OPEN ACCESS
ISSN: 2835-6764
Piao H and Hou J. 2026; 5(6) 668-678
DOI: 10.71010/2835-6764/ajser-e329
Research Article
American Journal of Science Education Research

Advancing Toward Higher-Order Thinking: Empirical Evidence from Instructional Interventions in Undergraduate Critical Thinking

Huiyan Piao[1*], Jia Hou[2]

[1]College of Educational Science, Shenyang Normal University
[2]A postgraduate student from the College of Educational Science at Shenyang Normal University

***Corresponding author:** Huiyan Piao

Citation: Piao H and Hou J (2026) Advancing Toward Higher-Order Thinking: Empirical Evidence from Instructional Interventions in Undergraduate Critical Thinking. American J Sci Edu Re: AJSER-329.

Received Date: 06 May, 2026; **Accepted Date:** 24 May, 2026; **Published Date:** 01 June, 2026

Abstract
Against the backdrop of the knowledge economy and globalization, fostering undergraduate students' critical thinking has become an urgent priority in the higher education sector. Based on a rigorous screening and meta-analysis of 59 international experimental or quasi-experimental studies on instructional interventions for undergraduate critical thinking over the past ten years, this study aims to examine how different instructional types, models, and strategies impact critical thinking. The results indicate that, overall, all types of instructional interventions significantly promote undergraduates' critical thinking. With regard to instructional types, discipline-independent courses exhibit the most salient effects. In terms of instructional models, the behavioral model is more conducive to facilitating students' cognitive development. Regarding instructional strategies, the integrated application of "interaction+experience" generates the most substantial outcomes. By contrast, the moderation effects of variables such as intervention duration and instructional tools are relatively limited; however, medium-term interventions (17-32 weeks) and the use of appropriate digital resources correlate with modestly greater improvements in critical thinking. Based on the findings, the following recommendations are proposed: Firstly, Higher education institutions should offer specialized critical thinking courses or explicitly embed critical thinking objectives within disciplinary curricula; Secondly, Priority should be given to instructional models that emphasize observable behaviors and self-reflection, alongside the effective integration of interaction and experiential strategies; Thirdly, Diversified assessment and feedback systems should be improved to continuously monitor and guide students' critical thinking development.

1. Introduction

1.1 Background

Driven jointly by the knowledge economy and the wave of globalization, higher education has increasingly been entrusted with the strategic mission of fostering higher-order thinking and developing citizens' core competencies. Higher-order thinking is characterized by multi-dimensionality, complexity, and creativity [1] and represents a crucial capacity of individuals to engage in in-depth inquiry, integrative judgment, and innovative practice when confronting uncertain situations. As a core component of higher-order thinking, critical thinking is accorded a central position in both international education reform agendas and domestic talent development systems.

International organizations such as United Nations Educational, Scientific and Cultural Organization (UNESCO) [2] and Organization for Economic Co-operation and Development (OECD) [3] have successively advocated the cultivation of critical thinking and creativity. Furthermore, the Framework for 21st Century Skills in the United States has fully integrated the "4C skills" (critical thinking, communication, collaboration, creativity) and ICT literacy into the educational process, aiming to foster future-oriented interdisciplinary talents [4]. In China, in pursuit of building a strong nation in science and technology and cultivating top-notch innovative talents, universities have successively introduced specific policies that identify critical thinking, innovative thinking, and related capacities as core demands in developing high-level talent cultivation systems [5]. Against this backdrop, the question of how to effectively cultivate critical thinking in higher education has become a critical issue for both researchers and practitioners.

1.2 Conceptualization of Critical Thinking

The concept of critical thinking has been interpreted in diverse ways across disciplinary contexts. The path of philosophical speculation centers on the rational standards of human reason, the spirit of skepticism, and the rationality of cognitive processes, and emphasizes the thinking chain of doubting, analyzing and evaluating [6]. Cognitive psychology, by contrast, places greater emphasis on the internal cognitive structure and metacognitive monitoring function of critical thinking, viewing it as reliant on cognitive skills such as logical reasoning, inductive and deductive thinking, and reflective judgment in processes of information selection and judgment [7]. Educational researchers, meanwhile, mostly adopt disciplinary content or instructional processes as the carrier to provide a functional definition of critical thinking [8], and emphasize how to stimulate students' abilities of critical thinking, reverse thinking, and reflective inquiry in teaching practice [9].

The six-dimensional framework of "interpretation, analysis, inference, evaluation, explanation, and self-regulation" proposed by Facione (1990) [10] is widely regarded as the fundamental framework for instructional interventions in critical thinking. In addition, some scholars juxtapose formal logic with informal logic [11], arguing that mastery of logical knowledge and contextualized reflection are equally important [12]. Overall, these multidimensional interpretations at the conceptual and structural levels provide rich theoretical support for subsequent instructional design and practical interventions, while also laying a multidimensional foundation for researchers in selecting intervention components and evaluation indicators.

1.3 Instructional Interventions for Critical Thinking

In studies of instructional interventions, pathways for cultivating critical thinking primarily involve three key elements: instructional type, instructional model, and instructional strategy.

First, with regard to instructional type, Ennis (1989) [13] categorized critical thinking instruction into discipline-independent, discipline-engrafted, discipline-embedded, and discipline-mixed approaches. Subsequently, Çeviker and Orhan (2020) [14] demonstrated via a meta-analysis that across different cultural and educational contexts, both skill-based dedicated training and content-based integrated cultivation have yielded relatively positive outcomes. However, no scholarly consensus has been reached on which instructional type is more effective in practice, as existing empirical findings often diverge with variations in curriculum design, instructors' pedagogical beliefs, and relevant cultural and educational contextual factors. Second, concerning instructional models. From a more macro perspective, the four major teaching models proposed by Joyce [15], namely the Information Processing Model, the Social Interaction Model, the Behavioral Model, and the Personal Model, serve as a critical theoretical basis for the development of classroom teaching frameworks and the combination of teaching strategies. Existing literature indicates that in a social interaction model that emphasizes interaction and cooperation, students' participation levels have increased [16]; nevertheless, no consistent conclusions have been reached regarding whether it can significantly foster critical thinking [17]. In contrast, several studies have found that flipped classrooms or project-based learning (PBL), which are designed based on behaviorally oriented theories, can not only focus on observable behavioral changes in students but also promote the development of their critical thinking capacities [18]. Third, with respect to instructional strategies, more specific instructional strategies can be categorized into direct, indirect, interactive, experiential, and independent learning approaches, and they can also be implemented in a "hybrid" form [19]. Among these, indirect instructional strategies, as classified in various typologies, prioritize student-centeredness and do not overemphasize student-student or teacher-student interaction processes; instead, they focus on fostering students' learning experiences and independent learning capabilities. Yoon et al. (2017) [20] compared direct methods (teacher-led instruction) with indirect methods (student-centered inquiry) and found that different types of students may exhibit greater adaptability to particular strategies. Building on this foundation, the integration of interaction and experiential learning would not only afford students sufficient space for critical thinking but also create opportunities for diversified communication, which demonstrates considerable potential for fostering critical thinking skills. Nevertheless, more empirical research is required to examine under what curricular and teacher-student conditions different instructional strategies can yield optimal results.

1.4 Research Gaps in Existing Studies

In recent years, several meta-analytic studies have sought to synthesize empirical findings regarding instructional interventions targeting college students' critical thinking, thereby enhancing the objectivity and comparability of research conclusions. Some studies have focused on whether technological is incorporated into the teaching process [21], while others have centered on the duration of interventions and disciplinary backgrounds [22]. Although these studies have provided valuable insights for subsequent practices, they also have several limitations.

Firstly, there exist discrepancies in the intervention participants and literature inclusion criteria. Some studies have conflated senior high school students with undergraduates or included graduate students as research subjects, which compromises the targeted relevance of the findings to the undergraduate population. Secondly, there is a lack of granularity in the classification of instructional approaches. Many meta-analyses have only adopted a simplistic dichotomous classification (e.g., direct vs. indirect instruction, technology-integrated vs. non-technology-based instruction), which fails to capture the multidimensional nature of instructional models and strategies. Thirdly, the consideration of moderating variables is inadequate. Existing studies have insufficiently examined the interaction effects among curriculum domains, instructional tools, and instructional types and models [23], thus failing to fully address the core question of "which instructional interventions are most effective under which conditions."

1.5 Research Questions

To address the above gaps, the present study will further focus on the context of undergraduate education, and based on 59 experimental or quasi-experimental studies, conduct a more systematic meta-analysis on the effectiveness of instructional interventions for critical thinking by introducing a more granular classification of instructional models, instructional strategies, and moderating factors. Through high-quality quantitative synthesis, the study seeks not only to provide actionable guidance for higher education curriculum reform but also to academically enrich the empirical understanding of the educational imperative of advancing higher-order thinking.

2. Method

2.1 Research Design

This study adopts a meta-analytic approach to synthesize the findings of empirical research on instructional interventions for undergraduate critical thinking. Meta-analysis is a data analysis method that conducts a secondary analysis of the results of multiple independent quantitative studies addressing the same research question, thereby drawing more generalizable conclusions [24]. This study employs CMA 3.0 software as the analytical tool. Sample sizes, standard deviations, means, and other numerical data from the literature are inputted into CMA for statistical computation. The standardized mean difference (SMD) is selected as the effect size to represent the extent to which different teaching approaches influence the development of critical thinking among undergraduates.

2.2 Data Sources and Search Strategy

A comprehensive literature search was conducted across both Chinese and international academic databases. Chinese databases included China National Knowledge Infrastructure (CNKI) and Wanfang Data. International databases included the Web of Science Core Collection, specifically SCI-EXPANDED, SSCI, A&HCI, and ESCI. The search terms consisted of combinations of keywords such as "Critical Thinking" and "Methods for Cultivating Critical Thinking", using Boolean operators (e.g., AND, OR) to maximize coverage. The retrieval period covered publications indexed from January 2014 to December 2024. To ensure consistency across databases, articles published in 2025 were not included. In addition, manual searches of reference lists were conducted to identify potentially relevant studies not captured through database searches.

2.3 Inclusion and Exclusion Criteria

This study further conducted a secondary screening of the literature based on the following inclusion criteria: (1) the research topic was limited to the instructional effectiveness of critical thinking, specifically focusing on classroom teaching practices; (2) the research design was experimental, quasi-experimental, or pretest–posttest control group design; (3) the articles were open-access and available in full text; and (4) the instructional objective targeted the development of students' critical thinking skills and/or dispositions, with teachers implementing relevant teaching methods or strategies in classroom instruction. The exclusion criteria were as follows: (1) duplicate publications; (2) studies that did not report means or standard deviations (SD); (3) experiments lacking a control group; and (4) control groups with a standard deviation of 0.00. The literature screening was conducted independently by two researchers in duplicate, yielding a final set of 59 studies.

2.4 Coding of the Studies

This study first identified experimental and quasi-experimental studies on methods for cultivating critical thinking and coded their variables. The selection of studies was conducted collaboratively by two researchers. Initially, the abstracts of the preliminarily screened studies were independently reviewed and analyzed. For items where disagreements arose, the full text of the papers was consulted, with a focus on research objectives and methods, and discrepancies were resolved through discussion. The full texts of the studies passing the secondary screening were then examined, and effect sizes or reported statistics were verified to confirm the final sample of studies, after which a coding framework was developed. Based on this framework, one researcher coded the data while the other independently verified and reviewed the codes. By referring to the review of research summaries in previous meta-analyses and the content of the coding framework, this study focused on coding the teaching process. This aspect was often overlooked in previous meta-analysis studies, thus forming a meta-analysis coding framework.

This study treats critical thinking as the dependent variable, with course type, intervention duration, and the use of instructional aids as moderator variables. Course types are classified into humanities and social sciences courses, natural sciences courses, and integrated courses integrating humanities and social sciences with natural sciences. Intervention duration is categorized into 1-8 weeks, 9-16 weeks, 17-32 weeks, and 33 weeks or longer. The use of instructional tools includes various forms such as multimedia technology, computer-assisted instruction, visualization tools, and virtual reality (VR) devices. Given the complexity and diversity of the implementation of instructional aids, it is not feasible to classify them into fixed categorical dimensions; therefore, this study only assessed the use or non-use of instructional aids. Additionally, this study treats instructional type, teaching model, and teaching strategy as independent variables to examine how different types, models, and strategies contribute to the development of students' critical thinking.

2.5 Data Analysis and Interpretation

The meta-analysis was conducted using Comprehensive Meta-Analysis (CMA) software. The standardized mean difference (SMD) was calculated for each study to represent the effect of instructional interventions on critical thinking.

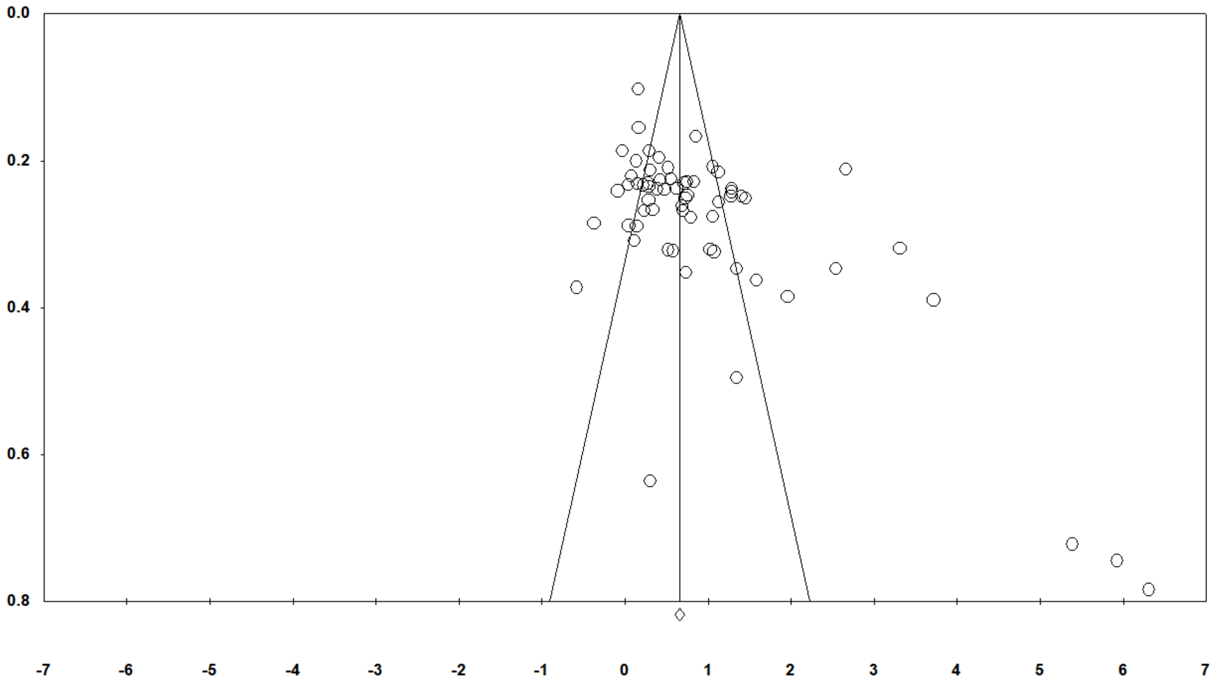
Heterogeneity and homogeneity testing are essential components of meta-analysis. Given that the included studies differ in participant characteristics, publication time, and experimental conditions, a certain level of heterogeneity is inevitable. Hedges and Olkin (2014) suggested that the primary test for effect size heterogeneity is Q statistics. The homogeneity test is based on the Q statistic; a statistically significant Q value indicates a heterogeneous distribution [25]. Heterogeneity is primarily assessed using the I^2 statistic. When I^2 exceeds 50%, substantial heterogeneity is considered to be present [26], and a random-effects model is required. The results of the heterogeneity and homogeneity tests in this study revealed that I^2 was 90.020%, exceeding the 50% threshold, and the Q value was 254,621 ($p < 0.001$), indicating a high degree of heterogeneity among the included studies. Therefore, a random-effects model was adopted to examine the effect of different teaching methods on the development of critical thinking skills among university students.

3. Results

3.1 Results on Publication Bias

In this study, both the funnel plot and the fail-safe N were employed to examine the presence of publication bias. Symmetrically is tributed ESs around the general ES in the funnel plot indicate that publication bias does not exist. As indicated by the visual inspection of the funnel plot (see Figure 1), the effect sizes of the included studies were symmetrically distributed around the central axis, suggesting a relatively low likelihood of publication bias. Also, fail-safe N number (9719) was much greater than the number (305) calculated by the formula of 5k+10 and this indicated a lack of publication bias. Consistent results from both methods confirm that publication bias is unlikely to have occurred in this meta-analysis.

Figure 1. *Funnel Plot.*



3.2 Overall Influence of Different Teaching Methods on the Cultivation of Critical Thinking of College Students

After several rounds of literature screening, 63 sets of sample data from 59 literatures were finally included in the meta-analysis. As seen in Table 1, the SMD was calculated as 0.661 with a Z price of 8.951 (p<0.001) under the random effects model and lower limit of the SMD was 0.601 and upper limit was 0.721 with 95% confidence interval. Also, the ES was 0.894 (lower limit: 0.698 and upper limit: 1.090) under the fixed effects model. So, I used random effects model to calculate the SMD in this meta-analysis. According to Cohen's established criteria for evaluating effect sizes (based on his 1977 [27] and 1988 [28] research), standardized mean difference values≤0.2 indicate a small effect, values around 0.5 suggest a moderate effect, while≥0.8 signifies a large effect. Therefore, different teaching methods generally demonstrate varying degrees of positive impact on cultivating critical thinking skills among college students.

Table 1: *The Estimated ESs and the Heterogeneity Test Results.*

Model	k	SMD	95% CI		Z	P	I²	Q	df	P
			Lower	Upper						
Random	63	0.661	0.601	0.721	8.951	0.000	90.020	621.254	62.000	0.000
Fixed	63	0.894	0.698	1.090	21.523	0.000				

3.3 Results on the Moderator Analyses

To further investigate the differential effects of these teaching approaches, this study conducted a comprehensive analysis across three dimensions and 16 specific dimensions, with the results presented in Table 2.

Table 2. *Results of the Moderator Analyses.*

	Sub-groups			95% CI		Two-Tailed Test		Heterogeneity test	
		k	ES	Lower	Upper	Z	P	Q	P
Teaching Type	Discipline-Independent Type	4	3.212	1.165	5.259	3.075	0.002	6.153	0.046
	Discipline-Infused Type	24	0.694	0.403	0.985	4.679	0.000		
	Discipline-Embedded Type	35	0.875	0.628	1.122	6.939	0.000		
Teaching Model	Behavioral Model	4	1.645	0.342	2.947	2.475	0.013	5.133	0.274
	IP Model	35	0.965	0.681	1.250	6.655	0.000		
	IP+SI Model	6	0.887	0.340	1.435	3.178	0.001		
	SI Model	15	0.603	0.237	0.970	3.225	0.001		

	Individual Model	3	1.106	0.735	1.477	5.848	0.000		
Teaching Strategy	Direct Instruction	5	0.567	0.160	0.974	2.730	0.006	9.383	0.095
	Indirect Instruction	19	0.826	0.476	1.176	4.627	0.000		
	Interactive Strategy	14	0.684	0.285	1.082	3.362	0.001		
	Interactive + Experiential Strategy	12	1.652	1.051	2.252	5.394	0.000		
	Experiential Strategy	5	0.888	0.043	1.733	2.060	0.039		
	Independent Learning	8	0.785	0.528	1.043	5.974	0.000		
Course Domain	Humanities & Social Sciences Courses	19	1.031	0.664	1.397	5.510	0.000	0.833	0.659
	Natural Sciences Courses	16	0.815	0.474	1.155	4.685	0.000		
	Interdisciplinary Courses	28	0.843	0.524	1.163	5.177	0.000		
Teaching Auxiliary Tools	Not Used	41	0.801	0.582	1.020	7.159	0.000	1.038	0.308
	Used	22	1.028	0.651	1.404	5.347	0.000		
Intervention Duration	1-8 Weeks	23	0.737	0.372	1.102	3.954	0.000	1.298	0.730
	9-16 Weeks	31	0.988	0.721	1.255	7.262	0.000		
	17-32 Weeks	7	1.017	0.291	1.742	2.748	0.006		
	≥33 Weeks	2	0.934	0.679	1.190	7.176	0.000		

3.3.1 Effects of Different Instructional Types on the Cultivation of Critical Thinking

At the instructional type level, discipline-independent type (SMD=3.212, p=0.002) had the greatest mean SMD. Discipline-embedded type (SMD=0.875, p<0.001), discipline-infused type (SMD=0.694, p<0.001) followed it, respectively. The effect values of these three teaching types were all above 0.5, indicating that regardless of whether critical thinking teaching is explicitly emphasized in the classroom, classrooms with critical teaching types will improve students' critical thinking skills. The between-group effect tests yielded Q=6.153, p=0.046 (p<0.05), demonstrating a statistically significant difference in the effectiveness among the three types of teaching approaches. Among them, the discipline-independent teaching approach yielded the optimal results, which verified the important value of the explicit integration logic in instructional design for cultivating students' critical thinking.

3.3.2. Effects of Different Instructional Models on the Cultivation of Critical Thinking

At the instructional model level, behavioral model (SMD=1.645, p=0.013) had the greatest mean SMD. Personal model (SMD=1.106, p<0.001), information processing model (SMD=0.965, p<0.001), information processing+social interaction model (SMD=0.887, p=0.001), and social interaction model (SMD=0.603, p=0.001) followed it, respectively. The between-group effect test showed Q=5.133, p=0.274 (p>0.05), indicating no statistically significant difference among the instructional models in fostering critical thinking. However, the behavioral model demonstrated comparatively greater effectiveness.

3.3.3. Effects of Different Instructional Strategies on the Cultivation of Critical Thinking

At the instructional strategy level, interaction+experiential strategy (SMD=1.652, p<0.001) had the greatest mean SMD. Experiential strategy (SMD=0.888, p =0.039), indirect strategy (SMD=0.826, p<0.001), independent learning (SMD=0.785, p<0.001), interaction strategy (SMD=0.684, p=0.001) and direct strategy (SMD=0.567, p=0.006) followed it, respectively. The between-group effect test yielded Q=9.383, p=0.095 (p>0.05), suggesting that there was no statistically significant difference among different instructional strategies in enhancing critical thinking. Nevertheless, strategies that simultaneously emphasize interaction and experiential learning were found to be the most effective.

3.4 Moderator Analysis

3.4.1 Effects of Different Course Domains on the Cultivation of Critical Thinking

Regarding the course domains in which critical thinking development has been investigated, researchers appear to show greater interest in exploring critical thinking instruction within interdisciplinary curricula (SMD=0.843, p<0.001), though the overall effectiveness is not particularly strong. Overall, interventions applied in the humanities and social sciences demonstrated the most substantial impact (SMD=1.031, p<0.001), while those conducted in natural science courses showed slightly weaker effects (SMD=0.815, p<0.001).

The between-group effect test yielded Q=0.833, p=0.659 (p>0.05), indicating no statistically significant differences across course domains. This suggests that, although humanities and social science courses appear more conducive to fostering critical thinking, the development of critical thinking is not strongly restricted by disciplinary differences.

3.4.2 Effects of Different Intervention Durations on the Cultivation of Critical Thinking

In terms of intervention duration, the period of 17 to 32 weeks (SMD=1.017, p=0.006) had the greatest mean SMD. And 9 to 16 weeks (SMD=0.988, P<0.001), 33 weeks and above (SMD=0.934, P<0.001), and 1 to 8 weeks (SMD=0.737, P<0.001) followed it, respectively. Given the proximity of effect sizes between the 9 to 16 week and 17 to 32 week groups, a further comparison was conducted, revealing no statistically significant difference between them (p=0.942>0.05) (see Table 3). Therefore, although the 17 to 32 week duration appears optimal, any intervention lasting 9 to 32 weeks can achieve desirable effects.

Table 3: *Post Hoc Comparison Between Selected Intervention Duration Subgroups.*

	Sub-groups			95% CI		Two-Tailed Test		Heterogeneity test	
		k	ES	Lower	Upper	Z	P	Q	P
Intervention Duration	9-16 Weeks	31	0.988	0.721	1.255	7.262	0.000	0.005	0.942
	17-32 Weeks	7	1.017	0.291	1.742	2.748	0.006		

3.4.3 *Effects of the Use of Instructional Support Tools on the Cultivation of Critical Thinking*

An analysis of the included literature showed that the presence or absence of instructional support tools did not exert a notable impact on critical thinking development (Q=1.038, p=0.308, p>0.05). However, the effect size of using multimedia, computer-assisted, and visualization tools (SMD=1.028, p<0.001) was slightly greater than that of teaching approaches not incorporating such tools (SMD=0.801, p<0.001).

This result indicates that, although the use of instructional tools does not produce a statistically significant difference, computer-assisted instructional tools demonstrate a marginally stronger effect on enhancing students' critical thinking than traditional teaching aids.

4 Research Conclusions and Discussion

Based on the results of empirical analysis, it can be concluded that the teaching interventions targeting college students' critical thinking have an overall significant facilitating effect. Meanwhile, different dimensions of teaching types, teaching models and teaching strategies also exhibit certain characteristics of difference and complexity.

4.1 Significance of the Overall Effect: The Improvement Effect of Critical Thinking Is Conclusive

The results of the meta-analysis showed that after pooling 63 effect sizes, the overall effect size (SMD) was approximately 0.661, which reached a statistically significant level (p<0.001). According to the effect size criteria proposed by Cohen, a value of 0.661 is close to or even exceeds the lower limit of a large effect. This indicates that well-designed college students' critical thinking teaching activities with a certain intensity of intervention can bring about a considerable improvement in students' cognitive skills and thinking dispositions. Compared with the viewpoints of some previous studies that questioned whether "critical thinking cannot be directly taught" or that "the improvement of critical thinking is limited" [29], the quantitative analysis in this paper provides more robust empirical support. On the one hand, this shows that regardless of the differences in teaching interventions in terms of curriculum content, teaching tools or time span, as long as the objectives are clear and the implementation is appropriate, students' critical thinking and reflective abilities can be promoted to a certain extent. On the other hand, this also echoes the call of international teaching reforms for the cultivation of higher-order thinking, namely, shaping students' more diverse and profound academic and thinking literacy through systematic teaching interventions [30].

4.2 Teaching Types: The Discipline-Specific Model Demonstrates a More Prominent Effect on Critical Thinking Cultivation

When comparing three teaching models, this study found the subject-independent model had the highest effect size (SMD=3.212), followed by the subject-integrated model (SMD=0.875) and the subject-attached model (SMD=0.694). The intergroup difference test also reached a statistically significant level (Q=6.153, p=0.046). Between-group difference tests were statistically significant (Q=6.153, p=0.046). This result is consistent with some prior studies but conflicts with others [31], reflecting the complexity of the field.

Firstly, the advantages of the discipline-specific model. The results of this study indicate that offering standalone courses with critical thinking cultivation as the primary objective can deliver more targeted training and conceptual instruction, thereby achieving relatively notable effects in the short term. Such courses typically include specialized training in thinking skills (e.g., reasoning, argumentation, and logical analysis) as well as intensive practice for nurturing thinking dispositions. The key factors contributing to its marked effectiveness may lie in students' clear awareness of learning goals and teachers' consistent focus on thinking training throughout the instructional process.

Secondly, the values and limitations of discipline-integrated models (the grafted and inherent models). Existing literature suggests that embedding critical thinking into disciplinary knowledge can enhance students' depth of thinking in more authentic scenarios [32]. However, the findings of this study show that the discipline-grafted model has a notably lower effect size, which may imply that teachers fail to fully integrate disciplinary content with thinking skills in practice, or that they are frequently constrained by curriculum schedules and examination pressures. Although the discipline-inherent model holds relative advantages, its implicit cultivation approach may compromise the systematicity of thinking training in the absence of clear goal guidance and evaluation criteria.

In conclusion, discipline-integrated courses are not inherently ineffective; instead, their effectiveness depends largely on whether curriculum design can achieve in-depth alignment between disciplinary content and thinking skills. The findings of this study offer an implication for colleges and universities in designing critical thinking training courses: clear thinking objectives and dedicated intervention pathways are generally more likely to yield prominent outcomes than incidental or implicit approaches.

4.3 Teaching Models: Behavioral and Personal Models Are More Conducive to Strengthening Thinking Transformation

The results of the meta-analysis on multiple teaching models, including information processing, social interaction, behavioral, and personal models, showed that the behavioral model achieved the optimal effect (SMD=1.645), followed by the personal model (SMD=1.106), the information processing model (SMD=0.965), the combined information processing and social interaction model (SMD=0.887), with the social interaction model ranking last (SMD=0.603). However, the intergroup differences did not reach a statistically significant level (Q=5.133, p=0.274).

First, the prominent effect of the behavioral model. This model emphasizes explicit behaviors and observable feedback, and it highlights the use of a series of teaching tasks such as stimulation, reinforcement, and control to enable students to internalize the transformation of thinking patterns in the process of behavioral change. Its subordinate teaching methods, such as flipped classrooms and project-based learning, tend to yield more significant learning outcomes if designed in accordance with the logic of thinking training. The findings of this study have once again verified the potential of such teaching models in critical thinking cultivation from a quantitative perspective.

Second, the internal driving force of the personal model. The personal model focuses on the development of individuals' self-concept, with self-reflection and self-monitoring as its core tenets. In contrast, the social interaction model places greater emphasis on group cooperation and interpersonal communication. Yet, the results of this study indicated that the role of social interaction in promoting critical thinking is relatively limited (SMD=0.603). This does not negate the value of cooperative learning; instead, it suggests that university teachers should avoid superficial exchanges when designing interactive activities. Instead, they need to guide students to conduct in-depth speculation and questioning on arguments, logic, and evidence during group discussions or debates. Simple group discussions without carefully designed thinking challenges and feedback sessions will hardly achieve the desired effect of thinking training.

Therefore, the implication for university teaching is that the selection of teaching models should align with different teaching objectives and student characteristics. Behavioral and personal models are more likely to drive thinking reconstruction at both the explicit behavioral and internal self levels, and thus may have greater advantages in intervention-oriented teaching that prioritizes critical thinking. Meanwhile, to effectively improve students' critical thinking through the social interaction model, teachers need to embed key elements such as logical analysis and evidence evaluation into the interaction process.

4.4 Teaching Strategies: The Combined "Interaction+Experience" Deserves Attention

In the analysis of teaching strategies, the combined "interaction+experience" strategy yielded the optimal effect (SMD=1.652), followed by the experience strategy (SMD=0.888), the indirect strategy (SMD=0.826), the independent learning strategy (SMD = 0.785), the interaction strategy (SMD=0.684), with the direct strategy ranking last (SMD=0.567). The result of the intergroup effect size test for teaching strategies was Q=9.383, p = 0.095 (p>0.05), which did not reach a statistically significant level.

First, the synergistic effect of the combined strategy. The integration of interaction and experience strategies can bring about multiple learning experiences. On the one hand, interaction enables students to engage in criticism and analysis amid the collision of ideas; on the other hand, experiential learning allows students to personally experience the confusion and breakthroughs in the thinking process through scenario simulation, practical operation and other approaches. The mutual complementarity between "external communication" and "internal experience" may greatly stimulate students' thinking potential.

Second, reflections on other strategies. Pure indirect strategies (e.g., group cooperative inquiry) or experience strategies are sometimes susceptible to factors such as classroom atmosphere and student participation, and may fail to yield tangible results if improperly implemented. In contrast, direct strategies (e.g., teachers' one-way instruction of thinking skills) can instill knowledge rapidly under certain conditions, but they lack the cultivation of students' abilities for active thinking and questioning, thus producing the most limited effects.

This result verifies the conclusion of existing literature that student-centeredness can effectively stimulate thinking potential [33], and also suggests that teachers need to focus on complementarity and integration when selecting teaching strategies.

4.5 Moderating Variables: The Impacts of Course Discipline, Intervention Duration and Teaching Tools Are Relatively Limited

First, regarding course discipline. In this study, courses were categorized into three types: humanities and social sciences, natural sciences, and integrated courses. Although the humanities and social sciences courses yielded a slightly higher effect size (SMD=1.031) than natural science courses (SMD=0.815) and integrated courses (SMD=0.843), the intergroup differences were not statistically significant (Q=0.833, p=0.659). This indicates that the cultivation of critical thinking does not necessarily depend on specific disciplines, but rather on the quality of curriculum design and teacher guidance.

The second point is about the intervention time. Intervention durations of 1 to 8 weeks, 9 to 16 weeks, 17 to 32 weeks, and over 33 weeks all exert a positive facilitative effect on critical thinking; among these, the 9 to 16 weeks interval (SMD=0.988) and the 17 to 32 weeks interval (SMD=1.017) yield slightly stronger effects than other groups, while no substantial overall difference is observed (Q=1.298, p=0.730). This finding is basically consistent with the conclusion of some existing literature that "medium-term interventions (10 to 20 weeks) are more effective" [34]. Therefore, in university teaching, it is necessary to avoid both the blind pursuit of quick-fix short-term interventions and unplanned lengthy interventions, so that students can complete the transfer and consolidation of thinking skills within a moderate time frame.

Third, regarding teaching auxiliary tools. The statistical analysis showed no significant difference between the tool-use group and the non-tool-use group (Q=1.038, p=0.308), yet the effect size of the tool-use group (SMD=1.028) was slightly higher than that of the non-tool-use group (SMD=0.801). It can thus be seen that the integration of digital or visual tools can provide certain support, but they are not the decisive factor in cultivating critical thinking. In other words, technology is merely a carrier; if the teaching model and content do not meet the needs of critical thinking training, the use of tools per se will be difficult to produce a significant effect.

5 Implications and Recommendations

Based on the empirical analysis of teaching interventions for college students' critical thinking in this study, and by synthesizing the core findings of Part IV Research Conclusions and Discussion, several highly operable implications and recommendations are proposed. These are intended to provide practical pathways for thinking cultivation to support curriculum reform and teaching practice in institutions of higher education. Overall, it is recommended that efforts should be made to advance in a coordinated manner from five dimensions: curriculum design, teaching models, teaching strategies, technology and resources, and evaluation mechanisms, with an emphasis on the systematicness and in-depth nature of thinking cultivation.

5.1 Construct Context-Appropriate "Specialized" or "Embedded" Thinking Courses

First, offer specialized thinking courses. The research findings indicate that the discipline-specific teaching model yields the most significant effect on enhancing critical thinking, which suggests that students are more likely to achieve notable improvements in a short period when course objectives focus on strengthening thinking skills and dispositions. In this regard, institutions of higher education may provide students with a dedicated environment for thinking training by offering courses such as General Critical Thinking or Critical Thinking Training. Course syllabi should clearly define learning objectives for thinking skills including logic, argumentation, and reasoning, and incorporate specific training modules such as case analysis, thinking competitions, and concept discrimination, so as to ensure that students lay a solid foundation in a relatively pure "thinking training ground".

Second, promote embedded cultivation. For institutions or programs that are temporarily unable to offer standalone thinking courses, critical thinking objectives can also be explicitly embedded in professional courses and reflected in teaching design and evaluation criteria in a transparent manner. For example, integrate structured expression and critical analysis modules of "thesis-argument-evidence" into courses in literature and social sciences; emphasize hypothesis testing, exploration of alternative solutions, and critical questioning of experimental results in science and engineering courses. The key is to prevent "implicit infiltration" from becoming a mere formality, and to treat thinking cultivation as a goal of equal importance to the teaching of professional knowledge.

5.2 Focus on Observable Behaviors and Self-Reflection to Optimize the Application of Teaching Models

First, give play to the guiding role of the behavioral model. The Behavioral Model has demonstrated favorable intervention effects in empirical research, which suggests that teachers can utilize "visible behavioral change" methods, including project-based learning, flipped classrooms and simulated scenarios, to guide students to develop critical thinking in the process of continuously completing high-level tasks. At the operational level, teachers need to closely align tasks with thinking elements: for instance, design explicit thinking analysis tasks in flipped classrooms, requiring students to submit a concise argumentation framework after video learning and pre-class reading, then conduct demonstrations of observable behaviors and mutual criticism in class.

Second, emphasize the self-reflection mechanism of the personal model. The personal model highlights the importance of students' internal reflection, autonomous monitoring, and metacognitive regulation. Teachers can leverage tools such as thinking journals and thinking maps to help students record and analyze their own thinking processes, and regularly organize feedback discussions or one-on-one guidance focusing on the thinking process. For example, upon the completion of an experiment or a large-scale project, teachers can ask students to conduct in-depth self-criticism on potential flaws in the experimental design and limitations of data analysis, then summarize improvement strategies. This ensures that the cultivation of critical thinking is not only confined to the behavioral level but also penetrates into students' internal cognitive processes.

5.3 Ingeniously Integrate Interaction and Experience to Enhance Classroom Effectiveness

First, move from "single interaction" to "in-depth interaction". Interaction strategies can effectively improve student participation, but if they only stay at the level of free discussion or superficial information exchange, it will be difficult to touch the depth of thinking. Teachers can pre-design multi-level interactive sessions in class, such as "opinion debates" and "peer evaluation", with clear thinking tasks and evaluation criteria matched to each session. For example, students can be assigned roles as "advocates", "opponents", or "mediators" to conduct multi-dimensional debates on the same issue. Through role swapping and rule constraints, interactive sessions can be prevented from degenerating into casual chats.

Second, construct immersive experiential teaching. Experiential strategies enable students to go through the entire process of problem exploration in real or simulated scenarios. In practice, it can be combined with links such as case simulation, social surveys, enterprise research, and public welfare project practice to create realistic challenging scenarios where students are faced with complex and ambiguous problems. For instance, through "micro social hot topic cases", student teams can simulate the policy-making process, which requires dialectical thinking from multiple perspectives and evidence angles. If supplemented by timely teacher guidance and peer feedback, critical thinking can be strengthened under the "dual drive" of interaction and experience.

5.4 Rationally Utilize Technical Tools and Multi-Dimensional Resources to Promote the Upgrade of Blended Teaching

First, attach importance to the appropriateness and embedding degree of tools. The research shows that technical tools themselves are not decisive factors, but their rational use can bring certain gains. When selecting multimedia, visualization software, or artificial intelligence-assisted platforms, teachers should focus on "whether they can promote students' reflection on the thinking process and results". For example, use mind mapping software to help students sort out argument chains and mark logical loopholes in real time; set up online assessments containing critical guiding questions on MOOC platforms to help students continuously track their thinking progress.

Second, construct learning analysis and adaptive systems. With the development of big data and artificial intelligence technologies, colleges and universities can try to build data-driven learning analysis platforms to conduct visual diagnosis of students' real-time performance in critical thinking training and provide personalized feedback and resource recommendations. Through process-oriented learning data (such as online discussion speeches, homework logs, assessment scores, etc.), teachers can more accurately understand students' ability performance in various dimensions such as questioning, analysis, and reasoning, and dynamically adjust teaching strategies accordingly.

5.5 Establish a Long-Term Evaluation and Feedback Mechanism to Strengthen the Sustainability of Thinking Cultivation

First, implement diversified evaluation that emphasizes both process and outcome. To enable students to achieve sustainable progress in critical thinking, institutions of higher education need to construct a comprehensive assessment system integrating formative evaluation and summative evaluation. Formative evaluation can adopt diverse forms such as thesis writing, opinion presentations, and case analysis to provide phased feedback on students' reasoning structure, evidence application, and depth of reflection. Summative evaluation, on the other hand, focuses on students' critical thinking proficiency within a complete knowledge framework, such as through graduation theses and comprehensive project research. This approach not only helps students continuously adjust their thinking habits, but also allows teachers to improve their teaching design in real time.

Second, encourage self and peer assessment to expand the sources of feedback. Beyond professional teacher assessment, self-assessment and peer mutual assessment also play a critical role in the development of critical thinking. A "peer review" function can be designed on the course platform, with review criteria jointly formulated by teachers and students, such as the clarity of arguments, the rationality of supporting evidence, and the coherence of logical reasoning, to enable students to conduct mutual reviews and provide constructive feedback for each other. For self-assessment, it guides students to review their own thinking process, identify cognitive blind spots and biases, and develop stronger learning motivation.

In conclusion, the cultivation of critical thinking is by no means an isolated teaching segment, but should be rooted in the broader ecosystem of higher education. Institutions of higher education should make coordinated efforts across multiple dimensions, including talent cultivation objectives, curriculum system layout, teachers' professional development, resource investment, and supporting policies, to promote the systematic cultivation of students' higher-order thinking abilities. In the era of digital intelligence, critical thinking is not only a key indicator of college students' personal competitiveness, but also a core driving force for universities to seize the initiative in educational reform and improve the quality of teaching. Only through top-level design and continuous exploration in teaching practice can the outcomes of critical thinking teaching interventions truly take root, helping college students go deeper and farther on the path toward developing higher-order thinking.

6 Limitations

Despite the valuable insights provided by this meta-analysis, several limitations should be considered. First, the studies included in this review were primarily experimental and quasi-experimental in design, which may limit the generalizability of the findings to real-world classroom settings where controlled experimental conditions are difficult to maintain. Second, the sample size of the studies included varied, with some studies involving small sample sizes, which could introduce bias and affect the overall robustness of the findings. Additionally, while the meta-analysis accounted for a range of instructional models, strategies, and types, the variability in implementation methods across different studies may lead to differences in outcomes that are not fully accounted for. Third, the majority of studies focused on short-to medium-term interventions, and long-term effects of instructional interventions on critical thinking remain unclear. Finally, the limited number of studies that incorporated digital tools and resources in the interventions may also affect the generalizability of the results in the increasingly digital landscape of higher education. Future research could explore these limitations by conducting large-scale, long-term studies that more effectively capture the diverse instructional practices in varied educational contexts.

References

1. Ma, S. F., & Yang, X. D. (2021). Cooperative reasoning-based learning to promote higher-order thinking. Research on Educational Development, 41(24), 64-73.
2. Conrad, H. (2025, February 15). The Universal Learning Programme: Educating future-ready citizens. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000372613
3. Vincent-Lancrin, S., et al. (2019). Fostering students' creativity and critical thinking: What it means in school, educational research and innovation. OECD Publishing.
4. Shi, Y. N., Shi, Y. H., & Wu, D. (2018). The development of new competencies for future education: ICT literacy connotation, evolution, and its implications. Comparative Education Research, 40(03), 3-11.
5. Li, Z. Q., & Li, Y. X. (2014). The cultivation of critical thinking in American universities and its implications for English teaching in China. Chinese Foreign Language, 11(06), 14-20.
6. Wu, H.Z. (2016). The soul of critical thinking — Rational standards. Studies in Logic, 9(03), 25-41.
7. Yu, T. F. (2021). The debate on the "generalizability" of critical thinking and its cognitive psychology explanation. Journal of Northwest Normal University (Social Sciences Edition), 58(05), 110-118.
8. Zhuo, L. (2024). The integration of ideological and political education in foreign language courses under the perspective of critical thinking: A case study of the Basic Japanese course. Journal of Tibet University (Social Sciences Edition), 39(02), 234-240.
9. Luo, R., & Wei, W. (2024). The application of critical thinking in English and American literature — A review of Contemporary Western Literary Criticism Methods and Practices. Science and Technology Management Research, 44(02), 223-224.
10. Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Millbrae: The California Academic Press.
11. Salmon, M. (2002). Introduction to Logic and Critical Thinking (4th ed.). Wadsworth, New York: Thomson Learning.
12. Cogan, R. (1998). Critical Thinking: Step by Step. Lanham: University Press of America, Inc.
13. Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. Educational Researcher, 18(3), 4-10.
14. Çeviker, A. Ş., & Orhan, A. (2020). The effect of different critical thinking teaching approaches on critical thinking skills: A meta-analysis study. Pamukkale University Journal of Education Faculty, 49, 88-111.
15. Joyce, B., Weil, M., & Calhoun, E. (2015). Models of Teaching (9th ed.). New York: Pearson Education, Inc.
16. Hajhosseini, M., Zandi, S., et al. (2016). Critical thinking and social interaction in active learning: A conceptual analysis of class discussion from Iranian students' perspective. Cogent Education, 3(1), 1-9.

17. Ye, Y. H., & Yin, Y. M. (2019). The cognitive characteristics and cultivation strategies of college students' critical thinking — An empirical study based on group cooperative inquiry. Research in Educational Development, 39(11), 66-74.
18. Etemadfar, P., Soozandehfar, S. M. A., & Namaziandost, E. (2020). An account of EFL learners' listening comprehension and critical thinking in the flipped classroom model. Cogent Education, 7(1), 1-22.
19. Du, X., Zhang, L., et al. (2024). Analyzing the effects of instructional strategies on students' on-task status from aspects of their learning behaviors and cognitive factors. Journal of Computing in Higher Education, 36(1), 29-56.
20. Yoon, C. H., Na, J. H., & Bak, B. G. (2017). A meta-analysis of the effects of critical thinking instruction. The Korean Journal of Educational Psychology, 31(1), 1-34.
21. Wang, H. C., Zheng, Y. L., & Ma, Y. (2024). Does technology promote the development of students' critical thinking? A meta-analysis of 42 experimental and quasi-experimental studies. Higher Education Exploration, 01, 120-128.
22. Leng, J., & Lu, X. X. (2020). Can critical thinking truly be taught? A meta-analysis of 79 experimental or quasi-experimental studies. Open Education Research, 26(06), 110-118.
23. Huber, C. R., & Kuncel, N. R. (2016). Does college teach critical thinking? A meta-analysis. Review of Educational Research, 86(2), 431-468.
24. Lipsey, M. W., & Wilson, D. B. (2019). Meta-Analysis Methods and Applications (Liu, J., & Wu, C. Y., Trans.). Chongqing: Chongqing University Press.
25. Kulinskaya, E., & Dollinger, M. B. (2015). An accurate test for homogeneity of odds ratios based on Cochran's Q-statistic. BMC Medical Research Methodology, 15, 49.
26. Borenstein, M. (2023). How to understand and report heterogeneity in a meta-analysis: The difference between I-squared and prediction intervals. Integrative Medicine Research, 12(4), 1-8.
27. Cohen, J. (1977). Statistical Power Analysis for the Behavioral Sciences (Rev. Ed.). New York: Academic Press.
28. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Erlbaum.
29. Zhang, Q. G., & Tang, H. L. (2021). Course learning and the value-added development of undergraduates' critical thinking abilities: An analysis based on the 2016-2019 National Undergraduate Ability Tracking Survey data. Higher Education Research, 42(08), 79-88.
30. Global Education Monitoring Report Team, International Task Force on Teachers for Education 2030. (2025, February 15). Inclusive teaching: Preparing all teachers to teach all students. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000374447/PDF/374447eng.pdf.multi
31. Ren, X. Z., & Jia, Y. J. (2023). The status, effectiveness, and influencing factors of critical thinking teaching in Chinese universities: A meta-analysis based on 20 years of research on critical thinking teaching in national universities. Higher Education Research, 44(10), 64-76.
32. Abrami, P. C., Bernard, R. M., Borokhovski, E., et al. (2008). Instructional interventions affecting critical thinking skills and dispositions: A Stage 1 meta-analysis. Review of Educational Research, 78(4), 11 02–1134.
33. Huang, W., Cao, R., Qi, Y., et al. (2019). Strategies to enhance critical thinking ability in the age of artificial intelligence: An empirical study on the application of thinking mapping to improve primary school students' critical thinking ability. China Educational Technology, 10, 102-108.
34. Niu, L., Behar-Horensten, L. S., & Gauvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. Educational Research Review, 9, 114-128.
35. Cai, X., Li, Z., Zhang, J., et al. (2022). Effects of ARCS model-based motivational teaching strategies in community nursing: A mixed-methods intervention study. Nurse Education Today, 119, 1-8.
36. Cooper, H. (2020). Meta-Analysis Methods (5th ed.). Li, C., & Zhang, Y. (Trans.). Beijing: China Renmin University Press.
37. Dong, Y. (2019). Ten Lectures on Critical Thinking: From Inquiry and Argument to Open Creativity. Shanghai Education Press.
38. Lin, S. Q., & Luo, L. (2024). A summary of the establishment of the Critical Thinking Professional Committee of the Chinese Logic Society and the first academic seminar. Journal of Sichuan Normal University (Social Sciences Edition), 51(02), 193-196.
39. Orhan, A., & Çeviker, A. Ş. (2023). How to teach critical thinking: An experimental study with three different approaches. Learning Environments Research, 26(1), 199-217.