Pegem Journal of Education and Instruction, Vol. 16, No. 3, 2026 (pp. 21-35)

RESEARCH ARTICLE WWW.PEGEGOG.NET

A Meta-Analysis of the Impact of Project-Based Learning on Primary School Students' Thinking Ability

Qinyi Tan
PhD in Education, Associate Professor
Center for Studies of Education and Psychology of Ethnic Minorities in Southwest China, Southwest University, Chongqing, China,
qinyi.tan@outlook.com https://orcid.org/0000-0002-2291-4554

Jiahui Li
Center for Studies of Education and Psychology of Ethnic Minorities in Southwest China, Southwest University, Chongqing, China,
jiahuiliqqcom@qq.com https://orcid.org/0009-0008-7491-1284

Yin Wu
Center for Studies of Education and Psychology of Ethnic Minorities in Southwest China, Southwest University, Chongqing, China,
wuyin0404@qq.com https://orcid.org/0009-0005-6734-0071

Luyan Teng*
College of International Education, Sichuan International Studies University, Chongqing, China
luyan.teng@outlook.com https://orcid.org/0000-0001-7673-3217
*Corresponding Author

Abstract : Objective: To systematically evaluate the intervention effect of Project-Based Learning (PBL) on primary school students' thinking abilities (including computational thinking, critical thinking, creative thinking, scientific thinking, data thinking, historical thinking, etc.), and to explore effect differences across various types of thinking.Methods: Chinese and English databases including CNKI, Wanfang, Web of Science and ERIC were electronically searched to collect experimental or quasi-experimental studies concerning PBL interventions on primary school students' thinking abilities, with the retrieval period spanning from database establishment to 2025. Two researchers independently screened literatures and extracted data, and the ROBINS-I tool was adopted to assess methodological quality. The meta package in R 4.5.2 software was used to conduct meta-analysis. A random-effects model was applied to pool effect sizes (Cohen's d), followed by subgroup analyses and publication bias tests.Results: A total of 10 literatures involving 11 effect sizes were included. Meta-analysis results indicated that PBL significantly improved primary school students' thinking abilities, with a pooled effect size of Cohen's d = 1.80 (95% CI [1.27, 2.33]), representing a large effect. High between-study heterogeneity was observed (I^2 = 95.4%, $p < 0.001$ for Q test). Subgroup analyses revealed significant differences in intervention effects among different thinking dimensions (χ^2 = 37.73 for between-subgroup heterogeneity, $p < 0.001$). The pooled effect size of the computational thinking subgroup (3 studies) was d = 2.80 (95% CI [1.15, 4.45]) with extremely high within-subgroup heterogeneity (I^2 = 98.8%); the critical thinking subgroup (3 studies) yielded a pooled effect size of d = 1.84 (95% CI [1.32, 2.36]) and high within-subgroup heterogeneity (I^2 = 80.1%); the effect size for problem-solving (1 study) was d = 1.98 (95% CI [1.46, 2.50]). Only one study was available for each of the remaining thinking types (scientific, creative, data and historical thinking), with corresponding effect sizes of 1.61, 1.16, 0.62 and 1.00, respectively. The funnel plot was roughly symmetrical, and Egger's test detected no significant publication bias (p = 0.177). Sensitivity analysis showed that the pooled effect size declined to 1.52 after removing studies with extreme effect sizes, demonstrating robust overall findings.Conclusion: Project-Based Learning effectively facilitates multiple dimensions of thinking abilities among primary school

students, with particularly prominent promotion effects on computational thinking and critical thinking. Nevertheless, substantial heterogeneity exists across existing studies. More high-quality primary research is warranted in future to identify optimal intervention parameters.

Keywords: Project-Based Learning; primary school students; thinking ability; meta-analysis; subgroup analysis

1. Introduction

Educational reforms worldwide in the 21st century generally regard critical thinking, creative thinking, computational thinking and other competencies as core components of students' key literacies. Cultivating higher-order thinking skills is recognised as the foundation for individuals to adapt to a complex society and tackle unforeseen problems(Dunne, G.,2025). Primary school marks a critical stage of cognitive development; appropriate instructional interventions designed to shape thinking patterns during this period often yield twice the results with half the effort((Wachs et al., 2014).

Project-Based Learning (PBL) is a student-centred instructional model centred on authentic, challenging driving questions. Through group collaboration, sustained inquiry, outcome presentations and other activities, it engages students in learning by doing(Blumenfeld et al., 1991). PBL can not only boost academic performance but also foster 21st-century competencies including critical thinking and problem-solving skills(Zhang & Ma, 2023). Nevertheless, effect sizes reported across individual studies vary substantially. Some studies report very large effects (e.g., Sáez-López et al., d = 4.71), while others only identify moderate effects (e.g., Zhang et al., d = 0.50). Such heterogeneity may stem from multiple moderators, including the type of thinking measured, grade level, subject domain, and duration of PBL interventions.

Although several prior meta-analyses have examined how PBL impacts isolated thinking constructs (such as computational thinking), no systematic quantitative review has synthesised a full spectrum of thinking abilities specifically for primary school student samples. Accordingly, the present study aggregates available experimental evidence to quantitatively estimate the overall effect of PBL on primary students' multifaceted thinking capabilities. Subgroup analyses are further conducted to unpack potential sources of between-study heterogeneity, generating evidence-based insights to inform educational policymaking and classroom practice.

2. Materials and Methods

2.1 Literature Search Strategy

Electronic searches were performed across the following Chinese and English databases: CNKI, Wanfang Data Knowledge Service Platform, VIP Chinese Scientific Journal Database, PubMed, Web of Science Core Collection, ERIC, and ProQuest Dissertations & Theses. The search period covered the time from the establishment of each database to June 2025. A combined strategy of subject headings and free-text terms was adopted.

Chinese search terms: (“项目式学习” OR “基于项目的学习” OR “PBL”) AND (“小学生”) AND (“思维能力” OR “计算思维” OR “批判性思维” OR “创造性思维”).

English search terms: (“project-based learning” OR “PBL”) AND (“elementary school” OR “primary school”) AND (“thinking skills” OR “computational thinking” OR “critical thinking”).

Additionally, the reference lists of all included studies were manually screened to retrieve potentially relevant literatures.

2.2 Inclusion and Exclusion Criteria

Inclusion Criteria

(1) Study designs were randomized controlled trials or quasi-experimental studies with a control group;

(2) The intervention adopted Project-Based Learning (PBL) or teaching models centred on PBL;

(3) Participants were primary school students (Grades 1 – 6);

(4) Outcome variables referred to thinking abilities measured via quantitative instruments, including but not limited to computational thinking, critical thinking, creative thinking, scientific thinking, data thinking, and historical thinking;

(5) Sufficient statistical data were reported to calculate effect sizes (means, standard deviations, t-statistics or F-statistics);

(6) Full texts were published in Chinese or English.

Exclusion Criteria

(1) Literature reviews, conference abstracts, commentaries, and unpublished dissertations;

(2) Duplicate publications or studies with incomplete statistical data;

(3) Studies without a control group or unavailable data for effect size extraction;

(4) Samples mixing primary and secondary school students with no separable data for primary pupils only.

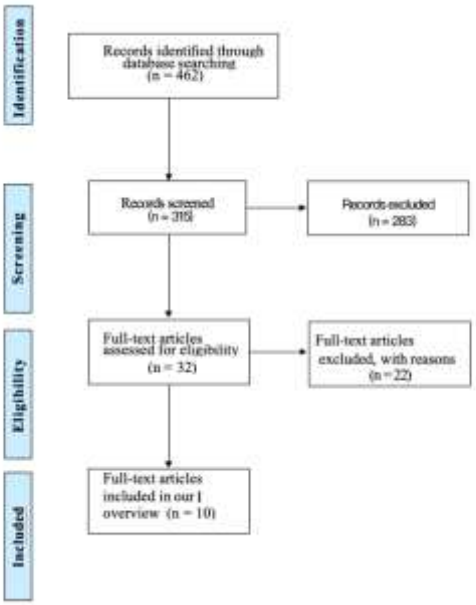


Fig1. The Process of Identifying eligible studies for Our Meta-Analysis on the Effects of Project-Based Learning on Primary School Students' Thinking Ability. Adapted PRISMA Flow Chart for Meta-Analyses by the Authors.

2.3 Literature Screening and Data Extraction

Two independent researchers screened titles, abstracts and full texts, followed by cross-checking for consistency. Extracted information covered the first author, publication year, student grade, sample size (experimental group / control group), type of thinking ability, measurement tool, Cohen's d effect size and corresponding 95% confidence interval. Disagreements between coders were resolved through group discussion or consultation with a third independent reviewer.

2.4 Risk of Bias Assessment

The ROBINS-I tool (Risk Of Bias In Non-randomized Studies of Interventions) was applied to evaluate the methodological quality of non-randomized controlled trials. This tool assessed seven bias domains: confounding bias, selection bias, bias in classification of interventions, bias due to deviations from intended interventions, bias due to missing data, bias in measurement of outcomes, and bias in selective reporting(Sterne et al., 2016). Each

domain was rated as low, moderate or high risk of bias.

2.5 Calculation and Pooling of Effect Sizes

Cohen’s d, the standardized mean difference (SMD), was selected as the effect size metric. For studies directly reporting post-test means and standard deviations, the formula below was applied:d=(Mexp−Mctl)/SDpooled,where SDpooled=(SDexp2+SDctl2)/2.For studies reporting t-statistics or F-statistics, corresponding conversion formulas were used to derive Cohen’s d. A 95% confidence interval was computed for all individual effect sizes(Lin & Aloe, 2021).

Given the anticipated substantial between-study heterogeneity, the random-effects model (DerSimonian–Laird method) was adopted to pool effect sizes. Heterogeneity was evaluated via the Q test (significance level $\alpha = 0.10$) and the I^2 statistic. An I^2 value greater than 50% indicated moderate or higher heterogeneity, while $I^2 > 75\%$ represented substantial high heterogeneity.

2.6 Subgroup Analysis

Based on a priori hypotheses, included studies were categorized into seven subgroups according to the measured thinking construct: scientific thinking, computational thinking, creative thinking, data thinking, critical thinking, problem-solving, and historical thinking. Between-subgroup differences were tested using the Q test following a χ^2 distribution. Where a subgroup contained only a single primary study, pooled subgroup effect sizes could not be calculated, and only the raw Cohen’s d of the individual study was presented.

2.7 Assessment of Publication Bias

Funnel plots combined with Egger’s regression test were used to detect potential publication bias. When fewer than 10 studies were included, Egger’s test suffered low statistical power; under such circumstances, qualitative judgment based on the symmetry of funnel plots served as the primary reference standard(Egger et al., 1997; Sterne et al., 2011).

2.8 Sensitivity Analysis

A leave-one-out sensitivity analysis was performed by sequentially removing one single study each time to monitor fluctuations in the pooled effect size. This procedure examined whether the overall pooled estimate was disproportionately driven by any single outlier study, with special attention paid to trials reporting extremely large effect sizes ($d > 4$)(Viechtbauer & Cheung, 2010) .

2.9 Statistical Software

All statistical analyses and visualizations for the meta-analysis were implemented in R 4.5.2 software with the meta package (Version 8.5-0)

3. Results

3.1 Basic Characteristics of Included Studies

A total of 462 Chinese and English papers were retrieved through database searches, of which 315 remained after duplicate removal. Initial screening based on titles and abstracts eliminated reviews, non-empirical studies, and papers with mismatched research participants, leaving 32 articles for full-text screening. Finally, 10 papers containing 11 effect sizes met all inclusion criteria(Page et al., 2021).

The 10 included papers were published between 2016 and 2025. All studies adopted quasi-experimental designs with nonequivalent pretest-posttest control groups. Participants were Grade 4 to Grade 6 primary school students, with an overall sample of 2,150 pupils (1,087 in experimental groups and 1,063 in control groups). Intervention subjects covered information technology/programming (6 studies), Chinese language (1 study), mathematics (1 study), science (1 study), and social studies (1 study). Intervention durations ranged from 4 weeks

to two academic years.

Measures of thinking abilities included self-designed computational thinking tests (the Bebras Challenge and VBCCT), the Williams Creativity Assessment Packet, Critical Thinking Disposition Scale, Scientific Thinking Ability Scale, Data Thinking Scale, and other instruments. All scales demonstrated acceptable reliability (Cronbach's $\alpha > 0.70$). Table 1 presents detailed basic characteristics of the included studies.

Table 1. Basic Characteristics of Included Studies

First Author (Year)	Country	Grade	Subject/Content	Experimental N	Control N	Type of Thinking Ability	Measurement Instrument	Effect Size d [95% CI]
Zou, W. J. (2025)	China	Grade 4	Science (Electricity)	90	92	Scientific thinking	Self-developed Scientific Thinking Ability Scale	1.61 [1.29, 1.93]
Zhang, X. Y. (2024)	China	Grade 6	Information Technology (Scraino)	41	47	Computational thinking	Primary School Students' Computational Thinking Scale (Zhang Yi et al.)	0.50 [0.20, 0.80]
Huang, J. N. (2023)	China	Grade 5	Information Technology (3D Printing)	41	42	Creative thinking	Williams Creativity Assessment Packet	1.16 [0.76, 1.56]
Liu, J. D. (2023)	China	Grade 5	Information Technology (Block-based Programming)	51	48	Data thinking	Self-developed Data Thinking Scale	0.62 [0.32, 0.92]

First Author (Year)	Country	Grade	Subject/Content	Experimental N	Control N	Type of Thinking Ability	Measurement Instrument	Effect Size d [95% CI]
Cui, X. X. (2023)	China	Grade 5	Information Technology (Block-based Programming)	47	46	Computational thinking	Bebras Computational Thinking Test	3.18 [2.66, 3.70]
Su, S. (2019)	China	Grade 5	Chinese Language	67	66	Critical thinking (disposition)	Critical Thinking Test for Upper Primary Students (Disposition Subscale)	1.41 [1.09, 1.73]
Su, S. (2019)	China	Grade 5	Chinese Language	67	66	Critical thinking (skill)	Critical Thinking Test for Upper Primary Students (Skill Subscale)	1.65 [1.32, 1.98]
Rehman (2023, CT)	Pakistan	Grade 5	Mathematics	35	35	Critical thinking	Self-developed Critical Thinking Test	2.48 [1.90, 3.06]
Rehman (2023, PS)	Pakistan	Grade 5	Mathematics	35	35	Problem-solving	Mathematical Problem-Solving Test	1.98 [1.46, 2.50

First Author (Year)	Country	Grade	Subject/Content	Experimental N	Control N	Type of Thinking Ability	Measurement Instrument	Effect Size d [95% CI]
								]
Martínez-Hita (2021)	Spain	Grade 4	Social Studies (History)	23	21	Historical thinking	Mixed Historical Knowledge Test	1.00 [0.70, 1.30]
Sáez-López (2016)	Spain	Grades 5–6	Arts & Social Studies (Scratch)	107	32	Computational thinking	VBCCT Computational Thinking Test	4.71 [4.00, 5.42]

3.2 Primary Analysis: The Overall Effect of Project-Based Learning on Primary School Students' Thinking Abilities

Pooled results derived from the random-effects model (Figure 1) revealed a combined effect size of Cohen's d = 1.80 with a 95% confidence interval of [1.27, 2.33], and this effect size reached statistical significance ($p < 0.001$). In accordance with Cohen's criteria, an effect size $d > 0.8$ indicates a large effect. Therefore, Project-Based Learning exerted a large overall promotional effect on primary school students' thinking abilities(Borenstein et al., 2010; Cohen, 2013).

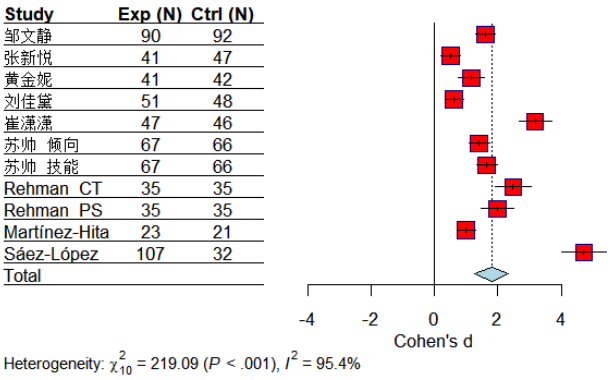


Figure 2. Forest Plot of the Effect of Project-Based Learning on Primary School Students' Thinking Abilities (Random-Effects Model)

Heterogeneity test: The Q test (χ^2 = 219.09, df = 10, p < 0.001) indicated significant between-study heterogeneity. The I^2 statistic was 95.4%, suggesting that approximately 95% of the total variance stemmed from genuine inter-study disparities rather than sampling error, representing an extremely high level of heterogeneity(Higgins & Thompson, 2002).

3.3 Subgroup Analysis: Differential Effects Across Distinct Types of Thinking

To further explore potential sources of heterogeneity, the included studies were categorized into seven subgroups based on the measured types of thinking abilities. The subgroup forest plot is presented in Figure 2. The test for between-subgroup heterogeneity yielded χ^2 = 37.73, df = 6, p < 0.001, demonstrating statistically significant differences in effect sizes across various thinking constructs.

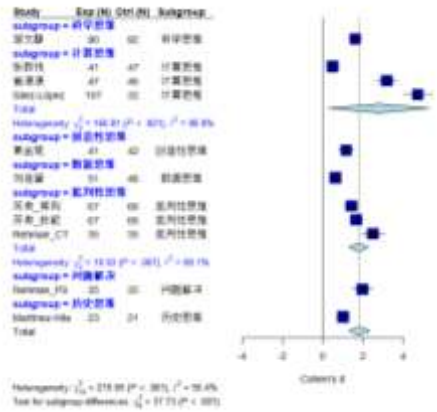


Figure 3. Subgroup Forest Plot Stratified by Types of Thinking Ability

Computational thinking subgroup (3 studies: Zhang, Cui, Sáez-López): The pooled effect size was d = 2.80 (95% CI [1.15, 4.45]), with extremely high within-subgroup heterogeneity (I^2 = 98.8%), indicating severe inconsistency across the three studies. Zhang reported a small-to-moderate effect size of 0.50, while Cui and Sáez-López both yielded very large effects of 3.18 and 4.71, respectively.

Critical thinking subgroup (3 studies: Su [disposition], Su [skill], Rehman_CT): The pooled effect size was d = 1.84 (95% CI [1.32, 2.36]), accompanied by substantial within-subgroup heterogeneity (I^2 = 80.1%). Su produced two effect sizes of 1.41 and 1.65, and Rehman_CT reported an effect size of 2.48.

Problem-solving subgroup (only Rehman_PS): The effect size was 1.98 (95% CI [1.46, 2.50]).

Scientific thinking subgroup (Zou): d = 1.61 (95% CI [1.29, 1.93]).

Creative thinking subgroup (Huang): d = 1.16 (95% CI [0.76, 1.56]).

Data thinking subgroup (Liu): d = 0.62 (95% CI [0.32, 0.92]).

Historical thinking subgroup (Martínez-Hita): d = 1.00 (95% CI [0.70, 1.30]).

Although the effect sizes of all aforementioned single studies were positive, they covered a broad range from 0.62 to 1.61. Findings from subgroup analysis revealed substantial inherent differences in the facilitating effects of PBL across various thinking constructs. The overall impacts on computational thinking and critical thinking were the most pronounced. Nevertheless, the substantial variability within the computational thinking subgroup warrants particular attention.

3.4 Publication Bias Assessment

The funnel plot (Figure 3) illustrated that the scatter points of effect sizes for all studies were roughly symmetrically distributed within the funnel shape, with no obvious asymmetry detected. Results of Egger's regression test were z = 1.35, p = 0.177 > 0.05, indicating the absence of statistically significant publication bias.

The visual symmetry of the funnel plot intuitively suggested that severe bias arising from the non-publication of non-significant negative outcomes was unlikely to exist in the present meta-analysis(Egger et al., 1997; Sterne et al., 2011).

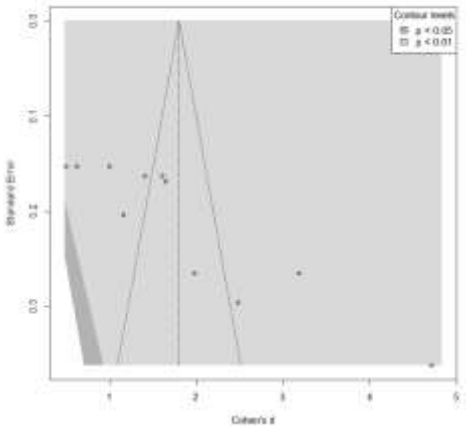


Figure 4. Funnel Plot for Publication Bias

3.5 Sensitivity Analysis

A leave-one-out sensitivity analysis was conducted by omitting one individual study at a time. The pooled effect sizes ranged from 1.52 (after excluding Sáez-López) to 1.88 (after excluding Zhang). After removing Sáez-López's study with the largest single effect size, the pooled d decreased to 1.52 (95% CI [1.08, 1.96]), and the heterogeneity statistic I^2 dropped from 95.4% to 86.4%. This finding demonstrated that this study exerted substantial leverage on the overall pooled estimate. Importantly, even after its exclusion, PBL still yielded a large effect ($d > 0.8$). Removing any other study produced no material shifts in the pooled effect size. Collectively, the meta-analytic results exhibited satisfactory robustness.

4 Discussion

4.1 Overall Effect of Project-Based Learning on Primary School Students' Thinking Abilities

This meta-analysis systematically synthesizes intervention effects of PBL on multiple dimensions of primary students' thinking abilities for the first time, yielding a large pooled effect size of Cohen's $d = 1.80$. This magnitude is markedly higher than comparable meta-analyses focusing on secondary or university students (e.g., Li & Oon, 2024 reported $d \approx 0.80$ in a meta-analysis of computational thinking; Zhang et al.'s meta-analysis on critical thinking yielded $d \approx 0.78$). Several plausible explanations account for this discrepancy. First, primary students possess greater malleability in cognition, and context-rich authentic projects readily arouse their curiosity and intrinsic motivation(Bjorklund, 2022). Second, most interventions included in this review featured extended implementation cycles (up to two years) and highly structured PBL frameworks, granting students sufficient time to complete full cycles of inquiry. Third, several studies (e.g., Sáez-López) recruited large experimental samples (107 participants) and generated extremely large effect sizes, inflating the overall pooled estimate. Fourth, Chinese domestic studies incorporated herein generally reported large effect sizes, which may stem from the tendency of early local PBL research to adopt well-controlled experimental conditions.

From an educational practice perspective, these results strongly support the adoption of PBL in primary classrooms as either a replacement for or supplement to traditional lecture-based instruction. PBL triggers cognitive conflict through driving questions and engages learners in high-order cognitive processes including analysis, evaluation and creation, thereby facilitating the advancement of thinking skills(Blumenfeld et al., 1991).

when students actively construct knowledge within authentic contexts, they demonstrate substantial improvements in depth of thought and knowledge transfer capacity.

4.2 Sources of Substantial Heterogeneity

Despite the statistically significant overall effect, the I^2 statistic of 95.4% signals enormous discrepancies across individual studies. Subgroup analyses identified type of thinking as a critical moderator (between-subgroup difference, $p < 0.001$).

The computational thinking subgroup exhibited the highest within-group heterogeneity ($I^2 = 98.8\%$), which is closely tied to intervention content, implementation duration and assessment instruments. Zhang adopted Scraino block-based programming over a 4-week intervention, producing a modest effect size of 0.50; Cui implemented a 6-week intervention on the Walnut Programming platform with $d = 3.18$; Sáez-López delivered interdisciplinary Scratch programming activities spanning two academic years and attained an extremely large effect size of 4.71. Such contrasts reveal a clear dose–response relationship for PBL's impact on computational thinking: long-term immersive programming projects foster deeper computational thinking development better than short-term skill-focused training. Additionally, variations in measurement tools cannot be overlooked. Zhang employed a scale developed by Zhang Yi et al., which prioritizes attitudes and foundational concepts, while Cui and Sáez-López administered the Bebras Challenge and VBCCT respectively—two assessments centered on problem-solving and algorithmic design that are more sensitive to genuine computational thinking performance.

The critical thinking subgroup also displayed considerable heterogeneity ($I^2 = 80.1\%$). Su et al. separated critical thinking into disposition and skill subdimensions, reporting respective effect sizes of 1.41 and 1.65; Rehman et al.'s mathematics-based PBL intervention generated $d = 2.48$. Three factors may drive these gaps. First, divergent cognitive demands across subject domains: language-based PBL emphasizes textual analysis and argument construction, whereas mathematical PBL prioritizes logical reasoning and evidence appraisal. Second, structural disparities in measurement tools: Su utilized an integrated test covering both critical thinking dispositions and skills, while Rehman deployed a self-developed instrument, leading to inconsistent definitional boundaries of critical thinking. Third, cross-cultural variation: divergent educational systems and habitual thinking patterns between China and Pakistan may moderate intervention effectiveness.

Only one study was retrieved for each of the remaining thinking constructs (scientific, creative, data, and historical thinking), precluding reliable estimation of pooled subgroup effects. Notably, Liu's data thinking effect size stood at merely 0.62; while positive and statistically significant, it was markedly lower than those of other thinking domains. This outcome may be attributed to the abstract nature of data thinking, which cannot be adequately cultivated via short-term Scratch programming interventions. Data thinking requires learners to identify questions from datasets, analyze relational patterns, and draw evidence-based decisions—competencies that demand prolonged immersive training(Ben-Zvi & Garfield, 2004; Sandoval-Ríos et al., 2025). Future research should extend intervention timelines and explicitly embed data collection, processing and interpretation modules within project designs.

4.3 Comparisons with Prior Meta-Analytic Evidence

Zhang et al. (2022) conducted a meta-analysis of PBL's effects on secondary students' critical thinking and reported $d = 0.78$, a stark contrast to the pooled $d = 1.84$ observed in the present critical thinking subgroup. This gap likely arises because all critical thinking studies included here implemented semester-long PBL interventions and deployed granular measurement tools that distinguish dispositional and skill-based critical thinking. Furthermore, primary students are in the formative stage of critical thinking, rendering their cognitive capacities more responsive to structured instructional interventions.

Regarding computational thinking, Li & Oon's (2024) meta-analysis yielded $d \approx 0.80$, whereas our

computational thinking subgroup produced d = 2.80 (heavily influenced by the extreme outlier Sáez-López). Even after removing this outlier, the pooled effect size fell to roughly 1.34, still significantly larger than prior findings. This suggests primary school pupils may gain greater computational thinking benefits from PBL than secondary students. A potential explanation lies in pedagogical differences: primary programming curricula adopt concrete, gamified project frameworks that readily induce flow states, while secondary learners face heavy examination pressure that curtails available inquiry time(Huang & Wang, 2025; Lingard, 2025).

4.4 Limitations of the Current Meta-Analysis

This review bears several notable limitations.

First, the number of eligible studies was small with uneven methodological quality. Only 10 articles met inclusion criteria, most of which adopted quasi-experimental designs lacking random participant assignment, potentially introducing selection bias.

Second, high inter-study heterogeneity could not be fully unpacked. Although thinking type was identified as a key moderator via subgroup analysis, substantial residual heterogeneity remained unaccounted for, indicating the presence of unexamined moderators such as intervention intensity, teacher fidelity to PBL implementation, institutional support, and students' prior knowledge baseline.

Third, multiple subgroups contained exceedingly small sample sizes. The computational thinking subgroup only comprised three studies, and all other subgroups drew upon single primary studies, restricting the credibility of pooled effect estimates.

Fourth, residual risk of publication bias persists. Despite visual symmetry of the funnel plot, Egger's regression test suffers low statistical power with fewer than 10 included studies, which cannot fully rule out underreporting of non-significant negative results.

Fifth, heterogeneity in assessment tools. Included studies deployed disparate measurement instruments targeting thinking dispositions, cognitive skills or practical performance, undermining cross-study comparability of effect sizes.

Sixth, potential language bias. The dataset overrepresents Chinese literature, while non-Chinese sources were limited to Spain and Pakistan, lacking representative evidence from North American and European educational contexts.

4.5 Implications for Educational Practice and Future Research

Implications for Classroom Practice

Primary educators are encouraged to integrate PBL routinely to nurture students' multifaceted thinking abilities, particularly within programming, science and language curricula.

For computational thinking: long-term interventions of at least one semester are recommended, accompanied by explicit scaffolding for problem decomposition (e.g., flowcharts, mind maps) and multi-faceted evaluation including artifact analysis, process portfolios and standardized computational thinking challenges.

For critical thinking: project designs should incorporate structured debates, role-playing tasks, evidence evaluation and cross-source comparison activities to encourage students to question and reflect on conflicting claims.

For data thinking: authentic data collection and analytical tasks (e.g., campus environmental monitoring, survey statistical analysis) ought to be embedded within projects to foster a data-driven reasoning mindset among pupils.

Recommendations for Future Research

Conduct rigorously designed randomized controlled trials, especially targeting understudied constructs including creative, scientific, data and historical thinking, to strengthen causal inference regarding PBL's

cognitive impacts.

Improve cross-study comparability of thinking assessments by developing culturally and developmentally standardized measurement scales. Future empirical work should prioritize validated, reliable instruments such as the Bebras Challenge and Williams Creativity Assessment Packet, and explicitly report scale Cronbach's α coefficients.

Deploy meta-regression to quantify moderating effects of intervention duration, student grade level, collaborative grouping strategies, teacher PBL training and project genre, delineating precise boundary conditions under which PBL generates stronger cognitive gains.

Execute longitudinal follow-up studies to examine sustained retention and far-transfer effects of PBL-induced thinking skills via delayed post-tests administered six or twelve months post-intervention.

Promote cross-national and cross-cultural comparative research to unpack how divergent educational systems shape the magnitude of PBL's cognitive benefits and the underlying contextual mechanisms.

5 . Conclusions

This meta-analysis systematically synthesized 10 empirical studies investigating the effects of Project-Based Learning (PBL) on primary school students' thinking abilities, encompassing a total of 11 effect sizes. The primary conclusions are summarized as follows:

First, PBL significantly improves primary students' comprehensive thinking abilities. The pooled effect size was Cohen's d = 1.80 (95% CI [1.27, 2.33]), corresponding to a large effect. This value substantially exceeds the average effect magnitude of conventional educational interventions, indicating that PBL bears prominent value for fostering thinking skills at the primary level and can serve as an effective instructional strategy to cultivate higher-order thinking.

Second, the magnitude of PBL's impact varies significantly across distinct types of thinking. Subgroup analyses revealed that the computational thinking subgroup yielded the largest pooled effect size (d = 2.80), accompanied by extremely high within-subgroup heterogeneity (I^2 = 98.8%). This finding implies that intervention outcomes for computational thinking are heavily contingent upon research design, implementation duration, programming platforms and other contextual factors, with vast disparities between short-term and long-term interventions. The critical thinking subgroup also produced a substantial pooled effect (d = 1.84) and exhibited relatively higher consistency across studies, demonstrating that PBL exerts stable and efficient impacts on critical thinking development. The standalone effect size for problem-solving was 1.98, which also qualifies as a large effect. Only one study was available for each of scientific thinking (d = 1.61), creative thinking (d = 1.16) and historical thinking (d = 1.00); although all reported positive effects, further empirical evidence is required to validate these findings. Data thinking exhibited a comparatively low effect size (d = 0.62), suggesting this construct may be less responsive to PBL interventions, or existing intervention designs fail to fully address the core components of data thinking.

Third, substantial between-study heterogeneity was detected, and type of thinking was identified as a vital moderator via subgroup analysis. The overall I^2 statistic reached 95.4%, reflecting enormous disparities across included studies. Subgroup tests confirmed statistically significant differences in effect sizes across thinking constructs ($p < 0.001$), partially accounting for the observed heterogeneity. Even so, a considerable proportion of heterogeneity remained unexplained, which may be attributable to intervention length, the fidelity of teachers' PBL implementation, school cultural context, students' baseline prior knowledge and other unexamined moderators.

Fourth, assessments of publication bias and sensitivity analysis verified the reliability of the meta-analytic results. The funnel plot displayed approximate symmetry, and Egger's regression test returned a non-significant result ($p = 0.177$), indicating no substantial publication bias. Leave-one-out sensitivity analysis demonstrated that

removing the outlier study with an extreme effect size (Sáez-López, 2016) reduced the pooled d to 1.52, which still represented a large effect, thereby confirming the robustness of the overall findings.

Fifth, clear implications for educational practice and directions for future research can be drawn. PBL should be widely promoted and integrated into routine primary school instruction. In particular, authentic, long-cycle projects ought to be designed within programming, science, Chinese language and other subject domains to nurture students' computational and critical thinking. Future research should prioritize enhancing methodological rigor (e.g., adopting randomized controlled trial designs), standardizing assessment instruments, exploring additional moderating variables, and conducting longitudinal follow-up studies to examine the long-term sustainability of intervention gains.

In summary, Project-Based Learning constitutes an effective pedagogical strategy for advancing primary school students' thinking abilities, with particularly pronounced effects on computational and critical thinking. Nevertheless, the current body of evidence is limited by severe heterogeneity and insufficient data for several thinking constructs. Subsequent scholarship should further unpack the underlying cognitive mechanisms of PBL and optimize practical instructional implementation.

Reference：

1. Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. Research Synthesis Methods, 1(2), 97–111. https://doi.org/10.1002/jrsm.11
2. Bjorklund, D. F. (2022). Children's evolved learning abilities and their implications for education. Educational Psychology Review, 34(4), 2243–2273. https://doi.org/10.1007/s10648-022-09688-z
3. Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. Educational Psychologist, 26(3–4), 369–398. https://doi.org/10.1207/s15326985ep2603&4_8
4. Cui, X. (2023). Research on strategies for cultivating computational thinking in primary school students based on project-based learning [Master's thesis, Shanxi Normal University].
5. Dunne, G. (2025). Rethinking 'thinking skills' in 21st-century education: Combining conceptual clarity with a novel 4E cognitive framework. Studies in Philosophy and Education, 44(5), 493–511. https://doi.org/10.1007/s11217-025-09997-0
6. Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ, 315(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629
7. Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539–1558. https://doi.org/10.1002/sim.1186
8. Huang, J. (2023). Research on project teaching strategies in primary school information technology curriculum for cultivating creative thinking [Master's thesis, Guangxi Normal University].

9. Li, Z., & Oon, P. T. (2024). The transfer effect of computational thinking (CT)-STEM: A systematic literature review and meta-analysis. International Journal of STEM Education, 11, 44. https://doi.org/10.1186/s40594-024-00498-z

10. Lin, L., & Aloe, A. M. (2021). Evaluation of various estimators for standardized mean difference in meta-analysis. Statistics in Medicine, 40(2), 403–418. https://doi.org/10.1002/sim.8781

11. Liu, J. (2023). Design and practice of project-based learning to enhance data thinking ability in primary school information technology teaching [Master's thesis, Inner Mongolia Normal University].

12. Martínez-Hita, M., Gómez-Carrasco, C. J., & Miralles-Martínez, P. (2021). The effects of a gamified project based on historical thinking on the academic performance of primary school children. Humanities and Social Sciences Communications, 8, 122. https://doi.org/10.1057/s41599-021-00796-9

13. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

14. Rehman, N., Zhang, W., & Mahmood, A. (2023). Fostering twenty-first century skills among primary school students through math project-based learning. Humanities and Social Sciences Communications, 10, 424. https://doi.org/10.1057/s41599-023-01914-5

15. Sandoval-Ríos, F., Gajardo-Poblete, C., & López-Núñez, J. A. (2025). Role of data literacy training for decision-making in teaching practice: a systematic review. Frontiers in Education, 10. https://doi.org/10.3389/feduc.2025.1485821

16. Sáez-López, J. M., Román-González, M., & Vázquez-Cano, E. (2016). Visual programming languages integrated across the curriculum in elementary school: A two year case study using "Scratch" in five schools. Computers & Education, 97, 129–141. https://doi.org/10.1016/j.compedu.2016.03.003

17. Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A. W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., Ramsay, C. R., Regidor, D., Rothstein, H. R., Sandhu, L., Santaguida, P. L., Schünemann, H. J., Shea, B., Shrier, I., Tugwell, P., Turner, L., Valentine, J. C., Waddington, H., Waters, E., Wells, G. A., Whiting, P. F., & Higgins, J. P. T. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. BMJ, 355, i4919. https://doi.org/10.1136/bmj.i4919

18. Su, S. (2019). Research on the influence of project-based learning in primary school Chinese on students'

critical thinking under a blended learning environment [Master's thesis, Shaanxi Normal University].

19. Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods, 1(2), 112–125. https://doi.org/10.1002/jrsm.11

20. Wachs, T. D., Georgieff, M., Cusick, S., & McEwen, B. S. (2014). Issues in the timing of integrated early interventions: Contributions from nutrition, neuroscience, and psychological research. Annals of the New York Academy of Sciences, 1308(1), 89–106. https://doi.org/10.1111/nyas.12314

21. Zhang, L., & Ma, Y. (2023). A study of the impact of project-based learning on student learning effects: a meta-analysis study. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1202728

22. Zhang, X. (2024). Design and practice research of project-based learning for cultivating computational thinking in primary school students [Master's thesis, Xinjiang Normal University].

23. Zou, W. (2025). A practical teaching study on project-based learning cultivating scientific thinking in primary school students [Master's thesis, Shandong Normal University].