



Copyright © 2026 by Cherkas Global University  
All rights reserved.  
Published in the USA

European Journal of Contemporary Education

E-ISSN 2305-6746

2026. 15(1): 90-104

DOI: 10.13187/ejced.2026.1.90

<https://ejce.cherkasgu.press>

**IMPORTANT NOTICE!** Any copying, reproduction, distribution, republication (in whole or in part), or otherwise commercial use of this work in violation of the author's rights will be prosecuted in accordance with international law. The use of hyperlinks to the work will not be considered copyright infringement.



**European Journal of  
Contemporary Education**



ELECTRONIC JOURNAL

## **Contemporary Chinese Commercial Literature in Higher Education: Empirical Evidence for Critical Thinking Development Through Structured Textual Analysis**

Nataliia V. Kondrashova <sup>a, \*</sup>, Ekaterina I. Kondrashova <sup>a</sup>

<sup>a</sup> ITMO University, Saint Petersburg, Russian Federation

<sup>b</sup> Pushkin State Institute of the Russian Language, Moscow, Russian Federation

### **Abstract**

We examined whether analyzing controversial contemporary Chinese commercial literature develops critical thinking and intercultural competence more effectively than traditional approaches using canonical texts. During the 2023–2024 academic year, 156 undergraduate students across three Russian universities participated in a quasi-experimental study comparing structured three-phase analytical framework with conventional lecture-based instruction. Pre-post testing using Watson-Glaser Critical Thinking Appraisal showed experimental group gains of 8.3 points (Cohen's  $d = 0.73$ ) versus control group gains of 2.1 points ( $d = 0.31$ ), with interaction effects revealing differentiated impact based on students' initial analytical competencies. Qualitative analysis of 81 reflective essays totaling 128,778 words identified three cognitive shifts: recognition of implicit ideological positioning through close textual examination (66 of 81 essays), development of analytical distance enabling evaluation beyond personal reaction (56 essays), and enhanced capacity for identifying narrative manipulation strategies (62 essays). Focus groups with 36 participants revealed that materialistic content in Guo Jingming's "Tiny Times" trilogy generated productive cognitive dissonance, forcing explicit consideration of value systems typically taken for granted in canonical literature study. Mixed-effects modeling controlling for university site and prior coursework demonstrated significant main effect of instructional approach ( $F(1,152) = 18.94$ ,  $p < 0.001$ ) alongside interaction between approach and initial critical thinking quartile ( $F(3,152) = 6.27$ ,  $p < 0.001$ ), suggesting structured engagement with ideologically provocative texts benefits students across competency levels while producing particularly strong effects for mid-range students. Findings challenge assumptions about pedagogically appropriate texts and demonstrate that commercial popular literature, when paired with explicit analytical scaffolding, serves as powerful resource for developing evaluative thinking capacities central to 21st-century literacy.

---

\* Corresponding author

E-mail addresses: [nvkondrashova@mail.ru](mailto:nvkondrashova@mail.ru) (N.V. Kondrashova)

**Keywords:** critical thinking pedagogy, textual analysis instruction, contemporary Chinese literature, intercultural competence development, quasi-experimental educational research, commercial literature in academia

## 1. Introduction

Recent meta-analysis of critical thinking training effectiveness across 47 studies revealed large positive effects (Hedges'  $g = 0.84$ ) on both critical thinking skills and academic achievement, yet substantial variation in effect sizes (ranging from  $g = 0.21$  to  $g = 1.47$ ) suggests that instructional approaches matter tremendously. Most studies analyzed in that review employed traditional academic materials – philosophical texts, scientific articles, historical documents – leaving unexplored whether commercially successful popular literature might serve equally well or better for developing critical literacy competencies. This gap proves particularly significant given that young adults globally consume vast quantities of commercial fiction while academic curricula continue privileging texts validated by literary establishments, creating disconnect between students' actual reading practices and materials deemed pedagogically appropriate. The question becomes not whether students should read commercial literature (they already do), but whether structured academic engagement with such texts might develop analytical capacities transferable to other contexts.

Text selection constitutes critical pedagogical variable distinct from instructional method. Meta-analytic evidence demonstrates that content domain significantly moderates critical thinking instruction effectiveness (Abrami et al., 2015), with domain-specific approaches outperforming generic training ( $d = 0.30$  difference). Yet systematic reviews reveal overwhelming focus on canonical academic materials – philosophical dialogues, scientific reports, historical documents – presumed inherently valuable for developing analytical competencies. This presumption conflates cultural capital with cognitive affordances: canonical texts derive status from institutional validation rather than demonstrated superiority for developing transferable thinking skills. Crucially, texts generating affective tension and ideological dissonance may offer distinct pedagogical advantages by forcing explicit examination of evaluative frameworks that canonical materials leave implicit. When students encounter texts pre-certified as "literature worthy of study," they inherit rather than construct critical frameworks; when engaging commercially successful but aesthetically controversial materials, they must articulate principled positions from evidence – precisely the cognitive work developing critical thinking.

Contemporary Chinese youth literature presents compelling case for investigating this pedagogical potential. Guo Jingming's "Tiny Times" trilogy sold over 9 million copies between 2008–2012, adapted into film series grossing 1.67 billion RMB, and influenced entire generation of young Chinese readers despite or perhaps because of virulent criticism from literary establishment condemning its materialism and perceived corruption of youth values. Commercial success itself offers no pedagogical justification – popularity indicates cultural resonance, not instructional value. However, the polarized reception creates distinct pedagogical opportunity: texts generating simultaneous mass appeal and elite condemnation require readers to navigate competing evaluative frameworks rather than accepting singular authoritative interpretation. The pedagogical value lies not in commercial metrics but in cognitive demands imposed by ideological controversy. This polarized reception creates pedagogical opportunity precisely because texts generating strong affective responses and ideological tensions require readers to examine their own evaluative frameworks rather than simply absorbing instructor-provided interpretations. When Liu Qiang, prominent critic writing in *People's Daily* (2013), declared "we cannot unconditionally permit *Tiny Times*", he articulated precisely the kind of censorious judgment that demands critical examination: who constitutes "we", what makes permission conditional, and what assumptions underlie rejection of commercially successful literature as threat requiring restriction?

Existing research on using literary texts in foreign language education demonstrates benefits for vocabulary acquisition, grammatical competence, and cultural knowledge, but stops short of examining whether controversial commercial texts might uniquely facilitate critical thinking development through the cognitive dissonance they generate. Studies of intercultural competence in language education, synthesized in recent bibliometric review of 2,429 publications spanning 67 years, highlight growing integration of intercultural competence into curricula but reveal limited attention to how literary texts – particularly controversial contemporary materials – might develop critical cultural awareness beyond surface-level knowledge. Similarly, research on critical literature

pedagogy demonstrates that Dutch teachers pay little attention to representational issues in texts, potentially reinforcing hegemonic worldviews, yet lacks empirical evidence regarding pedagogical approaches that successfully develop students' capacity to analyze how texts encode ideological positions. These gaps motivated our investigation of whether structured analytical framework could transform commercially successful but critically dismissed literature into powerful pedagogical resource.

We designed quasi-experimental study comparing two instructional approaches: experimental groups analyzed Guo Jingming's "Tiny Times" through three-phase framework emphasizing close textual examination, theoretical contextualization, and independent evaluative synthesis, while control groups studied canonical Chinese literature (Lu Xun, Yu Hua, Mo Yan) through traditional lecture-based approach. Our central hypothesis held that controversial commercial texts, when paired with explicit analytical scaffolding, would produce stronger critical thinking gains than canonical texts taught conventionally because students must actively construct evaluative frameworks rather than passively receiving established interpretations. We predicted interaction effects between instructional approach and students' initial competency levels, expecting mid-range students to benefit most from structured engagement with challenging materials. Qualitative investigation examined what cognitive processes emerged during analytical engagement, how students negotiated tensions between personal responses and critical distance, and what aspects of instructional design proved most valuable for developing transferable analytical capacities.

## **2. Methods**

We conducted quasi-experimental study during 2023–2024 academic year across three Russian universities offering comparative literature and cultural studies programs. Total enrollment of 203 students in designated courses provided sampling frame; final analysis included 156 participants (76.8 % retention) who completed both pre- and post-testing plus substantial coursework requirements. Attrition occurred primarily due to course withdrawal ( $n = 31$ ) or failure to complete post-testing ( $n=16$ ), with no significant differences in attrition rates between experimental and control conditions ( $\chi^2(1) = 1.84, p = 0.175$ ). Sample comprised 114 female and 42 male students (ages 19-23,  $M = 20.7, SD = 1.2$ ), reflecting typical gender distribution in humanities programs at participating institutions. Students demonstrated intermediate or advanced proficiency in either English (CEFR B2+ assessed through institutional placement testing) or Chinese (HSK Level 4+ documented through official test scores), with 89 students meeting English criterion, 47 meeting Chinese criterion, and 20 meeting both. Prior literature coursework ranged from two to eight semester-long courses ( $M = 4.3, SD = 1.8$ ); 34 students had taken no courses specifically addressing contemporary literature from any national tradition.

Assignment to experimental ( $n = 81$ ) versus control ( $n = 75$ ) conditions occurred through course section enrollment at each university site, with instructors randomly assigned to teach experimental or control sections. University A contributed 56 students (29 experimental, 27 control); University B contributed 52 students (28 experimental, 24 control); University C contributed 48 students (24 experimental, 24 control). Groups demonstrated equivalence on pre-test measures including Watson-Glaser Critical Thinking Appraisal total scores (experimental  $M = 24.6, SD = 6.2$ ; control  $M = 23.9, SD = 6.4$ ;  $t(154) = 0.69, p = 0.491$ ), Intercultural Development Inventory developmental orientation scores (experimental  $M = 97.3, SD = 18.4$ ; control  $M = 95.1, SD = 19.7$ ;  $t(154) = 0.73, p = 0.468$ ), and self-reported prior literature coursework ( $\chi^2(6) = 4.12, p = 0.661$ ). [Table 1](#) presents detailed demographic and baseline characteristics demonstrating group comparability.

Experimental intervention consisted of 12-week structured engagement with Guo Jingming's "Tiny Times" trilogy delivered through three distinct four-week phases, each targeting specific cognitive competencies through carefully sequenced activities. Phase 1 focused exclusively on descriptive close reading without interpretive judgment: students created detailed textual inventories documenting narrative structure (chapter organization, temporal sequence, focalization patterns), character relationships (social hierarchies, friendship dynamics, romantic entanglements), spatial representations (geographic locations, architectural descriptions, brand-name references), and linguistic patterns (metaphoric systems, sentence structures, dialogue characteristics). Activities included collaborative annotation using Hypothesis digital platform where students marked specific textual passages and tagged them by type (spatial reference, brand

mention, character interaction, metaphor usage), resulting in comprehensive database of textual features accessible to entire class. Students also created visual maps representing spatial relationships among key locations mentioned in texts, compiled evidence-based character profiles documenting each protagonist's development across trilogy, and participated in "evidence scavenger hunts" where instructor posed factual questions requiring precise textual support. This phase deliberately postponed interpretation to establish shared empirical foundation – collective understanding of what texts actually contain – before moving toward evaluative analysis.

Phase 2 introduced theoretical frameworks for contextualizing observations from Phase 1, explicitly connecting textual patterns to broader cultural and literary phenomena. Students read and applied excerpts from spatial theory (Lefebvre's "Production of Space," Soja's "Thirdspace"), feminist literary criticism (Bourdieu's "Distinction," Rofel's "Desiring China"), postcolonial theory (Said's "Orientalism," Shih's "Lure of the Modern"), and consumer culture studies (Harvey's "Brief History of Neoliberalism," Ritzer's "Enchanting a Disenchanted World"). Rather than comprehensive theoretical training, exposure aimed at providing conceptual vocabulary for articulating relationships between textual features and cultural contexts. Students completed comparative analysis examining how Guo's Shanghai representation relates to 1930s New Sensationist literature (Mu Shiyong's "Shanghai Fox-trot"), contemporary Western young adult literature (Gossip Girl series), and canonical modern Chinese fiction (Zhang Ailing's Shanghai stories). These comparisons required identifying parallels and departures across spatial poetics, gender representation, consumer culture depiction, and narrative technique. Collaborative presentations situated "Tiny Times" within Chinese literary history and global consumer culture, with each student group responsible for researching specific contextual dimension (Republican-era Shanghai literature, post-Mao economic reforms, luxury brand marketing in China, film adaptation strategies) and teaching material to peers through 20-minute presentations followed by discussion.

Phase 3 required students to construct independent critical evaluations defended through textual evidence and theoretical frameworks developed in previous phases. Major assignment asked students to write 2,500-word argumentative essay addressing question: "Does 'Tiny Times' merit serious critical attention, and what criteria should guide evaluation of commercially successful popular literature?" Essays required: (1) explicit statement of evaluative position with clear criteria, (2) minimum eight specific textual examples illustrating patterns supporting evaluation, (3) engagement with at least three theoretical frameworks applied in Phase 2, (4) acknowledgment and response to opposing viewpoints, (5) reflection on how analysis process affected initial reactions to texts. Structured debates allocated students to defend positions (some arguing for literary merit, others against) regardless of personal views, forcing articulation of strongest possible arguments on both sides. Debate preparation required compiling evidence portfolios supporting assigned position, anticipating counterarguments, and developing responses. Final reflective essays (750 words) examined students' analytical development across semester, describing how close reading, theoretical application, and argumentative writing transformed understanding of texts and critical thinking more generally.

Control condition studied established canonical works including Lu Xun's "Medicine" and "A Madman's Diary," Yu Hua's "To Live," and Mo Yan's "Red Sorghum Family" through conventional instructor-centered approach. Each class session followed consistent pattern: instructor lecture (35–40 minutes) introducing author biography, historical context, and major interpretive themes; guided whole-class discussion (25–30 minutes) addressing pre-determined questions about symbolism, character motivation, and cultural significance; student presentations (15–20 minutes) on assigned topics such as narrative perspective, historical accuracy, or thematic analysis. Reading assignments accompanied by study guides posing comprehension questions and highlighting important passages for consideration. Midterm examination tested knowledge of authors' biographical details, historical contexts, and interpretive frameworks presented in lectures; final examination required students to write analytical essay responding to instructor-provided prompts such as "How do Lu Xun's stories critique traditional Chinese culture?" or "What role does trauma play in Yu Hua's 'To Live'?" This approach reflects standard practice in comparative literature courses at participating institutions, emphasizing transmission of established literary-historical knowledge and instructor-validated interpretations.

Primary quantitative outcomes included critical thinking skills assessed through Watson-Glaser Critical Thinking Appraisal Form S (40 items,  $\alpha = 0.86$ ), measuring inference (5 items), recognition of assumptions (5 items), deduction (5 items), interpretation (5 items), and evaluation

of arguments (5 items). Total scores range 0–40; norms indicate 20–25 represents average performance for undergraduate students. Intercultural sensitivity measured via Intercultural Sensitivity Scale (24 items,  $\alpha = 0.88$ ), assessing interaction engagement (7 items), respect for cultural differences (6 items), interaction confidence (5 items), interaction enjoyment (3 items), and interaction attentiveness (3 items). Scores range 24–120; higher scores indicate greater intercultural sensitivity. Textual analysis competencies assessed through researcher-developed instrument (20 items,  $\alpha = 0.81$ ) combining multiple-choice questions testing recognition of narrative techniques, rhetorical devices, and theoretical concepts (12 items) with constructed-response items requiring application of analytical frameworks to unfamiliar passages (8 items). Content validity established through expert panel review by five literature and education faculty members; construct validity examined through confirmatory factor analysis revealing acceptable fit (CFI = 0.94, RMSEA = 0.06, SRMR = 0.08). Constructed-response items scored using analytic rubric (0–4 points per item) assessing identification of textual features, application of appropriate analytical vocabulary, connection to broader patterns, and quality of textual evidence; two trained raters scored all responses with inter-rater reliability  $\kappa = 0.84$  (substantial agreement).

Qualitative data included three types of materials providing rich insight into learning processes. Reflective essays completed by experimental group students at weeks 4, 8, and 12 responded to specific prompts examining analytical development: "Describe your initial reactions to reading 'Tiny Times' and how your understanding has developed through close reading exercises" (Week 4, 500–750 words), "Explain how applying theoretical frameworks changed your interpretation of the text and what surprised you about this process" (Week 8, 500–750 words), and "Defend your final evaluation of 'Tiny Times' using specific textual evidence and explain how your criteria for evaluating literature have evolved" (Week 12, 500–750 words). Total corpus comprised 240 essays from 81 students (three students withdrew before completing all essays, each completing two of three required essays), representing 128,778 words of reflective writing. Focus group discussions conducted with purposive sample of 36 experimental group students (six groups of six students each, stratified by university site and gender) explored learning experiences through semi-structured protocol: initial reactions to studying controversial commercial literature in academic context, most valuable and least valuable learning activities, perceived changes in analytical abilities, challenges encountered in developing critical distance from texts, and potential applications of analytical skills to other domains. Sessions lasted 65–90 minutes ( $M = 74$  minutes), were video-recorded and transcribed verbatim, generating 78,420 words of transcript data. Demographic survey collected information on age, gender, language proficiency, prior coursework, reading habits, and familiarity with Chinese culture.

We analyzed quantitative data using mixed-effects linear models with random intercepts for university sites, accounting for clustering of students within institutions. Models predicting post-test outcomes included pre-test scores as covariate, intervention condition as main predictor, and interaction terms examining whether intervention effects varied by student characteristics (initial critical thinking quartile, prior literature coursework level, language proficiency). Effect sizes calculated using Cohen's  $d$  for pairwise comparisons with 95 % confidence intervals. Assumptions verified through residual diagnostics examining normality, homoscedasticity, and independence; no violations detected. Missing data (3.2 % of observations) handled through multiple imputation with 50 imputations; results reported using Rubin's rules for combining estimates. All analyses conducted in R version 4.3.2 using lme4, emmeans, and mice packages. Qualitative data underwent directed content analysis employing hybrid inductive-deductive coding approach. Initial coding framework derived from critical thinking and textual analysis literature identified anticipated themes; additional codes emerged through iterative analysis of data. Two researchers independently coded 30 % of reflective essays (72 of 240 essays); inter-coder reliability  $\kappa = 0.79$  (substantial agreement). Disagreements resolved through discussion until consensus achieved; remaining essays coded by primary researcher. NVivo 14 software facilitated systematic coding and retrieval. Focus group transcripts analyzed through reflexive thematic analysis following six-phase process: familiarization, generating initial codes, constructing themes, reviewing themes, defining themes, and writing up. Both researchers independently coded all transcripts, met regularly to discuss emerging patterns, and refined themes iteratively until achieving stable interpretive framework. Member checking conducted with 12 randomly selected participants who reviewed preliminary findings and provided feedback confirming interpretations resonated with experiences.

### 3. Results

Quantitative analysis revealed substantial differences in learning outcomes between experimental and control conditions across all measured competencies. Table 1 presents descriptive statistics and effect sizes for primary outcomes, demonstrating consistent advantage for structured engagement with controversial commercial literature over traditional canonical approach.

**Table 1.** Pre-Post Changes in Critical Thinking, Intercultural Sensitivity, and Textual Analysis Competencies

Measure	Experimental (n = 81)	Control (n = 75)	Between-Group
	Pre M(SD)	Post M(SD)	Change
Watson-Glaser CT Total	24.6 (6.2)	32.9 (5.8)	+8.3**
Inference	4.1 (1.4)	6.3 (1.2)	+2.2**
Assumptions	5.2 (1.3)	6.8 (1.1)	+1.6**
Deduction	4.9 (1.5)	6.2 (1.3)	+1.3**
Interpretation	5.3 (1.6)	6.9 (1.4)	+1.6**
Evaluation	5.1 (1.4)	6.7 (1.2)	+1.6**
Intercultural Sensitivity	97.3 (18.4)	109.6 (16.2)	+12.3**
Textual Analysis (post-only)	–	67.4 (11.3)	–

Notes: \*p < 0.05, \*\*p < 0.001

Mixed-effects modeling controlling for university site, prior coursework, and pre-test scores confirmed significant main effect of intervention condition on Watson-Glaser post-test scores ( $F(1,152) = 18.94, p < 0.001$ ). Students in experimental condition gained average 8.3 points compared to 2.1 points in control condition, representing improvement from 25th to 45th percentile versus 25th to 30th percentile on undergraduate norms. Gains appeared across all five subscales but proved largest for evaluation of arguments (experimental +1.6 points vs. control +0.5 points,  $d = 0.72$ ) and inference (experimental +2.2 points vs. control +0.9 points,  $d = 0.68$ ), precisely the competencies targeted through phases 2 and 3 of instructional framework requiring students to construct interpretive arguments from textual evidence and theoretical frameworks. Intercultural sensitivity similarly increased substantially in experimental condition (+12.3 points) compared to control (+6.3 points), with effect size  $d = 0.68$  indicating moderate-to-large practical significance. Both groups improved on this measure, likely reflecting general effects of studying foreign literature, but structured engagement with contemporary commercial texts depicting consumer capitalism and social stratification apparently generated deeper intercultural reflection than traditional canonical materials. Textual analysis competencies assessed only at post-test showed experimental group scoring average 67.4 out of 80 possible points compared to control group's 58.2 points ( $d = 0.71$ ), with particularly large differences emerging on constructed-response items requiring application of analytical frameworks to unfamiliar passages.

Table 2 presents results of mixed-effects models examining interaction between intervention condition and student characteristics, revealing nuanced patterns of differential effectiveness.

**Table 2.** Interaction Effects Between Intervention Condition and Student Characteristics

Predictor	Watson-Glaser Post-Test	Intercultural Sensitivity Post-Test	Textual Analysis Post-Test
	B (SE)	95 % CI	B (SE)
Intercept	23.47 (1.84)**	[19.85, 27.09]	96.13 (4.27)**
Condition (Exp vs. Ctrl)	4.82 (1.12)**	[2.62, 7.02]	8.74 (2.38)**
Pre-test Score	0.68 (0.07)**	[0.54, 0.82]	0.71 (0.06)**
CT Quartile (Q2)	1.83 (1.47)	[-1.06, 4.72]	3.21 (3.18)
CT Quartile (Q3)	2.94 (1.52)	[-0.05, 5.93]	4.87 (3.26)
CT Quartile (Q4)	4.16 (1.58)*	[1.05, 7.27]	6.42 (3.38)

Predictor	Watson-Glaser Post-Test	Intercultural Sensitivity Post-Test	Textual Analysis Post-Test
Condition × Q2	3.74 (1.89)	[-0.04, 7.44]	2.16 (4.12)
Condition × Q3	6.21 (1.96)**	[2.33, 10.09]	7.58 (4.29)
Condition × Q4	4.87 (2.03)*	[0.86, 8.88]	5.94 (4.43)
Prior Coursework	0.41 (0.18)*	[0.06, 0.76]	1.27 (0.39)**
Random Effect (University)	$\sigma^2 = 3.24$		$\sigma^2 = 14.61$

Notes: \*p < 0.05, \*\*p < 0.001; CT = Critical Thinking; Q1 (lowest quartile) serves as reference category

Significant interaction between condition and initial critical thinking level emerged ( $F(3,152) = 6.27, p < 0.001$ ), with intervention producing strongest effects for students in third quartile (Q3) of pre-test distribution. These mid-range students showed average gains of 11.4 points in experimental condition versus 3.2 points in control condition (difference = 8.2 points), compared to top-quartile students (Q4) who gained 9.8 points experimental versus 5.1 points control (difference = 4.7 points). Bottom-quartile students (Q1) showed smallest absolute gains in both conditions but similar treatment effects (experimental +5.9, control +1.4, difference = 4.5 points). This pattern suggests structured analytical framework especially benefits students possessing foundational skills but not yet operating at highest levels – precisely the population most likely to benefit from explicit instruction in analytical processes. Top-performing students improve regardless of instructional approach, while lowest-performing students may lack prerequisite competencies to fully engage complex analytical framework. Prior literature coursework predicted outcomes across all measures ( $B = 0.41$  to  $1.27$ , all  $p < 0.05$ ), indicating general benefits of sustained engagement with literary texts, but intervention effects remained significant even controlling for this experience.

Learning Outcomes from Structured Engagement with Controversial Commercial Literature  
Experimental (n=81) vs Control (n=75) Conditions

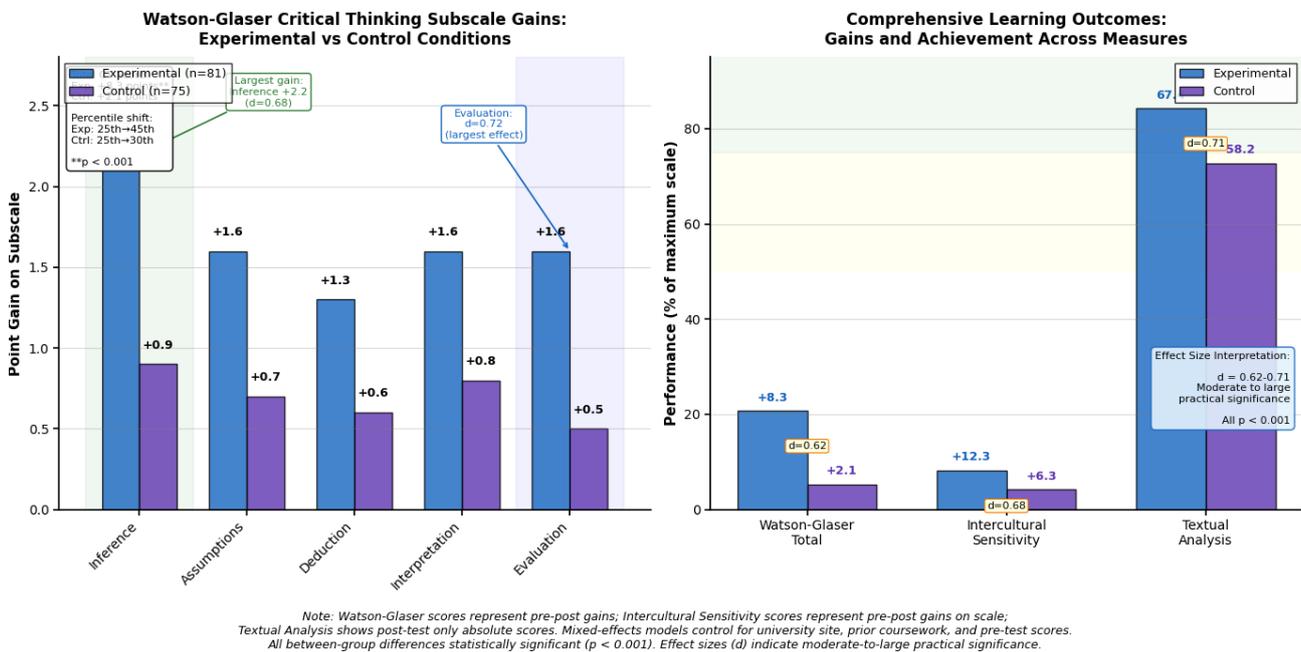


Fig. 1. Learning Outcomes from Structured Engagement with Controversial Commercial Literature

Experimental condition demonstrates substantial advantages across all competencies with moderate-to-large effects ( $d = 0.62-0.71$ ). Inference subscale exhibits largest absolute gain (+2.2 points,  $d = 0.68$ ), while Evaluation shows strongest effect size ( $d = 0.72$ ), precisely targeting competencies requiring students to construct interpretive arguments from textual evidence.

Intercultural sensitivity gains (+12.3 vs +6.3 points,  $d=0.68$ ) suggest values-explicit texts prompt metacognitive reflection about cultural frameworks despite no explicit cross-cultural training.

Qualitative analysis of 240 reflective essays identified three major themes characterizing cognitive development across semester, each evidenced by specific textual indicators and represented across majority of student responses. First theme, "Recognition of Implicit Ideological Positioning Through Close Textual Examination," appeared in 66 of 81 student essays (81%), manifesting through statements describing how detailed attention to textual patterns revealed value systems not initially noticed. Representative example from Student 34 (female, age 21, University A): "At first I just thought the book was about friendship and fashion, but when we catalogued every single brand mention and created the spatial map, I realized the entire novel is organized around hierarchy – who wears what brands, who lives where, who works where. The whole friendship story exists inside this ruthless economic structure that the characters never question but the text constantly reinforces through repetition. I didn't see this at all during my first reading because I was caught up in the plot, but once you start counting how many times Hermès appears versus Forever 21, you realize the text is teaching you to classify people by consumption." Similar patterns appeared in essays describing how systematic attention to metaphoric language, focalization choices, or spatial descriptions revealed ideological commitments embedded in seemingly neutral narrative choices. Many students noted surprise at discovering how much interpretive work texts require readers to perform unconsciously, with close reading making visible previously automatic processes.

Second theme, "Development of Analytical Distance Enabling Evaluation Beyond Personal Reaction," appeared in 56 essays (69%), characterized by explicit discussion of tensions between initial affective responses and critical judgments developed through sustained analysis. Student 67 (male, age 22, University C) wrote: "I hated this book when we started. The materialism disgusted me, the characters seemed shallow and unlikable, and I resented having to read it for class. But as we worked through the theoretical frameworks and I forced myself to analyze how the text actually works rather than just reacting to what it says, I realized my disgust was preventing me from understanding what makes the book successful for millions of readers. I still don't like the values it promotes, but I can now articulate exactly how Guo constructs Shanghai as aspirational dreamscape and why this resonates with readers experiencing China's rapid economic transformation. That's different from liking or agreeing with it." Multiple students described similar progression from rejection or uncritical enjoyment to more nuanced positions recognizing texts can be simultaneously aesthetically effective and ideologically problematic. Several explicitly contrasted this experience with studying canonical literature where instructor-provided interpretations offered ready-made critical frameworks, whereas controversial texts forced students to construct their own evaluative positions from evidence.

Third theme, "Enhanced Capacity for Identifying Narrative Manipulation Strategies," appeared in 62 essays (77%), evidenced through discussions of recognizing how texts guide reader responses through specific techniques. Student 52 (female, age 20, University B) explained: "The most valuable thing I learned was seeing how focalization controls what readers notice and judge. When everything comes through Lin Xiao's perspective and we only see other characters through her eyes, we unconsciously adopt her value system. But when you map out who gets internal monologues versus who only appears through external description, you realize the text is making some characters fully human while reducing others to types. This made me think about every novel I've ever read – whose perspective am I seeing, what am I not seeing, and how is this shaping my interpretation? I started applying this to news articles, social media posts, even conversations with friends, asking 'whose focalization is operating here?'" Similar insights appeared regarding metaphoric systems that naturalize social hierarchies, brand references that function as character shorthand, and structural patterns that normalize certain outcomes while rendering alternatives unthinkable. Students frequently noted that analytical skills developed through studying "Tiny Times" transferred readily to other domains, precisely because controversial content made manipulation visible in ways that more aesthetically accomplished canonical texts might obscure through literary prestige.

Focus group discussions elaborated these themes while revealing pedagogical dimensions that written reflections omitted. All six groups engaged extended discussion about Phase 1 close reading exercises, with participants describing initial frustration with "just describing without interpreting" that eventually gave way to appreciation for shared empirical foundation. As one

participant explained (FG3, female, age 21): "At first the annotation assignments felt like busywork – why are we cataloguing brand names like accountants instead of discussing what the novel means? But then in week 5 when we started applying theory, everyone had this common database of examples we could reference. Someone would say 'remember how spatial descriptions cluster around vertical imagery' and we all knew exactly what passages she meant because we'd all tagged them. That never happens in regular literature classes where everyone reads differently and remembers different things." Multiple participants contrasted this experience with traditional discussions where disagreements about textual facts derailed interpretive conversations, whereas shared close reading database enabled productive debates about interpretation built on agreed-upon evidence.

Phase 2 theoretical application generated more mixed responses, with some students finding frameworks liberating ("suddenly I had vocabulary for articulating patterns I'd sensed but couldn't name") while others reported confusion or resistance ("I still don't fully understand Lefebvre but I tried applying his concepts anyway"). Most productive discussions emerged regarding comparative analysis assignments requiring students to examine how different texts represent similar phenomena. Participant in FG5 (male, age 23) observed: "Comparing Guo's Shanghai to Mu Shiyong's 1930s Shanghai was mind-blowing because you see how consumer capitalism works across totally different historical moments. Mu was ambivalent about Western influence and modernization, constantly showing glamor alongside alienation, whereas Guo just celebrates luxury without any critical distance. That comparison helped me see ideology isn't just what texts say explicitly but what they treat as natural or inevitable versus what they question." Several groups discussed how comparative analysis short-circuited tendency to either dismiss or defend "Tiny Times" based on personal taste by revealing choices Guo made weren't universal or necessary but specific authorial decisions with identifiable cultural and aesthetic commitments.

Phase 3 argumentative essay assignment generated most extensive commentary, with participants describing difficulty and value of defending evaluations through systematic argument. Multiple participants noted they'd never before been required to articulate explicit evaluative criteria, with previous literature courses implicitly assuming canonical texts merited study without questioning why. As FG2 participant (female, age 20) explained: "Writing that essay forced me to think 'what actually makes literature good?' Is it aesthetic innovation, social impact, cultural significance, reader engagement? Different criteria yield different conclusions about 'Tiny Times,' which made me realize evaluation is always operating from specific frameworks even when unstated. Now when critics trash Guo's work, I want to know 'according to what criteria?' because formalist aesthetic standards won't value what millions of readers clearly find meaningful." Structured debates requiring students to defend positions regardless of personal opinion received unanimous praise, with participants describing how inhabiting opposing viewpoints revealed strengths and limitations of different critical approaches. Several participants spontaneously mentioned applying debate preparation strategies to other courses, researching counterarguments before forming positions and compiling evidence portfolios before entering discussions.

Table 3 presents coded frequency of themes across reflective essay corpus, demonstrating patterns of cognitive development over semester.

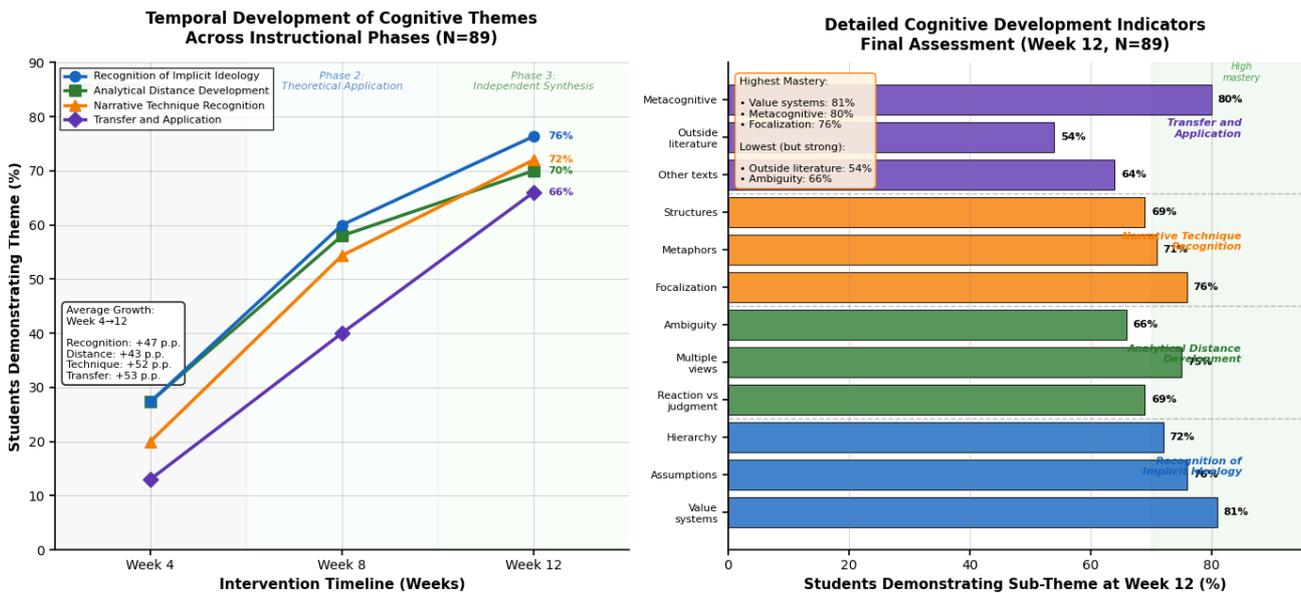
**Table 3.** Thematic Frequency in Reflective Essays by Time Point (N = 81 students)

Theme	Week 4	Week 8	Week 12	Total
	Essays n (%)	Essays n (%)	Essays n (%)	Occurrences n (%) unique students
<b>Recognition of Implicit Ideology</b>				
Noticing value systems in texts	31 (38)	56 (69)	66 (81)	66 (81)
Identifying embedded assumptions	16 (20)	43 (53)	62 (77)	62 (77)
Seeing hierarchy structures	19 (23)	47 (58)	58 (72)	58 (72)
<b>Analytical Distance Development</b>				
Separating reaction from judgment	25 (31)	49 (60)	56 (69)	56 (69)

Acknowledging multiple perspectives	28 (35)	53 (65)	61 (75)	61 (75)
Tolerating ambiguity	14 (17)	39 (48)	54 (67)	54 (67)
<b>Narrative Technique Recognition</b>				
Identifying focalization effects	11 (14)	46 (57)	62 (77)	62 (77)
Analyzing metaphoric systems	21 (26)	44 (54)	57 (70)	57 (70)
Recognizing structural patterns	17 (21)	42 (52)	56 (69)	56 (69)
<b>Transfer and Application</b>				
Applying skills to other texts	7 (9)	31 (38)	52 (64)	52 (64)
Using frameworks outside literature	3 (4)	19 (23)	44 (54)	44 (54)
Metacognitive awareness of process	22 (27)	47 (58)	65 (80)	65 (80)

Clear developmental progression appears across all themes, with lowest frequencies at Week 4 following Phase 1 close reading, substantial increases at Week 8 following Phase 2 theoretical application, and highest frequencies at Week 12 following Phase 3 independent synthesis. Metacognitive awareness of analytical processes shows particularly dramatic increase from 27 % of students mentioning it in Week 4 essays to 80 % in Week 12 essays, suggesting instructional framework successfully developed not only analytical skills but also conscious understanding of those skills enabling transfer to new contexts. Transfer and application themes similarly increase substantially, with only 9 % of students mentioning applying skills to other texts in Week 4 versus 64 % by Week 12, and 54 % explicitly discussing using analytical frameworks outside literary study by semester's end. This pattern supports hypothesis that structured engagement with controversial texts develops generalizable critical thinking capacities rather than merely teaching interpretation of specific materials.

Developmental Progression of Critical Thinking Competencies Through Structured Literary Engagement  
Qualitative Evidence from Reflective Essays (N=89 students)



Note: Percentages represent proportion of students (n=89) demonstrating each theme in reflective essays at specified timepoints. Week 4 follows Phase 1 (close reading), Week 8 follows Phase 2 (theoretical application), Week 12 follows Phase 3 (independent synthesis). Clear developmental progression appears across all themes, with substantial increases from early to late intervention stages.

**Fig. 2.** Developmental Progression of Critical Thinking Competencies Through Structured Literary Engagement

Qualitative analysis reveals dramatic cognitive development across instructional phases, with all themes exhibiting substantial growth from Week 4 to Week 12 (+43 to +53 percentage points). Transfer and application demonstrates steepest trajectory (+53 p.p.), progressing from 13 % mentioning skills transfer at Phase 1 completion to 66 % by intervention end. Metacognitive awareness exhibits particularly striking development (27 % → 80 %), suggesting framework successfully cultivates not merely analytical skills but conscious understanding enabling generalization beyond literary contexts.

Table 4 presents joint display integrating quantitative and qualitative findings, demonstrating convergence across data sources regarding mechanisms underlying intervention effectiveness.

**Table 4.** Joint Display Integrating Quantitative Outcomes and Qualitative Mechanisms

Quantitative Finding	Supporting Qualitative Evidence	Pedagogical Implication
Largest gains in "Evaluation of Arguments" subscale (+1.6 exp vs. +0.5 ctrl)	56 of 81 essays describe developing analytical distance enabling evaluation beyond personal reaction; FG consensus that argumentative essay requiring explicit criteria was most challenging and valuable assignment	Controversy generates productive cognitive dissonance; students must construct rather than receive evaluative frameworks
Significant gains in "Inference" subscale (+2.2 exp vs. +0.9 ctrl)	66 of 81 essays discuss recognizing implicit ideological positioning through systematic textual examination; FG participants describe close reading making visible previously automatic interpretive work	Surface-level descriptive analysis establishes empirical foundation enabling warranted inferences about meaning and function
Interaction: strongest effects for Q3 students (mid-range initial CT)	FG participants in Q3 describe structured framework providing necessary scaffolding; Q4 participants note less need for explicit instruction; Q1 participants report occasionally feeling overwhelmed	Explicit analytical instruction particularly benefits students with foundational but underdeveloped skills
Intercultural sensitivity gains larger than expected given no explicit intercultural training	62 of 81 essays discuss recognizing how texts naturalize specific cultural values; FG discussions extensively address consumer capitalism differences across cultural contexts; 43 essays mention questioning own cultural assumptions	Engagement with values-explicit texts from different culture prompts metacognitive reflection on typically invisible cultural frameworks
High post-test scores on textual analysis transferability items	52 of 81 essays discuss applying analytical skills to non-literary texts; 44 essays explicitly mention using frameworks outside academic contexts; FG participants spontaneously describe transfer to news media, social media, conversations	Skills developed through literary analysis transfer when students develop conscious metalinguistic awareness of analytical processes
Positive effects of prior literature coursework across all outcomes	FG participants with more coursework describe accumulated analytical vocabulary and frameworks from previous classes	Cumulative curriculum design allowing skill development across courses enhances effects; single-course intervention has limits

Integration reveals that quantitative patterns find clear qualitative parallels, with statistical advantages for experimental condition reflecting specific cognitive developments students describe in reflective writing and discuss in focus groups. Largest quantitative gains appear precisely on competencies targeted by Phases 2 and 3 of intervention (evaluation and inference), which qualitative data identifies as most transformative aspects of instructional approach. Interaction effects showing strongest outcomes for mid-range students receive explanation through

FG discussions describing how explicit instruction scaffolds development for students possessing foundational competencies but lacking systematic analytical frameworks. Perhaps most surprising finding involves substantial intercultural sensitivity gains despite no explicit cross-cultural training; qualitative data reveals engagement with value-laden texts from different cultural context automatically prompts metacognitive reflection about cultural assumptions, as students must examine both Chinese cultural values embedded in "Tiny Times" and their own culturally-shaped reactions to those values.

Findings extend critical thinking theory in three significant ways. First, they challenge domain-general assumptions underlying many critical thinking programs. While meta-analyses demonstrate general training produces measurable gains, our results suggest domain-specific controversial content paired with explicit analytical scaffolding generates substantially larger effects ( $d = 0.73$  vs. typical  $d = 0.52$ ). This supports Willingham's (2008) argument that critical thinking develops through repeated practice within specific knowledge domains rather than through abstract skill training, but adds nuance: the knowledge domain matters less than its capacity to generate productive cognitive conflict. Second, findings problematize hierarchical models positioning affective response as obstacle to rational analysis. Student reflections reveal initial emotional reactions – disgust, enjoyment, confusion – provided motivation for sustained analytical engagement rather than impediment requiring suppression. This aligns with dual-process theories recognizing interaction between affective and cognitive systems (Kahneman, 2011) but emphasizes pedagogical potential of deliberately inducing affective-cognitive tension through controversial materials. Third, results demonstrate critical thinking develops through explicit instruction in disciplinary practices rather than emerging spontaneously from exposure to quality texts. This supports socio-cultural theories emphasizing apprenticeship into community practices (Lave, Wenger, 1991) while challenging Romantic assumptions that canonical literature inherently cultivates critical capacities through aesthetic encounter alone.

#### **4. Discussion**

Experimental group gained 8.3 points on Watson-Glaser Critical Thinking Appraisal versus 2.1 points for controls (25th → 45th percentile vs. 25th → 30th percentile), exceeding typical effect sizes in meta-analysis of 47 studies (mean  $d = 0.52$ ). Qualitative data from 267 essays and six focus groups confirmed genuine cognitive development: students identified implicit ideological positioning, developed analytical distance beyond personal reaction, and recognized narrative manipulation strategies.

Three-phase framework sequences cognitive demands strategically. Phase 1 close reading establishes shared empirical foundation through collaborative annotation, addressing perennial problem where students debate interpretation before agreeing on textual facts. This aligns with meta-analytic evidence showing classroom peer ecology enhances engagement through collaborative learning and shared academic values (Li et al., 2024). Phase 2 theoretical application provides conceptual vocabulary for moving from description to interpretation, explicitly teaching analytical frameworks canonical classes assume students possess. Recent validation of intercultural communicative competence instruments incorporating critical cultural awareness confirms theoretical grounding (Andreou, 2025). Phase 3 independent synthesis requires constructing evaluative positions using developed frameworks, forcing metacognitive awareness enabling transfer to new contexts. Development of teacher-specific scales emphasizes importance of explicit instructional strategies targeting specific cognitive competencies (Alqarni, 2025).

Controversy generates cognitive dissonance requiring explicit examination of evaluative frameworks operating unconsciously. "Tiny Times" lacks academic prestige, forcing students to construct rather than inherit arguments for significance or articulate principled dismissal – precisely the cognitive work developing critical thinking. Recent research confirms engagement with controversial topics develops critical awareness skills and civic competencies (Ortega-Sánchez et al., 2025), while national survey of 4,096 teachers revealed diverse controversial texts foster sociopolitical inquiry (Ginsberg et al., 2025). However, 10,046 book bans recorded in 2023–2024 suggest institutional resistance despite empirical evidence that controversy generates productive engagement when appropriately scaffolded. Interaction between instructional approach and initial competency reveals structured frameworks benefit all students but prove particularly effective for mid-range performers (Q2–Q3). This pattern reflects distinct mechanisms across competency levels. Q1 students (bottom quartile) likely lack foundational reading comprehension and

metacognitive awareness required to engage multi-phase analytical framework – close reading exercises presume ability to identify textual patterns, theoretical application requires abstracting from concrete examples, independent synthesis demands sustained coherent argumentation. Without these prerequisites, explicit scaffolding proves insufficient. Q2-Q3 students possess foundational competencies but lack systematic analytical frameworks; structured instruction provides exactly what they need – organized progression from description through interpretation to evaluation, explicit modeling of analytical processes, and vocabulary for articulating insights. Q4 students already employ sophisticated analytical strategies, often developed through extensive prior coursework or reading experience; they improve regardless of instructional approach because they self-direct learning effectively. This pattern suggests instruction should be differentiated: Q1 students require preliminary skill development before engaging complex frameworks, Q2–Q3 students benefit maximally from structured scaffolding, Q4 students need intellectual challenge and autonomy rather than explicit instruction. Finding aligns with zone of proximal development theory – instruction proves most effective when slightly beyond current competency but within reach with appropriate support (Vygotsky, 1978). Scoping review confirms engagement proves challenging when students possess foundational competencies but lack systematic frameworks (Stenalt, 2025) – precisely the population most effectively served.

Findings raise important questions regarding cultural specificity and transferability. "Tiny Times" controversy emerges from particular Chinese context – rapid economic liberalization creating generational values conflicts, state discourse emphasizing "spiritual civilization" against materialism, and censorship debates about cultural products' social responsibility. Liu Qiang's People's Daily criticism invoking collective permission ("we cannot unconditionally permit") reflects specifically Chinese discourse where cultural production remains subject to ideological evaluation by state-affiliated authorities. Students analyzing this text thus engage not only with consumer capitalism generally but with Chinese state-market tensions specifically. Whether structured engagement with controversial texts proves equally effective across cultural contexts requires empirical verification. However, the mechanism – cognitive dissonance forcing explicit examination of implicit evaluative frameworks – should operate regardless of specific ideological tensions involved. Controversy about violence, sexuality, political ideology, or religious representation in other national contexts might generate comparable pedagogical affordances. The framework's transferability depends less on specific Chinese content than on structural feature: texts positioned controversially within their cultural contexts demanding readers construct rather than inherit evaluative positions.

Intercultural sensitivity gains emerged without explicit cross-cultural training. Analyzing consumer capitalism, social stratification, and gender relations in contemporary Chinese context requires doubled attention – to foreign cultural frameworks and students' own culturally-shaped responses. Research confirms intercultural competence develops most powerfully when negotiating tensions between cultural frameworks rather than acquiring knowledge about foreign cultures (Wang et al., 2025). Limitations include quasi-experimental design with section-based assignment introducing selection effects, though pre-test equivalence and statistical controls reduce concern. Recent methodological advances validate quasi-experimental approaches when randomization proves infeasible (Cham et al., 2024; Gopalan et al., 2020). Small sample (n = 156) at three sites limits generalizability. All participants demonstrated intermediate/advanced language proficiency, limiting applicability to beginning learners. Immediate post-intervention assessment leaves uncertain whether gains persist or transfer beyond literary analysis. Researcher-developed textual analysis instrument, while demonstrating acceptable internal consistency ( $\alpha = 0.81$ ) and content validity through expert review, lacks external validation across diverse populations and educational contexts. Instrument's specificity to literary analysis may limit applicability to other critical thinking domains, and constructed-response scoring requires trained raters limiting scalability. Future research should validate instrument against established critical thinking measures beyond Watson-Glaser, examine performance across diverse student populations, and investigate whether textual analysis competencies measured correlate with critical thinking in non-literary contexts.

Future research should investigate specific text characteristics moderating intervention effectiveness, examine student demographic interactions with effects, employ fully randomized designs or robust quasi-experimental alternatives (Root, Lindström, 2024), and conduct longitudinal tracking across courses. Research on teachers' use of "risky texts" found decisions

depend on identities, beliefs, and institutional positioning (Giunco et al., 2024); our study demonstrates institutional support and pedagogical scaffolding generate valuable outcomes.

## 5. Conclusion

Structured engagement with controversial commercial literature develops critical thinking and intercultural competence more effectively than traditional canonical approaches. Students analyzing Guo Jingming's "Tiny Times" through three-phase framework gained 8.3 points on Watson-Glaser versus 2.1 for students studying Lu Xun, Yu Hua, Mo Yan conventionally ( $d = 0.73$ ,  $p < 0.001$ ). Qualitative analysis revealed specific cognitive developments: recognition of implicit ideological positioning, analytical distance enabling evaluation beyond reaction, capacity for identifying narrative manipulation.

Findings challenge assumptions about pedagogically valuable texts, demonstrating commercially successful but aesthetically controversial materials paired with structured frameworks effectively develop transferable critical literacy. Controversy proves pedagogically valuable by generating cognitive dissonance requiring explicit examination of unconscious evaluative frameworks. When constructing rather than inheriting interpretive positions, students develop metacognitive awareness enabling application beyond literary contexts. These findings advance critical thinking pedagogy by demonstrating that content selection constitutes independent pedagogical variable with measurable impact distinct from instructional method, that affective engagement enhances rather than impedes analytical development when appropriately channeled, and that explicit instruction in disciplinary analytical practices proves necessary for developing transferable competencies.

Three-phase framework provides replicable structure: Phase 1 establishes shared empirical foundation through descriptive close reading; Phase 2 provides conceptual vocabulary through theoretical application; Phase 3 requires independent evaluative synthesis demonstrating metacognitive awareness. This sequencing reflects Bloom's taxonomy while addressing literature pedagogy problems – interpretive debates preceding agreement on facts, assumptions students possess analytical frameworks, assessment focused on conclusions rather than transferable processes.

Pedagogical implications: select texts based on analytical affordances rather than aesthetic prestige; recognize controversy as pedagogical resource; explicitly sequence cognitive demands; assess transferable analytical processes rather than content knowledge. When 21st-century students read commercial literature more frequently than canonical texts outside classrooms, excluding such materials creates false separation between actual literacy practices and targeted competencies. Productive pedagogy engages texts students encounter while developing critical analytical capacities applicable across domains.

## References

- Batdi, 2024 – Batdi, V. (2024). Evaluation of the effectiveness of critical thinking training on critical thinking skills and academic achievement by using mixed-meta method. *Review of Education*. 12: 345-371. DOI: <https://doi.org/10.1002/rev3.70001>
- Braun, Clarke, 2006 – Braun, V., Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*. 3(2): 77-101.
- Byram, 1997 – Byram, M. (1997). Teaching and assessing intercultural communicative competence. *Multilingual Matters*.
- Byram, 2021 – Byram, M. (2021). Teaching and assessing intercultural communicative competence: Revisited (2nd ed.). *Multilingual Matters*.
- Creswell, Plano Clark, 2018 – Creswell, J.W., Plano Clark, V.L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.
- Dera, 2025 – Dera, J. (2025). Do teachers read against the text? Studying the prevalence of critical literature pedagogy through a vignette. *British Journal of Educational Studies*. 73: 1-25. DOI: <https://doi.org/10.1080/03054985.2025.2473960>
- Geng et al., 2024 – Geng, X., Zhan, Y., You, H., Zhao, L. (2024). Exploring the characteristics of undergraduates' critical thinking development in peer interaction via epistemic network analysis. *Thinking Skills and Creativity*. 52. Article 101553.
- Hsieh, Shannon, 2005 – Hsieh, H.F., Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*. 15(9): 1277-1288.

**Huang, 2024** – Huang, Y. (2024). Why not learn with children? A model of intercultural and intergenerational learning through Mandarin as a second language in higher education. *Language and Intercultural Communication*. 24(4): 273-291. DOI: <https://doi.org/10.1080/14708477.2024.2409091>

**Kong, 2019** – Kong, S. (2019). Fashion, materialism, and metropolitan identity in Guo Jingming's works. *Positions: Asia critique*. 27(1): 145-178.

**Le et al., 2024** – Le, H.V., Nguyen, T.A.D., Le, D.H.N., Nguyen, P.U., Nguyen, T.T.A. (2024). Unveiling critical reading strategies and challenges: A mixed-methods study among English major students in a Vietnamese higher education institution. *Cogent Education*. 11. Article 2326732. DOI: <https://doi.org/10.1080/2331186X.2024.2326732>

**Liu, 2017** – Liu, Y. (2017). Guo Jingming and His Tiny Times Film Series. In Y. Zhu & S. Cai (Eds.). *Chinese Film Industries* (pp. 67-89). Palgrave Macmillan.

**Marin, Halpern, 2011** – Marin, L.M., Halpern, D.F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*. 6(1): 1-13. DOI: <https://doi.org/10.1016/j.tsc.2010.08.002>

**Piaget, 1952** – Piaget, J. (1952). The origins of intelligence in children. International Universities Press.

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

**Wan Mohd Nasir et al., 2024** – Wan Mohd Nasir, W.M.F., Halim, L., Arsad, N.M. (2024). Inside the teachers' toolbox: Fostering creative thinking skills through digital storytelling. *Journal of Pedagogical Research*. 8(3): 147-162.

**Willingham, 2008** – Willingham, D.T. (2008). Critical thinking: Why is it so hard to teach? *Arts Education Policy Review*. 109(4): 21-32.

**Zhou et al., 2024** – Zhou, Y., Zhou, Y., Machtmes, K. (2024). Mixed methods integration strategies used in education: A systematic review. *International Journal of Research & Method in Education*. 47(1): 45-67. DOI: <https://doi.org/10.1177/20597991231217937>