth continues. These studies should measure how closely the study was implemented to determine the conditions under which the cooperative learning methodology had the greatest impact on developing critical thinking skills in pre-service teacher training.

REFERENCES

Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Sage Publications.

Baines, E., Blatchford, P., & Kutnick, P. (2017). *Promoting effective group work in the primary classroom: A handbook for teachers and practitioners* (2nd ed.). Routledge.

Baloche, L., & Brody, C. M. (2017). Cooperative learning: Exploring challenges, crafting innovations. *Journal of Education for Teaching, 43*(3), 274–283. https://doi.org/10.1080/02607476.2017.1319513

Banaruee, H., Farsani, D., & Khatin-Zadeh, O. (2025). Empowering teacher education with jigsaw techniques: Insights from language learners' listening tasks. *Frontiers in Education, 10*, Article 1503888. https://doi.org/10.3389/feduc.2025.1503888

Barahona, C., Nussbaum, M., Martin, V., Meneses, A., Arriagada, S., Di Serio, A., & Hilliger, I. (2023). Technology-scaffolded peer assessment for developing critical thinking in pre-service teacher training: The importance of giving feedback. *Educational Technology Research and Development, 71*(2), 667–688. https://doi.org/10.1007/s11423-022-10173-1

Beniche, M. (2025). Hurdles Moroccan university professors encounter in enhancing critical thinking skills. *International Journal of Language and Literary Studies, 7*(3), 61–70. https://doi.org/10.36892/ijlls.v7i3.2100

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). McGraw-Hill Education.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*(2), 81–109. https://doi.org/10.1017/S0261444803001903

Brookfield, S. D. (2012). *Teaching for critical thinking: Tools and techniques to help students question their assumptions*. Jossey-Bass.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.

Chan, S., Maneewan, S., & Koul, R. (2021). Cooperative learning in teacher education: Its effects on EFL pre-service teachers' content knowledge and teaching self-efficacy. *Journal of Education for Teaching, 47*(5), 654–667. https://doi.org/10.1080/02607476.2021.1931060

Chan, S., Maneewan, S., & Koul, R. (2024). Cooperative learning in teacher education: A means to foster learning motivation and academic engagement among EFL pre-service teachers. *European Journal of Teacher Education, 47*(4), 658–675. https://doi.org/10.1080/02619768.2022.2117026

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). SAGE Publications.

Ennis, R. H. (1991). Critical thinking: A streamlined conception. *Teaching Philosophy, 14*(1), 5–24. https://doi.org/10.5840/teachphil19911412

Evans, C. M. (2020). *Measuring student success skills: A review of the literature on critical thinking*. National Center for the Improvement of Educational Assessment. https://files.eric.ed.gov/fulltext/ED607780.pdf

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. The Delphi Report.* American Philosophical Association.

Facione, P. A. (2015). *Critical thinking: What it is and why it counts* (2015 update). Insight Assessment. https://www.ndm.edu/sites/default/files/pdf/Facione-2015.pdf

Farrell, T. S. C. (2015). *Promoting teacher reflection in second language education: A framework for TESOL professionals*. Routledge.

Fernández-Río, J., Rivera-Pérez, S., & Iglesias, D. (2022). Cooperative learning interventions and associated outcomes in future teachers: A systematic review. *Revista de Psicodidáctica (English ed.), 27*(2), 118–131. https://doi.org/10.1016/j.psicoe.2022.04.001

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology, 38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2–18. https://doi.org/10.1037/a0024338

Gelerstein, D., del Río, R., Nussbaum, M., Chiuminatto, P., & López, X. (2016). Designing and implementing a test for measuring critical thinking in primary school. *Thinking Skills and Creativity, 20*, 40–49. https://doi.org/10.1016/j.tsc.2016.02.002

Gillies, R. M. (2014). Cooperative learning: Developments in research. International Journal of Educational Psychology, 3(2), 125–140. https://doi.org/10.4471/ijep.2014.08

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.

Huang, J., & Sang, G. (2023). Conceptualising critical thinking and its research in teacher education: A systematic review. Teachers and Teaching, 29(6), 638–660. https://doi.org/10.1080/13540602.2023.2212364

Isaacs, M. A., Tondeur, J., & Vaesen, J. (2024). Digital storytelling in teacher education: Developing pre-service teachers' critical thinking. *Australasian Journal of Educational Technology, 40*(3), 92–109. https://doi.org/10.14742/ajet.9258

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365–379. https://doi.org/10.3102/0013189X09339057

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review, 19*(1), 15–29. https://doi.org/10.1007/s10648-006-9038-8

Karacop, A., & Diken, E. H. (2017). The effects of the jigsaw technique based on cooperative learning on prospective science teachers' science process skills. *Journal of Education and Practice, 8*(6), 86–97.

Karacop, A., & Doymus, K. (2013). Effects of jigsaw cooperative learning and animation techniques on students' understanding of chemical bonding and their conceptions of the particulate nature of matter. *Journal of Science Education and Technology, 22*(2), 186–203. https://doi.org/10.1007/s10956-012-9385-9

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151–160. https://doi.org/10.1007/BF02288391

Lipman, M. (2003). *Thinking in education* (2nd ed.). Cambridge University Press.

Liu, O. L., Mao, L., Frankel, L., & Xu, J. (2016). Assessing critical thinking in higher education: The HEIghten™ approach and preliminary validity evidence. *Assessment & Evaluation in Higher Education, 41*(5), 677–694. https://doi.org/10.1080/02602938.2016.1168358

Liu, Y., Thurston, A., & Ye, X. (2024). Technology-enhanced cooperative language learning: A systematic review. *International Journal of Educational Research, 124*, Article 102314. https://doi.org/10.1016/j.ijer.2024.102314

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Lorencová, H., Jarošová, E., Avgitidou, S., & Dimitriadou, C. (2019). Critical thinking practices in teacher education programmes: A systematic review. *Studies in Higher Education, 44*(5), 844–859. https://doi.org/10.1080/03075079.2019.1586331

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics, 18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Mapeala, R., & Siew, N. M. (2015). The development and validation of a test of science critical thinking for fifth graders. *SpringerPlus, 4*, Article 741. https://doi.org/10.1186/s40064-015-1535-0

Marangio, K., Carpendale, J., Cooper, R., & Mansfield, J. (2024). Supporting the development of science pre-service teachers' creativity and critical thinking in secondary science initial teacher education. *Research in Science Education, 54*(1), 65–81. https://doi.org/10.1007/s11165-023-10104-x

Mrabti, A., Nfissi, A., & Madani Alaoui, K. (2023). University EFL teachers' perceptions of critical thinking barriers in higher education. *International Journal for Multidisciplinary Research, 5*(5). https://doi.org/10.36948/ijfmr.2023.v05i05.7508

Önal, İ. (2025). Building critical thinking in pre-service teachers: Dispositions, skills, and standards. *PLOS ONE, 20*(8), e0330536. https://doi.org/10.1371/journal.pone.0330536

Ouahani, N. E. & Hiba, B. (2023). Critical thinking practice in Moroccan higher education: An evaluation. *International Journal of Linguistics, Literature and Translation, 6*(1), 91–97. https://doi.org/10.32996/ijllt.2023.6.1.12

Paul, R. (1993). *Critical thinking: How to prepare students for a rapidly changing world*. Foundation for Critical Thinking.

Paul, R., & Elder, L. (2006). *Critical thinking: Tools for taking charge of your learning and your life*. Pearson Prentice Hall.

Paul, R., & Elder, L. (2014). *The miniature guide to critical thinking concepts and tools* (7th ed.). Foundation for Critical Thinking.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Qasserras, M., & Qasserras, L. (2023). *Moroccan instructors' perception of creative teaching practice*. International Journal for Multidisciplinary Research, 5(1), 1–9. https://doi.org/10.36948/ijfmr.2023.v05i01.2398

Sağsöz, Ö., Karataş, Ö. F., Türel, V., Yıldız, M., & Kaya, E. (2017). Effectiveness of jigsaw learning compared to lecture-based learning in dental education. *European Journal of Dental Education, 21*(1), 28–32. https://doi.org/10.1111/eje.12174

Salido, A., Syarif, I., Sitepu, M. S., Suparjan, Wana, P. R., Taufika, R., & Melisa, R. (2025). Integrating critical thinking and artificial intelligence in higher education: A bibliometric and systematic review of skills and strategies. *Social Sciences & Humanities Open, 12*, Article 101924. https://doi.org/10.1016/j.ssaho.2025.101924

Salvador-Garcia, C., & Sanahuja Ribés, A. (2025). Pre-service teachers' views of the value of cooperative learning: How can it shape 21st-century skills? *Journal of Teaching and Learning, 19*(5), 226–248. https://doi.org/10.22329/jtl.v19i2.8728

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.

Sekano, G. S., Laubscher, D., & Bailey, R. (2026). Growing independent thinkers: The role of technology-supported cooperative learning in fostering self-directed learning in mathematics. *Frontiers in Education, 11*, Article 1724270. https://doi.org/10.3389/feduc.2026.1724270

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Silva, H., Lopes, J., Morais, E., & Dominguez, C. (2023). Fostering critical and creative thinking through the cooperative learning jigsaw and group investigation. *International Journal of Instruction, 16*(3), 261–282. https://doi.org/10.29333/iji.2023.16315a

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.). Allyn & Bacon.

Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de Psicología, 30*(3), 785–791. https://doi.org/10.6018/analesps.30.3.201201

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford University Press.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Vives, E., Poletti, C., Robert, A., Butera, F., Huguet, P., ProFAN consortium, & Régner, I. (2024). Learning with jigsaw: A systematic review gathering all the pieces of the puzzle more than 40 years later. *Review of Educational Research*. Advance online publication. https://doi.org/10.3102/00346543241230064

Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology, 79*(1), 1–28. https://doi.org/10.1348/000709908X380772

Yin, X., Saad, M. R. B. M., & Halim, H. B. A. (2023). A systematic review of critical thinking instructional pedagogies in EFL writing: What do we know from a decade of research. *Thinking Skills and Creativity, 49*, Article 101363. https://doi.org/10.1016/j.tsc.2023.101363