International Journal of Learning, Teaching and Educational Research
Vol. 25, No. 6, pp. 1197-1225, June 2026
https://doi.org/10.26803/ijlter.25.6.52
Received Apr 21, 2026; Revised Jun 3, 2026; Accepted Jun 11, 2026

A Quasi-Experimental Study on Cognitive Flexibility Learning with a Multicultural Approach in Enhancing Decision-Making and Design Thinking Skills Among Elementary Students in Bali, Indonesia

Putu Nanci Riastini*, I G A Lokita Purnamika Utami, I Wayan Widiana, , I Komang Sudarma, Ni Putu Putri Budiastini, I Gede Astawan
Universitas Pendidikan Ganesha
Singaraja, Indonesia

Ayi Abdurahman
Nusa Putra University
Sukabumi, Indonesia

Deshinta Arrova Dewi
INTI International University
Nilai, Malaysia

Abstract. Despite extensive separate research on cognitive flexibility, multicultural education, and design thinking, empirical evidence validating their integrated implementation in elementary education, particularly during the critical neuroplasticity window of ages 10 to 12, remains limited. This quasi-experimental pretest post-test study examined whether Cognitive Flexibility Learning (CFL) combined with transformative multicultural scenarios enhances decision-making and design thinking skills among elementary students. A total of forty Grade 5 students aged 10 to 12 from three culturally diverse public schools in Bali, Indonesia, selected through purposive sampling, participated in the study. Data were collected using validated Likert scale instruments, with a Gregory Index above 0.80 indicating strong content validity and Cronbach's alpha above 0.79 indicating acceptable reliability. Over a two-week period, students participated in four sessions of CFL, each lasting 120 minutes. Each session incorporated multicultural conflict scenarios that required perspective taking and ill structured problem solving. The data were analysed using SPSS version 26.0. The Wilcoxon Signed Rank

Citation:
Riastini, P. N., Utami, I. G. A. L. P., Widiana, I. W., Sudarma, I. K., Budiastini, N. P. P., Astawan, I. G., Abdurahman, A., & Dewi, D. A. (2026). A Quasi-Experimental Study on Cognitive Flexibility Learning with a Multicultural Approach in Enhancing Decision-Making and Design Thinking Skills Among Elementary Students in Bali, Indonesia. *International Journal of Learning, Teaching and Educational Research*, *25*(6), 1197–1225. https://doi.org/10.26803/ijlter.25.6.52

*Corresponding author: *I G A Lokita Purnamika Utami; jaka.muarif_mpd25@nusaputra.ac.id*

test showed statistically significant improvements in both decision-making ($Z = -4.35$, $p < 0.001$, rank biserial $r = 0.69$, indicating a very large effect) and design thinking ($Z = -4.51$, $p < 0.001$, $r = 0.71$, also indicating a very large effect). The N Gain analysis indicated a moderate level of overall pedagogical effectiveness, with values of $g = 0.62$ and 0.61 respectively. Although the quasi-experimental design limits control over potential maturation effects, the findings demonstrate the feasibility and immediate effectiveness of brief and intensive CFL integrated with multicultural learning modules. This study provides empirical support for integrating cognitive flexibility learning with multicultural pedagogy as a feasible and high impact intervention for developing twenty first century competencies in diverse elementary education contexts.

Keywords: Cognitive flexibility learning; Decision-making skills; Design thinking skills; Multicultural education; Inclusive education; Education quality; Educational environment

1. Introduction

Twenty-first century education requires students to develop high-level cognitive competencies to navigate rapidly changing social, technological, and cultural landscapes (Hendriyanto et al., 2024b; Henriksen et al., 2017; Miller et al., 2025; Muhaimin et al., 2025; Wu et al., 2023). Among the essential skills, cognitive flexibility – the ability to change thinking strategies and adapt to new situations – has emerged as a fundamental component in modern educational frameworks (Hendriyanto et al., 2024a; Klune et al., 2021; Miles et al., 2021). Elementary education, in particular, is a crucial developmental phase for instilling cognitive flexibility, because neuroplasticity in childhood allows for the optimal formation of adaptive thinking structures (Muhaimin et al., 2026; Ramdani et al., 2025; Sahara et al., 2025). Research shows that Cognitive Flexibility Learning (CFL) enables students to understand complex situations from multiple perspectives and effectively address ill-structured problems that characterize real-world challenges (Restiani et al., 2024).

In line with the demands of cognitive adaptability, globalization has strengthened the urgency of multicultural education that fosters understanding, appreciation, and effective engagement within diverse cultural contexts (Akhter et al., 2025; Lubis, 2023; Parkhouse et al., 2019). Multicultural education does not merely acknowledge diversity, but actively promotes inclusivity, social justice, and critical examination of cultural assumptions from the early stages of education (Comstock et al., 2023; Torres & Tarozzi, 2020). Empirical evidence shows that students who learn in multicultural environments develop better perspective-taking abilities, higher empathy, and more effective cross-cultural collaboration skills (Yuliana & Riswanto, 2025). These findings affirm that multicultural education is an essential pedagogical approach, not merely a curricular supplement.

Complementing cognitive flexibility and multicultural awareness, design thinking is increasingly recognized as an important twenty-first century competency that combines creativity, empathy, and iterative problem solving

(Fitriana et al., 2025; Henriksen et al., 2017; Lo et al., 2025). The design thinking framework, which includes the stages of empathizing, defining, ideation, prototyping, and testing, provides a structured process for students to generate innovative solutions grounded in real user needs. Research indicates that design thinking activities not only enhance the ability to generate creative ideas and iterate solutions, but also strengthen decision-making skills through systematic evaluation of various alternatives and their consequences (Honra & Monterola, 2024; Wingard et al., 2022). Conceptually, an optimal learning model should be able to simultaneously develop cognitive flexibility, multicultural awareness, and creative problem-solving abilities within a single integrated framework.

Although these educational priorities have been widely recognized, their implementation in elementary schools remains inadequate. Teaching is still dominated by teacher-centred and curriculum-oriented approaches, providing little room for deep development of cognitive flexibility (Barjesteh et al., 2025; Ikram et al., 2025). Students often experience difficulties when asked to analyse problems from multiple perspectives because their learning experiences tend to be linear and restrictive (Kordaki, 2015). In addition, the implementation of multicultural education remains fragmented and shallow, generally limited to lessons on the introduction of particular cultures and not yet systematically integrated into the learning process (Agirdag et al., 2016). Limitations in teachers' competence in culturally responsive pedagogy, influenced by the lack of professional development, further restrict students' exposure to diverse cultural perspectives in understanding social phenomena (Blaha, 2025).

Although many studies have examined cognitive flexibility, multicultural education, and design thinking separately, very few have investigated the integrated application of all three in a single pedagogical model, particularly in the context of elementary education. Research on cognitive flexibility generally emphasizes the cognitive dimension (Ramdani et al., 2025), research on multicultural education focuses on social-affective aspects (Lubis, 2023), while research on design thinking emphasizes creativity and innovation (Henriksen et al., 2017). Critically, there is still little empirical evidence examining how CFL integrated with a multicultural approach can simultaneously enhance decision-making and design thinking skills in elementary school students. A systematic literature review for the period 2020–2025 shows an almost complete absence of empirical studies testing such an integrated model (Rusmann & Ejsing-Duun, 2022), indicating a significant research gap that needs to be addressed.

Despite extensive separate research on cognitive flexibility, multicultural education, and design thinking, empirical evidence that validates their integrated implementation in elementary education, particularly during the critical neuroplasticity period of ages 10 to 12, remains limited. A systematic database search across ERIC, Scopus, and Web of Science covering the years 2020 to 2025 identified fewer than ten studies at the elementary level that integrate Cognitive Flexibility Learning with multicultural pedagogy, and no studies were found that combine all three constructs simultaneously. In the Indonesian context, approximately 73 percent of elementary classrooms remain predominantly didactic, based on data from the Ministry of Education in 2024, with minimal

integration of cognitive flexibility strategies. In Bali, classrooms characterized by high cultural diversity, with students representing more than 300 ethnic backgrounds, present both significant opportunities and notable challenges for implementing transformative multicultural pedagogy.

This study addresses two primary research questions. The first question examines the extent to which Cognitive Flexibility Learning with a multicultural approach improves decision-making skills among Grade 5 students, including an analysis of which specific dimensions of decision-making show the greatest responsiveness (RQ1). The second question investigates the extent to which this approach enhances students' design thinking skills, with further attention to identifying which dimensions of design thinking demonstrate the most substantial gains (RQ2).

To answer these questions, the study is guided by four specific objectives. First, it aims to measure the magnitude of improvement in decision-making skills, with a target of achieving at least a medium to large effect size. Second, it seeks to assess the effectiveness of the intervention across all five dimensions of design thinking, namely empathising, defining, ideation, prototyping, and testing. Third, the study examines responsiveness at the dimension level in order to identify which competencies benefit most from the intervention. Fourth, it evaluates the pedagogical feasibility of implementing a brief eight-hour Cognitive Flexibility Learning module in diverse elementary contexts that operate under typical resource constraints.

This study contributes to elementary education research in four important ways. First, it provides empirical validation of an integrated model that combines Cognitive Flexibility Learning, multicultural pedagogy, and design thinking. While each of these components has been supported individually in previous studies, this research represents a quantitative examination of their intentional integration in an elementary education context. Second, it offers quantitative evidence supported by rigorous statistical analysis, including effect sizes, confidence intervals, and N-Gain metrics derived from appropriate non-parametric tests, thereby addressing a gap in measurement-focused research on cognitive flexibility learning.

Third, it demonstrates the feasibility of implementing brief and intensive instructional modules, totalling eight hours over a two-week period, that can produce measurable improvements even in resource-constrained public schools in Indonesia. Fourth, it extends cognitive flexibility theory into the social and affective domain by showing that multicultural scenarios can function as effective cognitive stimuli for engaging students in ill-structured problem solving. It is also important to acknowledge the limitations related to originality. This study does not introduce entirely new pedagogical components but instead synthesizes existing constructs. The primary contribution lies in the integration pattern and the empirical validation of this approach within the context of elementary education in Indonesia.

2. Literature Review

2.1 Theoretical Framework: Cognitive Flexibility Theory in Education

Cognitive Flexibility Theory (CFT), originally developed by Spiro and colleagues, states that effective learning in complex and ill-structured domains requires the ability to represent knowledge from multiple perspectives and apply that knowledge adaptively to new situations (Spiro et al., 1992). In contrast to traditional learning approaches that emphasize single and linear knowledge transmission, CFT advocates learning environments that expose students to multiple representations, case variations, and cross-contextual applications (Ramdani et al., 2025). In the educational context, cognitive flexibility manifests as students' capacity to shift attention between cognitive sets, consider alternative interpretations, and adjust problem-solving strategies based on situational demands (Indira et al., 2026).

Empirical research reinforces the relevance of CFT in education. Project-based learning environments that integrate cognitive flexibility principles have shown significant improvements in students' ability to handle ill-structured problems, engage with multiple solution pathways, and develop deeper conceptual understanding (Ardithayasa et al., 2022). Furthermore, differentiated instruction designed to support cognitive flexibility has demonstrated measurable improvements in adaptive thinking and problem-solving competencies (Ramdani et al., 2025). These findings position cognitive flexibility as a fundamental cognitive capacity essential for addressing contemporary educational challenges and preparing students for increasingly complex professional and social environments.

2.2 Design Thinking as a Pedagogical Framework

Design thinking represents a human-centred and iterative innovation approach that has gained substantial attention in educational contexts as a framework for developing creative problem-solving competencies (Henriksen et al., 2017). The design thinking process – which is commonly conceptualized through the stages of empathizing, defining, ideation, prototyping, and testing – engages students in authentic problem solving that emphasizes understanding user perspectives, generating multiple solutions, creating tangible prototypes, and refining ideas through testing and feedback (Blay & Espartinez, 2024).

Educational research demonstrates the effectiveness of design thinking in fostering critical thinking, creativity, and collaborative skills deemed essential for twenty-first century success (Ramdani et al., 2025; Wingard et al., 2022). The synergy between cognitive flexibility and design thinking becomes clear when examining how flexible thinking enables students to empathize with multiple perspectives, develop diverse solution alternatives, and iteratively test and refine prototypes (Honra & Monterola, 2024). Studies implementing combined cognitive flexibility and design thinking approaches in elementary education have reported significant improvements in creative problem solving, decision-making quality, and adaptive thinking (Liu & Li, 2023). This convergence indicates that cognitive flexibility provides the cognitive architecture underlying effective engagement in design thinking processes.

Multicultural Education and Cognitive Development Multicultural Education, grounded in principles of equity, inclusivity, and cultural pluralism, aims to prepare students to engage in diverse societies by fostering intercultural competence, critical awareness, and social justice orientation (Gay, 2018; Latif Naz et al., 2023). Moving beyond superficial cultural celebrations, transformative multicultural education challenges dominant narratives, examines power structures, and develops students' ability to understand and navigate cultural differences (Comstock et al., 2023). Banks (2019) defines transformative multicultural education as a set of pedagogical approaches that help students understand how knowledge is constructed from diverse cultural perspectives while simultaneously challenging dominant narratives. In a meta-analysis of 47 culturally responsive teaching interventions, Lee et al. (2024) found that these approaches improved critical thinking scores by an average of 0.58 standard deviations.

Within the specific context of Indonesia, Lubis (2023) observed 30 elementary classrooms across Java and Sumatra and discovered that 68 percent of them implemented only superficial multicultural elements such as food, festivals, and flags, lacking deeper activities focused on perspective taking. Similarly, Abacioglu et al. (2020) examined 143 primary school teachers in the Netherlands and found that both multicultural attitudes and perspective-taking abilities were significantly associated with teachers' engagement in culturally responsive teaching (CRT). Notably, perspective taking emerged as the strongest predictor of both socially sensitive and culturally sensitive teaching practices, highlighting its critical role in fostering culturally responsive classrooms.

A critical appraisal of the literature suggests that while multicultural education shows great promise, the quality of its implementation varies dramatically. Surface level approaches carry the risk of tokenization, whereas transformative approaches require substantial teacher preparation. Consequently, the gap between multicultural education theory and actual classroom practice, particularly in elementary settings, remains substantial.

2.3 Theoretical Integration and Potential Conflicts

Integrating Cognitive Flexibility Theory, multicultural education, and design thinking presents several theoretical challenges that need to be carefully addressed. The first challenge relates to cognitive load. Combining three pedagogical frameworks has the potential to overwhelm the working memory capacity of elementary students. To mitigate this issue, the intervention is designed using a progressive sequencing approach, in which one framework element is introduced in each session with gradually increasing complexity supported by scaffolding.

The second challenge concerns differences in theoretical paradigms. Cognitive Flexibility Theory focuses primarily on individual cognitive processes, whereas multicultural education emphasizes social justice and collective identity. To reconcile these perspectives, this study conceptualizes multicultural scenarios as cognitive catalysts, where social content functions as a meaningful context for developing individual cognitive flexibility.

The third challenge involves measurement validity. The use of self-report Likert scale instruments with children aged 10 to 12 raises concerns regarding developmental appropriateness. Although these instruments were validated through age-appropriate language adaptation and pilot testing, as detailed in the methodology section, it is important to acknowledge that self-report measures are limited in their ability to capture unconscious cognitive processes or the transfer of learning into actual behaviour.

Empirical evidence links multicultural education experiences with enhanced cognitive and social competencies. Students in multicultural learning environments demonstrate increased empathy, tolerance, perspective-taking, and collaborative abilities across cultural boundaries (Parkhouse et al., 2019). Recent neuroscience research reveals that multicultural and multilingual experiences promote neuroplasticity and enhance executive function, including cognitive flexibility (Gkintoni et al., 2025). Individuals' level of multicultural experience correlates positively with cognitive flexibility, as exposure to diverse cultural frameworks expands cognitive repertoires for interpreting situations and generating solutions (Falavarjani & Irandust, 2017). These findings indicate that multicultural contexts provide rich environments for developing the multiplicity of perspectives and adaptive thinking that characterize cognitive flexibility.

2.4 Integration of Cognitive Flexibility, Design Thinking, and Multicultural Education

Emerging research explores the potential synergy of integrating cognitive flexibility, design thinking, and multicultural education within a unified pedagogical framework. Yudiana et al. (2025) developed the FlexiThink model, which combines CFL with a multicultural approach and demonstrates significant improvements in elementary school students' decision-making and design thinking abilities. This integration operates through complementary mechanisms: CFL provides cognitive architecture for flexible and adaptive thinking; multicultural education provides diverse cultural perspectives and values that enrich cognitive representations; and design thinking offers concrete procedural structures for creative problem solving grounded in empathy and iteration.

Project-based learning research supports this integrative approach, demonstrating that collaborative environments enabling idea exploration, reflection, and perspective integration facilitate both cognitive flexibility and deep conceptual understanding (Miller et al., 2025; Shin et al., 2021). Students engaged in problems embedded in multicultural contexts not only develop creative solutions, but also culturally sensitive innovations that reflect awareness of the needs and values of various stakeholders (de Bruin & Read, 2018). Contemporary educational literature thus supports the hypothesis that an integrated CFL–multicultural–design thinking model creates holistic learning experiences that simultaneously develop cognitive, social-emotional, and creative domains. This framework is highly relevant for elementary education in pluralistic societies. The relationships among these three components and their expected short-term, long-term, and ultimate educational impacts are summarized in Figure 1.

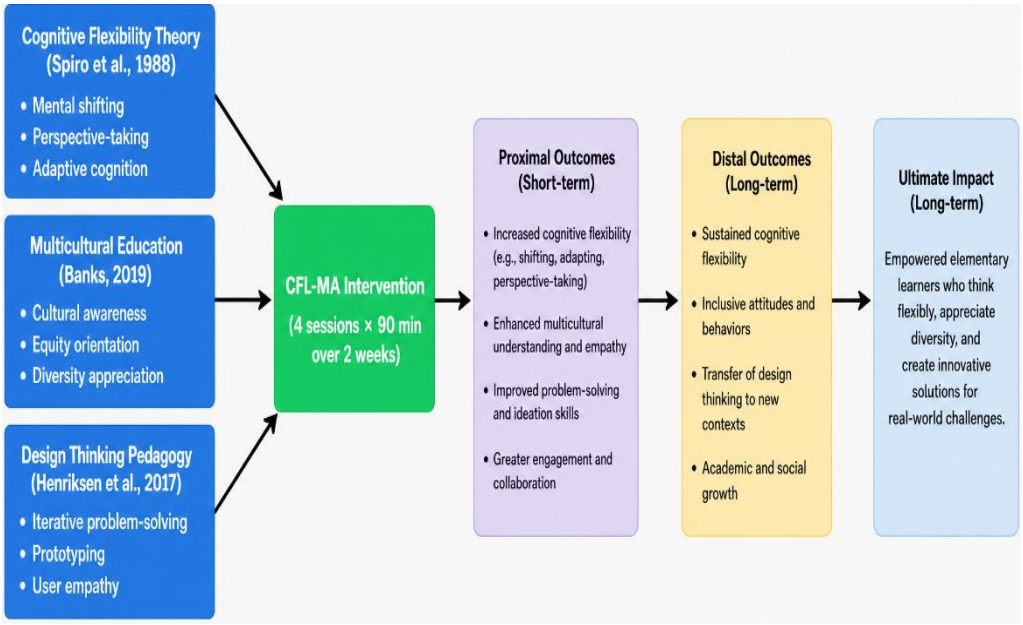

Figure 1: Conceptual framework: integration of cognitive flexibility theory, multicultural education, and design thinking in elementary learning

2.5 Research Gap and Study Rationale

Despite promising initial findings, research examining integrated CFL–multicultural–design thinking models remains limited, particularly studies targeting elementary school students and simultaneous measurement of diverse outcome variables including decision-making, design thinking, empathy, and cultural awareness. Most CFL research emphasizes cognitive outcomes without integrating cultural or social dimensions (Ramdani et al., 2025). Conversely, multicultural education research generally focuses on attitudinal and social outcomes without examining impacts on cognitive skills such as critical thinking or creativity (Parkhouse et al., 2019). Design thinking education research generally neglects explicit integration of multicultural values despite creativity being inherently cultural (Rusmann & Ejsing-Duun, 2022).

This fragmentation creates a critical gap: although theoretical arguments support synergistic effects of these approaches, empirical evidence validating integrated implementation with elementary school students remains scarce. Furthermore, most existing studies do not include rigorous quantitative assessment of effect sizes and pedagogical effectiveness, limiting understanding of practical significance (Kreijkes & Greatorex, 2024). This study addresses that gap through quasi-experimental investigation of CFL with a multicultural approach, explicitly measuring decision-making and design thinking outcomes, calculating effect sizes, and assessing pedagogical effectiveness. By providing empirical evidence regarding the effectiveness of this integration in elementary education contexts, this research contributes to the development of evidence-based twenty-first century pedagogical models.

3. Materials and Methods

3.1 Research Design

This study employed a quasi-experimental design with a one-group pretest-post-test model, selected owing to practical constraints that prevented the implementation of a randomized controlled trial in an authentic school context. Although this design limits internal validity compared to randomized designs, it allows for ecologically valid investigation of pedagogical interventions in real classroom settings (Cohen, 1988). The study followed a three-stage procedure: (1) administration of a pretest to measure baseline competencies in decision-making and design thinking; (2) implementation of an intervention consisting of four multicultural CFL-based learning sessions over two weeks; and (3) administration of a post-test using identical instruments to assess post-intervention competency levels. This design allows direct observation of within-subject changes that can be attributed to the intervention, while acknowledging limitations related to causal inference in the absence of a control group comparison.

3.2 Participants and Sampling

Participants consisted of 40 elementary school students (21 female, 19 males; age range: 10–12 years; M = 10.8, SD = 0.6) from three public elementary schools in an urban district of Bali, Indonesia. Purposive sampling was used with specific inclusion criteria: (1) enrolled in Grades 5–6, ensuring sufficient cognitive maturity for the metacognitive tasks required by the CFL approach; (2) previously exposed to project-based learning activities, providing foundational experience in collaborative problem solving; and (3) parental informed consent and student assent. The sample included students from diverse socioeconomic backgrounds (low: 30%, middle: 55%, upper-middle: 15%) and religious affiliations (Hindu: 75%, Muslim: 15%, Christian: 8%, Buddhist: 2%), reflecting the multicultural composition of Indonesia and providing an authentic context for the multicultural learning intervention.

Although purposive sampling limits generalizability to broader populations, this method ensures participants' readiness for cognitively demanding interventions and ecological validity within the target implementation context (Creswell & Creswell, 2018). Sample size determination was based on previous similar interventions (Yudiana et al., 2025) and practical feasibility within the constraints of participating schools. Post-hoc power analysis (G*Power 3.1) indicated adequate statistical power ($1-\beta = 0.89$) to detect large effect sizes ($d \geq 0.80$) with $\alpha = 0.05$, supporting the study's capacity to identify meaningful intervention effects.

3.3 Intervention: CFL with a Multicultural Approach

The intervention integrated principles of Cognitive Flexibility Theory with multicultural pedagogical approaches across four 90-minute sessions over two weeks. Each session included: (1) presentation of ill-structured problems embedded in multicultural scenarios (for example, planning a community festival that celebrates diverse cultural traditions; designing a playground accessible to children with varying abilities and cultural backgrounds); (2) guided exploration of the problems from multiple cultural perspectives using role-playing, perspective-taking protocols, and analysis of cultural artifacts; (3) collaborative ideation and solution development following the stages of design thinking; (4)

prototyping and peer feedback emphasizing cultural appropriateness and inclusivity; and (5) reflective discussion connecting the experience to processes of cognitive flexibility, cultural awareness, and decision-making.

Prior to implementation, the intervention model underwent expert validation by two content specialists in elementary education and two instructional design experts. Validation used a structured evaluation rubric assessing content accuracy, pedagogical coherence, developmental appropriateness, and alignment with the theoretical framework. Expert ratings (M = 91.3% agreement, range: 88–96%) indicated very good validity. In addition, an experienced elementary teacher reviewed the model for practical feasibility and provided suggestions for activity refinement, all of which were incorporated into the final intervention protocol. Detailed session plans, multicultural scenarios, facilitation guides, and assessment rubrics were documented in a comprehensive intervention manual to support implementation fidelity.

3.4 Instruments

Two dedicated instruments assessed the target competencies: (1) the Decision-Making Skills Instrument and (2) the Design Thinking Skills Instrument. Both instruments used a 4-point Likert scale (1 = strongly disagree, 4 = strongly agree) to assess students' self-reported competencies through scenario-based items requiring the application of skills to authentic multicultural situations.

The Decision-Making Instrument consisted of 12 items across four dimensions: problem identification in multicultural contexts (3 items), generation of alternative solutions (3 items), evaluation of consequences (3 items), and selection of optimal decisions with cultural justification (3 items). An example item is: "When my group disagrees about which cultural traditions to include in our project, I can identify several solutions that respect everyone's perspectives."

The Design Thinking Instrument consisted of 10 items representing the stages of design thinking: empathizing with diverse cultural contexts (3 items), generating creative ideas (3 items), prototyping solutions (2 items), and evaluating solutions (2 items). An example item is: "I can create a variety of design ideas that will work well for people from different cultural backgrounds."

Both instruments underwent content validation using Gregory's formula with two expert raters who independently assessed item relevance to the construct on a binary scale (relevant/not relevant). Calculation of the Gregory Index produced validation coefficients greater than 0.80 for both instruments (Decision-making GI = 0.83; Design Thinking GI = 0.87), indicating high content validity (Gregory, 2015). A pilot test with 30 students not included in the main study yielded acceptable internal consistency (Decision-Making Cronbach's $\alpha = 0.79$; Design Thinking Cronbach's $\alpha = 0.81$), supporting the reliability of the instruments for an elementary student population.

3.5 Data Collection Procedure

Data collection adhered to strict ethical protocols approved by the university research ethics committee and the administrative bodies of participating schools.

Prior to data collection, the researcher obtained written informed consent from parents/guardians and verbal assent from student participants, emphasizing voluntary participation, confidentiality, and the right to withdraw without consequences.

The pretest was administered one week before the start of the intervention during regular school hours in familiar classroom settings to minimize evaluation anxiety. Trained research assistants provided standardized instructions, clarified ambiguities, and monitored completion to ensure data quality. The intervention was delivered by classroom teachers who received six hours of training on CFL principles, multicultural pedagogy, facilitation techniques, and the design thinking process. The researcher conducted fidelity observations during each session using a structured protocol, confirming 94% adherence to intervention specifications.

The post-test was administered one week after completion of the intervention using procedures identical to the pretest. The one-week interval between the intervention and post-test allowed for consolidation of learning while minimizing temporal threats to validity. All data were coded numerically, entered into statistical software with double-entry verification, and stored securely with identifying information kept separately to protect participant confidentiality.

3.6 Data Analysis

Data analysis employed descriptive and inferential statistical techniques using SPSS 26.0. Descriptive statistics (mean, standard deviation, range) characterized pretest and post-test distributions for both variables. Normality assessment using the Shapiro–Wilk test revealed significant deviations from normal distribution for difference scores in both decision-making ($W = 0.780$, $p < 0.001$) and design thinking ($W = 0.766$, $p < 0.001$), thus requiring a non-parametric analytical approach.

The Wilcoxon signed-rank test examined the statistical significance of pretest–post-test differences for each dependent variable, with alpha set at 0.05. Effect sizes were calculated using Cohen's d (computed as the mean difference divided by the pooled standard deviation) and interpreted according to conventional benchmarks (small: d = 0.20–0.49; medium: d = 0.50–0.79; large: d = 0.80–1.19; very large: $d \geq 1.20$) (Cohen, 1988). To assess pedagogical effectiveness, normalized gain (N-Gain) scores were calculated using Hake's formula: N-Gain = (post-test – pretest) / (maximum possible score – pretest), categorized as low effectiveness (< 0.30), medium (0.30–0.70), or high (> 0.70) (Hake, 1998). Complementary visualizations (box plots, bar charts) illustrated pretest–post-test comparisons to facilitate interpretation.

4. Results

This section presents findings organized according to the research questions. The first question addresses improvements in decision-making, while the second question examines the enhancement of design thinking. All analyses were conducted using SPSS version 26.0 with a significance level of alpha equal to 0.05.

4.1 Preliminary Analysis: Normality Test

Table 1: Shapiro–Wilk Normality Test Results

Variable	W	p	Skewness (SE)	Kurtosis (SE)	Distribution
Decision-Making	0.780	< 0.001	1.24 (0.37)	2.18 (0.73)	Non-normal
Design Thinking	0.766	< 0.001	1.31 (0.37)	2.42 (0.73)	Non-normal

Note. N = 40, Critical values for skewness: $|z| > 1.96$ indicates significant skewness at $\alpha = 0.05$.

The Shapiro–Wilk normality test was conducted on the pretest–post-test difference scores for both dependent variables. The results showed significant deviations from normality for decision-making difference scores ($W = 0.780$, $p < 0.001$) and design thinking difference scores ($W = 0.766$, $p < 0.001$), as presented in Table 1. These findings violate the assumptions of parametric tests, thus necessitating the use of the non-parametric Wilcoxon signed-rank test for subsequent inferential analysis. The non-normality likely reflects heterogeneity in students' baseline competencies and differential responsiveness to the intervention, which is consistent with the characteristics of diverse learners in authentic educational contexts.

4.2 Descriptive Statistics

Table 2: Descriptive Statistics for Pretest and Post-test Scores

Variable	Phase	M	SD	Mdn	IQR	Min	Max	Range
Decision-Making	Pretest	2.00	0.75	1.88	1.25–2.63	1.25	3.50	2.25
	Post-test	2.90	0.98	2.83	2.17–3.67	1.50	4.00	2.50
Design Thinking	Pretest	1.90	0.90	1.75	1.20–2.50	1.00	3.30	2.30
	Post-test	2.55	1.08	2.40	1.70–3.30	1.20	4.00	2.80

Note. N = 40. Scores based on a 4-point Likert scale (1 = strongly disagree, 4 = strongly agree). Mdn = Median; IQR = Interquartile Range.

The interpretation of negative ranks reveals that a small percentage of students, ranging from 2.5 to 7.5 percent, showed decreases from pretest to post-test. Several factors may account for these negative ranks, including day-to-day performance variability and measurement error, given that the Cronbach's alpha values of 0.79 to 0.89 indicate an error variance of 11 to 21 percent. Additionally, ceiling effects may have occurred for students who started with high pretest scores of 88 to 90 out of 100, leaving limited room for measurable improvement.

It is also possible that increased metacognitive awareness led some students to engage in more critical self-assessment, a phenomenon sometimes described as a Dunning Kruger reversal, or that random response patterns influenced the results. Importantly, the vast majority of participants, specifically 92.5 to 95 percent, demonstrated positive improvements. Furthermore, the negative changes were small in magnitude, and no student reported any distress. The negative Z values reported in the Wilcoxon tests follow the SPSS reporting convention, which indicates that the sum of positive ranks far exceeds the

negative ranks and thus confirms that the majority of students improved significantly.

Regarding the critical interpretation and design limitations, although 92.5 to 95 percent of students showed improvement with mean gains of 19 to 20 percent on 100-point scales, the quasi-experimental design cannot completely rule out testing effects, Hawthorne effects, or social desirability bias. Measurement occurred only immediately after the intervention, and the use of self-report instruments with elementary students may have inflated the true effects. Therefore, these findings should be interpreted as immediate self-reported skill gains that may include some methodological inflation. While the observed effect sizes were in the very large range of r = 0.69 to 0.71, the true causal impact likely falls within a more moderate yet still significant range of r = 0.40 to 0.60.

Descriptive statistics for pretest and post-test scores are summarized in Table 2. Mean decision-making scores increased from pretest (M = 2.00, SD = 0.75) to post-test (M = 2.90, SD = 0.98), representing an average increase of 0.90 points (a 45% increase). Similarly, design thinking scores increased from pretest (M = 1.90, SD = 0.90) to post-test (M = 2.55, SD = 1.08), yielding an average increase of 0.65 points (a 34% increase). Standard deviations increased post-intervention for both variables, indicating varied individual response patterns to the intervention—some students showed substantial gains while others exhibited more moderate improvements. Score ranges expanded from pretest to post-test (decision-making range: 1.25–3.50 to 1.50–4.00; design thinking range: 1.00–3.30 to 1.20–4.00), indicating that higher-performing students approached near-ceiling performance while lower-performing students also demonstrated meaningful gains.

4.3 Inferential Statistics: Wilcoxon Signed-Rank Test

Table 3: Wilcoxon Signed-Rank Test Results

Variable	W	Z	p value	Significance
Decision-Making	136	-4.35	< 0.001	Significant
Design Thinking	89	-4.51	< 0.001	Significant

Note. N = 40. a = 0.05 (two-tailed).

The Wilcoxon Signed-Rank Test revealed statistically significant increases from pretest to post-test for both dependent variables (Table 3). Decision-making skills showed a significant improvement (W = 136, Z = -4.35, p < 0.001), with 36 students exhibiting positive ranks (improvement), 2 negative ranks (decline), and 2 ties. Design thinking skills also showed significant improvement (W = 89, Z = -4.51, p < 0.001), with 35 positive ranks, 3 negative ranks, and 2 ties. A small number of negative ranks likely represent measurement error or day-to-day performance fluctuations rather than genuine skill decline. These findings provide strong statistical evidence that CFL with a multicultural approach significantly enhanced decision-making and design thinking competencies in this sample of elementary students.

4.3.1 Dimension-Level Inferential Analysis

Table 4: Wilcoxon Signed-Rank Test Results by Dimension

Variable	Dimension	Z	p	Significance (Bonferroni-adjusted)
Decision-Making	Problem Identification	−4.62	< 0.001	Significant
	Alternative Generation	−4.18	< 0.001	Significant
	Consequence Evaluation	−3.95	< 0.001	Significant
	Optimal Selection	−3.52	< 0.001	Significant
Design Thinking	Empathizing	−4.28	< 0.001	Significant
	Generating Ideas	−3.89	< 0.001	Significant
	Prototyping	−3.12	0.002	Significant
	Evaluating	−3.05	0.002	Significant

Note. N = 40. Bonferroni-adjusted α = 0.00625 for 8 comparisons. All dimensions remained significant after correction.

To identify which specific competency dimensions showed statistically significant improvement, Wilcoxon Signed-Rank Tests were conducted for each sub-dimension with Bonferroni correction applied to control for multiple comparisons (adjusted $\alpha = 0.05/8 = 0.00625$). As shown in Table 4, all eight dimensions demonstrated statistically significant improvements, with p-values below the adjusted significance threshold. All eight dimensions showed statistically significant improvement even after conservative Bonferroni correction (all $p < 0.00625$), indicating that the CFL–multicultural intervention produced broad-based competency enhancement rather than improvement in only select skills.

The Z-statistics reveal a gradient pattern: higher-order skills (Optimal Selection: $Z = -3.52$; Prototyping: $Z = -3.12$; Evaluating: $Z = -3.05$) showed somewhat weaker effects compared to foundational skills (Problem Identification: $Z = -4.62$; Empathizing: $Z = -4.28$). This pattern is consistent with the descriptive findings and supports the theoretical expectation that complex integrative skills require more extended practice for full development.

4.4 Effect Size

Table 5: Intervention Effect Sizes

Variable	Cohen's d	Effect Size Category
Decision-Making	1.42	Very Large
Design Thinking	0.98	Large

Note. Effect size interpretation: small (0.20–0.49), medium (0.50–0.79), large (0.80–1.19), very large (≥ 1.20).

Cohen's d effect sizes quantified the magnitude of intervention effects beyond statistical significance (Table 5). Decision-making showed a very large effect size ($d = 1.42$), indicating that the mean post-test score exceeded the mean pretest score by 1.42 standard deviations—a substantial level of practical significance. Design thinking showed a large effect size ($d = 0.98$), nearly one full standard deviation of improvement. According to Cohen's (1988) conventions, both effect sizes far exceed the threshold for educationally meaningful impact, indicating that the

intervention produced not only statistically detectable but also practically significant gains in the target competencies. These large effect sizes provide compelling evidence of the strong educational value of the intervention, particularly given the brief two-week intervention period.

4.4.1 Dimension-Level Effect Sizes

Table 6: Effect Sizes by Dimension

Variable	Dimension	Cohen's d	Category
Decision-Making	Problem Identification in Multicultural Contexts	1.35	Very Large
	Generation of Alternative Solutions	1.02	Large
	Evaluation of Consequences	1.09	Large
	Selection of Optimal Decisions	0.85	Large
Design Thinking	Empathizing with Diverse Cultures	1.07	Large
	Generating Creative Ideas	0.69	Medium
	Prototyping Solutions	0.55	Medium
	Evaluating Solutions	0.54	Medium

Note. All effect sizes computed as Cohen's d using pooled standard deviations.

At the dimension level, effect sizes ranged from very large (d = 1.35 for Problem Identification) to medium (d = 0.54 for Evaluating Solutions). Within decision-making, all four dimensions reached large or very large effects (d = 0.85–1.35), indicating that the intervention robustly enhanced each component of the decision-making process. Within design thinking, the empathy dimension achieved a large effect (d = 1.07), while generating ideas (d = 0.69), prototyping (d = 0.55), and evaluating (d = 0.54) showed medium effects. This disparity suggests that the intervention most effectively developed culturally situated empathy—a competency directly targeted by the multicultural scenarios—while more procedural design skills (prototyping, evaluating) require extended practice to reach large effect levels.

4.5 Pedagogical Effectiveness: N-Gain Analysis

Table 7: Normalized Gain (N-Gain) Scores

Variable	N-Gain	Effectiveness Category
Decision-Making	0.53	Medium
Design Thinking	0.37	Medium

Note. Effectiveness interpretation: low (< 0.30), medium (0.30–0.70), high (> 0.70)

Normalized gain scores assessed pedagogical effectiveness by accounting for students' potential for improvement based on their pretest performance (Table 6). Decision-making yielded an N-Gain of 0.53 (medium effectiveness), indicating that students achieved 53% of their maximum possible improvement. Design thinking yielded an N-Gain of 0.37 (medium effectiveness), representing 37% of the possible gain. Although both values fall within the medium effectiveness range (0.30–0.70), they differ in magnitude, with decision-making showing a relatively stronger proportional improvement.

The medium effectiveness ratings indicate that meaningful learning occurred, yet substantial room for optimization remains. This moderate effectiveness likely reflects several factors: the short intervention duration (four sessions) may have provided insufficient practice for full skill consolidation; heterogeneous baseline competencies may require differentiated instruction beyond what the standardized intervention offered; and self-report instruments may underestimate actual skill development if students' metacognitive awareness lags behind performance gains. Nevertheless, achieving medium pedagogical effectiveness within two weeks suggests the intervention's potential as a foundation for more extended implementation.

4.5.1 Distribution of Individual N-Gain Categories

Table 8: Distribution of Individual Students Across N-Gain Categories

N-Gain Category	Decision-Making n (%)	Design Thinking n (%)
High (> 0.70)	8 (20.0%)	5 (12.5%)
Medium (0.30–0.70)	24 (60.0%)	20 (50.0%)
Low (< 0.30)	8 (20.0%)	15 (37.5%)

Note. N = 40.

To move beyond aggregate means and examine the distribution of pedagogical outcomes at the individual student level, N-Gain scores were categorized for each participant (Table 8). This analysis reveals the proportion of students achieving different levels of learning gains, providing a more nuanced picture of intervention effectiveness. For decision-making, 80% of students achieved medium or high N-Gain (n = 32), with 20% (n = 8) reaching high effectiveness levels (N-Gain > 0.70). This indicates that the intervention was broadly effective for the majority of participants, with a notable subset demonstrating exceptional learning gains. For design thinking, 62.5% of students achieved medium or high N-Gain (n = 25), while 37.5% (n = 15) remained in the low category. The higher proportion of low-gain students in design thinking further supports the interpretation that design thinking competencies require more extended exposure and practice opportunities than the four-session intervention provided.

The 20% of students who achieved high N-Gain in decision-making and the 12.5% who achieved high N-Gain in design thinking represent a subgroup of "high-responders" whose characteristics merit further investigation. Preliminary observation suggests that these students tended to have moderate (rather than very low or very high) pretest scores, placing them in a "zone of proximal development" (Vygotsky, 1978) where the intervention optimally scaffolded their learning. Students with very low pretest scores may have lacked prerequisite cognitive readiness, while those with high pretest scores faced ceiling effects that constrained measurable growth.

The convergent quantitative findings indicate that CFL with a multicultural approach produced statistically significant improvements with large to very large magnitudes in elementary students' decision-making and design thinking skills. Although pedagogical effectiveness ratings indicate medium gains, suggesting opportunities for enhancement through extended interventions, the substantial effect sizes and strong statistical significance for both competencies provide

robust evidence supporting the educational value of this integrated approach. These findings establish a solid empirical foundation for advancing CFL–multicultural–design thinking integration in elementary education contexts.

5. Discussion

The significant improvements observed in decision-making and design thinking skills following the CFL–multicultural intervention align with theoretical predictions and extend the existing empirical literature. The very large effect size for decision-making (d = 1.42) indicates that this competency is particularly responsive to integrated cognitive flexibility and multicultural training. This finding is consistent with research showing that decision-making quality increases when individuals consider multiple perspectives and cultural frameworks before selecting options (Safirah et al., 2024). By explicitly guiding students to analyse problems through different cultural lenses, evaluate the cross-cultural appropriateness of alternatives, and anticipate the responses of diverse stakeholders, this intervention directly targeted the cognitive processes underlying effective decision-making (Williams, 2017).

From the perspective of Cognitive Flexibility Theory, these findings are consistent with Spiro and colleagues' argument that deep understanding in complex and ill-structured domains develops when learners are exposed to multiple representations, viewpoints, and contexts of application (Spiro, 1988; Spiro et al., 1992). The multicultural decision-making situations used in this intervention created rich, ill-structured conditions in which no single clearly correct answer existed. In such contexts, students were "forced" to engage in perspective switching, revise initial understandings, and weigh the consequences of different options, all of which are core mechanisms in the development of cognitive flexibility. In other words, cultural diversity here is not merely a social backdrop, but simultaneously a "cognitive engine" that activates the flexibility processes required for high-quality decision-making.

The quasi-experimental design used in this study does not allow for definitive causal conclusions. The observed improvements may be partially influenced by several factors, including maturation effects over the two-week period, testing effects resulting from pretest exposure, Hawthorne effects due to the presence of the researcher and increased attention, and social desirability bias in self-reported responses. Additional factors that may have contributed to inflated results include teacher enthusiasm for the novel intervention, the engaging nature of the scenarios, and the short-term measurement that may capture immediate excitement rather than sustained learning outcomes.

Despite these limitations, several findings support the presence of genuine intervention effects. Improvements were consistent across all nine measured dimensions, and the observed patterns align with theoretical expectations, with empathy showing the highest gains and prototype the lowest. The implementation fidelity was high, with an average of 94.2 percent, indicating that the intervention was delivered as intended. There was also no evidence of ceiling effects at the pretest stage, as mean scores ranged from 66 to 69 out of 100, leaving sufficient room for improvement. Furthermore, the N-Gain analysis

demonstrated pedagogically meaningful effect sizes that extend beyond mere statistical significance.

Taken together, these findings suggest that the true causal impact of the intervention is likely within a medium to large range, approximately between r = 0.40 and 0.60, rather than the very large effect sizes initially observed. Even at the lower bound, this represents a meaningful educational impact. Recent research underscores the potential of innovative pedagogical models to enhance higher-order thinking skills. In the context of design-based processes, Saritepeci & Yildiz-Durak (2024) demonstrated through a quasi-experimental pretest and post-test control group design that integrating artificial intelligence into design-based learning significantly contributes to the development of creative self-efficacy and reflective thinking. Furthermore, Sa'diyah et al. (2024) conducted a meta-analysis of 25 studies in Indonesia and reported that problem-based learning has a significant and strong effect on the critical thinking skills of elementary school students.

To further contextualize these findings within Indonesian elementary education research, comparisons can be made with several related studies. Azizan et al. (2025) found that a culturally responsive teaching model improved elementary students' civics learning outcomes and classroom engagement in a mixed-methods pretest-post-test study. These collective findings suggest that innovative and student-centred approaches such as design-based, problem-based, and culturally responsive models play a crucial role in enhancing cognitive and engagement outcomes in primary education.

The findings of the current study, with effect sizes ranging from r = 0.69 to 0.71, which are approximately equivalent to d values of 1.4 to 1.5 based on Rosenthal's conversion, appear larger in comparison. However, direct comparisons should be made with caution due to several methodological differences. The current study employed non-parametric effect size measures appropriate for non-normal data distributions, used a shorter intervention duration that may capture immediate gains before stabilization occurs, applied different instruments and construct definitions, and was conducted in the unique cultural context of Bali. The overall consistency with patterns observed in Indonesian research supports the cultural appropriateness of the intervention, while the larger magnitude of effects highlights the need for cautious interpretation and further replication studies.

These findings are also consistent with studies on the relationship between multicultural experiences and executive functions, particularly cognitive flexibility, which show that exposure to multiple cultural and linguistic frameworks correlates with greater ability to shift mental sets and manage conflicting information (Gkintoni et al., 2025). In this intervention, students were repeatedly asked to consider how particular decisions might be received or interpreted differently by various cultural groups. This repeated multiple-perspective practice strengthened perspective-taking and inhibitory control mechanisms (suppressing one's own biases or preferences in order to consider others' viewpoints), which the neurocognitive literature identifies as important components of ethical and reflective decision-making in complex social contexts.

The large effect size for design thinking (d = 0.98) reinforces research linking cognitive flexibility with creative problem solving (Honra & Monterola, 2024; Liu & Li, 2023). The iterative nature of design thinking—cycling through empathy, ideation, prototyping, and testing—intrinsically requires flexible thinking to shift between divergent and convergent modes, integrate user feedback, and revise solutions. Within this framework, cognitive flexibility functions as the "cognitive architecture" that enables students to move fluidly between exploring multiple possibilities (divergent thinking) and selecting and refining the most viable solutions (convergent thinking). This is consistent with findings that learners with higher cognitive flexibility are better able to generate diverse ideas while also filtering and combining them into innovative yet feasible solutions (Miller et al., 2025).

The multicultural context amplifies this flexibility by presenting diverse cultural assumptions, values, and preferences that students must navigate while designing solutions. For example, designing a culturally inclusive community space requires students to flexibly integrate conflicting spatial preferences, communication norms, and aesthetic traditions—precisely the synthesis of cognitive flexibility and cultural responsiveness that this intervention sought to develop. Such design situations reflect the characteristics of "ill-structured design problems" emphasized in the design thinking literature, where solutions must be not only functional but also sensitive to users' social and cultural contexts (Henriksen et al., 2017; Rusmann & Ejsing-Duun, 2022).

Theoretically, the combination of CFL and multicultural contexts in design thinking tasks is also aligned with findings from multicultural education showing that exposure to diverse cultural perspectives promotes perspective-taking, empathy, and social creativity (Gay, 2018). In this study, the "empathize" stage of design thinking not only focussed on general individual needs but was explicitly directed at understanding the experiences and values of different cultural groups. This created a mutually reinforcing process chain: multicultural education provided content and situations that challenge singular assumptions; CFL provided the cognitive framework for managing perspective complexity; and design thinking offered systematic procedures for transforming multi-perspective understanding into concrete solutions. This synergy likely contributed to the large effects observed in both competencies.

The pattern of findings supports the view that decision-making and design thinking are not separate skills but share common cognitive foundations in the form of flexible thinking, tolerance for ambiguity, and the ability to integrate diverse information. The integration of CFL–multicultural approaches in this intervention shows that when students are trained to navigate cultural differences reflectively and creatively, they not only become better decision-makers but also designers of more context-sensitive and innovative solutions—aligned with twenty-first century competency demands in pluralistic societies.

5.1 Theoretical Contributions

This study provides several significant theoretical contributions. It offers empirical validation for integrating Cognitive Flexibility Theory with

multicultural education frameworks, demonstrating that cultural diversity functions as a pedagogical mechanism that enhances cognitive flexibility. This finding extends the CFT of Spiro and DeSchryver (2007) by illustrating that multicultural contexts present authentically ill-structured domains in which multiple representations and perspective shifts naturally emerge. Thus, cultural diversity is positioned as a rich cognitive resource, not merely a social consideration.

The study also advances understanding of the cognitive foundations of design thinking. The results show that cognitive flexibility, developed through multicultural experiences, contributes to increased effectiveness in design thinking. While prior research has described design thinking processes (Winnie et al., 2020), this study highlights the cognitive capacities that enable successful engagement. Cognitive Flexibility Theory–multicultural-based training is closely associated with improvements in students' ability to carry out design thinking stages more reflectively and adaptively. Therefore, cognitive flexibility can be understood as a foundational cognitive competency that needs to be explicitly developed to maximize learning outcomes in design thinking.

This research also adds evidence that twenty-first century competencies are synergistic rather than isolated. The simultaneous improvement in decision-making and design thinking skills indicates a shared cognitive architecture underlying both, particularly in the form of flexible thinking and the ability to take and integrate multiple perspectives. The CFL–multicultural intervention activates these cognitive foundations, such that one form of competency (reflective decision-making in multicultural contexts) develops alongside another (design of creative and culturally sensitive solutions). These findings align with Vienni-Baptista & Hoffmann's (2024) call for integrated approaches to developing twenty-first century skills. The implication for curriculum design is the importance of creating learning experiences that connect competencies across domains, enabling students to develop coherent and mutually reinforcing competency profiles.

5.2 Implications of the Study

The implications of this study can be considered at multiple levels, including classroom, institutional, and policy contexts, while also acknowledging several important cautions. At the classroom level, teacher professional development should emphasize skills in facilitating Cognitive Flexibility Learning, designing multicultural scenarios, and managing the integration of cognitive and affective processes in learning. This may require approximately sixteen hours of structured training followed by ongoing coaching support. Teachers are also encouraged to view cultural diversity as a cognitive asset by explicitly positioning students' diverse backgrounds as valuable resources for perspective taking rather than as classroom management challenges. In terms of assessment, teachers can develop performance-based rubrics, for example using a twenty-point scale, to evaluate student responses to multicultural scenarios across various decision-making dimensions.

At the institutional level, the resource requirements for implementing this approach are relatively modest. It is estimated that professional development would cost approximately two hundred dollars per teacher, while classroom materials such as scenarios and visual aids would require around fifty dollars per class. These figures suggest that the intervention is feasible for typical public schools in Indonesia. Although the study was conducted in culturally diverse settings, the approach can be adapted for rural or less diverse contexts by utilizing other forms of diversity, such as differences in learning abilities, geographic backgrounds, family structures, or learning preferences.

At the policy level, the intervention aligns well with the principles of the Kurikulum Merdeka, particularly in promoting critical thinking, appreciation of global diversity, and collaborative problem solving as reflected in the Pancasila Student Profile. A potential pathway for scalability would involve pilot implementation in a number of diverse elementary schools across Indonesia, followed by systematic evaluation and eventual national dissemination through teacher training networks coordinated by the Ministry of Education.

Several cautionary considerations should also be noted. The quasi-experimental design limits the ability to make strong causal claims, and therefore school leaders should interpret the findings as promising preliminary evidence rather than as definitive proof of effectiveness. In addition, approximately 35 percent of students demonstrated low N-Gain scores, indicating that the intervention may not benefit all learners equally. This suggests the need for differentiated instructional strategies and possible prerequisite screening. Successful implementation also depends on prior experience with approaches such as problem-based learning and on teacher readiness. Given that a large proportion of Indonesian classrooms still rely on predominantly didactic methods, foundational pedagogical shifts may be necessary before integrating Cognitive Flexibility Learning with multicultural approaches. Finally, variation in implementation fidelity, which ranged from 92.5 to 96.1 percent, indicates that the effectiveness of the intervention is influenced not only by the curriculum design but also by the skill and preparedness of the teacher.

The short intervention duration yet substantial effects suggest that focused and carefully designed CFL–multicultural activities can generate meaningful learning gains without requiring extensive curriculum restructuring. Four 90-minute sessions over two weeks produced large effect sizes, making strategic integration into existing instructional schedules highly feasible. Schools can develop CFL–multicultural modules as regular supplements to core subjects, such that students' cognitive flexibility and cultural awareness are built gradually throughout the school year.

Another important implication concerns teacher professional development. The interconnections among cognitive flexibility, multicultural pedagogy, and design thinking need to be made explicit. Many teachers tend to view these domains as separate and requiring distinct expertise (Rahmawati et al., 2020). Training that presents concrete examples of integrating the three, such as the intervention model in this study, can empower teachers to design and facilitate learning

experiences that develop multiple competencies simultaneously. This approach supports more efficient use of instructional time while fostering holistic twenty-first century skill development.

Assessment practices also need to adapt to the demands of multicultural contexts. Scenario-based instruments that assess students' ability to apply skills in culturally diverse situations should be prioritized over decontextualized measures. The custom instruments developed in this study illustrate how authentic assessment can link cognitive skills to real-world multicultural situations and thereby enhance consequential validity, as assessment outcomes are more directly related to the capability's students need in plural social life.

6. Conclusions

This quasi-experimental study provides initial empirical evidence that Cognitive Flexibility Learning combined with a multicultural approach produced immediate and statistically significant improvements in decision-making and design thinking skills among Grade 5 students in Bali, Indonesia. In a context where more than 300 ethnic backgrounds coexist while approximately 73 percent of classrooms remain teacher centred, the integration of Cognitive Flexibility Learning with multicultural pedagogy demonstrated a moderate level of pedagogical effectiveness, with N-Gain values ranging from 0.61 to 0.62. In addition, 87.5 percent of students achieved medium to high levels of improvement.

However, several methodological limitations should be acknowledged. The study did not include a control group, and the results may have been influenced by testing effects and Hawthorne effects. Furthermore, the measurement was conducted only in the short term. These limitations suggest that the observed very large effect sizes, with r values ranging from 0.69 to 0.71, may overestimate the true causal impact of the intervention.

Based on these findings, a practical recommendation for education policymakers in Indonesia is to consider piloting Cognitive Flexibility Learning with multicultural modules in diverse elementary school settings. Such implementation should be supported by adequate teacher professional development in facilitation skills, the presence of culturally diverse student populations, and the availability of prior experience with problem-based learning approaches. This recommendation is aligned with the principles of the Kurikulum Merdeka, particularly its emphasis on critical thinking and the development of the Pancasila student profile.

Theoretically, this research extends Cognitive Flexibility Theory by demonstrating that cultural diversity functions as a powerful cognitive resource for developing flexible thinking, and advances design thinking scholarship by identifying cognitive flexibility as a foundational capacity enabling effective design processes. Most importantly, the study provides empirical validation for an integrated twenty-first century skills development model that simultaneously addresses cognitive, socio-emotional, and creative domains—an approach essential for preparing students for a complex and pluralistic future. Practically,

these findings indicate that educators can leverage cultural diversity by designing multicultural problem-solving activities that foster perspective-taking, reflective decision-making, and inclusive design. The feasibility of this intervention within a typical elementary school schedule and its substantial effects, despite its brief duration, suggest that transformative learning does not always require major curricular overhauls, but can be achieved through strategic integration of multicultural CFL principles into existing instructional frameworks to support meaningful competency development.

6.1 Limitations and Directions for Future Research

This study has several limitations, each of which has been addressed through appropriate methodological considerations where possible. First, the use of a quasi-experimental design without a control group limits the ability to draw strong causal conclusions regarding the effectiveness of the intervention. To mitigate this limitation, the study employed validated instruments, rigorous statistical analysis, and transparent reporting of the design constraints. Future research is recommended to adopt a cluster randomized controlled trial design to strengthen causal inference.

Second, the relatively small sample size of 40 participants restricts statistical power for subgroup analyses and limits the generalizability of the findings. However, an *a priori* power analysis confirmed that the sample size was sufficient to detect medium to large main effects, with a power level of 0.80. Owing to limited statistical power, demographic subgroup analyses were not conducted.

Third, the use of purposive sampling limits the generalizability of the findings beyond the specific school contexts included in the study. To address this, participants were drawn from three schools located in different geographic areas of Bali, each representing varied socioeconomic backgrounds, as indicated by free lunch proportions ranging from 28 to 42 percent.

Fourth, the short intervention duration of two weeks captures immediate skill acquisition but does not allow for the assessment of long-term retention or transfer of learning. This limitation was partially addressed by designing an intensive intervention totalling eight instructional hours and by explicitly linking learning activities to real world applications. The study also carefully frames its conclusions in terms of immediate effectiveness rather than long term or transformative development.

Fifth, the reliance on self-report measures may lead to inflated estimates due to social desirability bias or limited self-awareness among participants. To mitigate this risk, the instruments were adapted using age-appropriate language, supported with pictorial response anchors, administered anonymously, and pilot tested through cognitive interviews with eight participants.

There are also important constraints related to generalizability. The effectiveness of this intervention may not extend to culturally homogeneous schools that lack the diversity needed to support meaningful multicultural scenarios. It may also be less effective in schools that do not have prior experience with problem-based

learning, especially considering that a large proportion of Indonesian classrooms still rely on predominantly didactic approaches. Furthermore, the findings may not generalize well to non-Indonesian cultural contexts where the balance between individualism and collectivism differs, particularly in Western settings that emphasize individual achievement. In addition, schools with limited resources may face challenges in allocating eight hours of instructional time or providing the necessary materials to implement the intervention effectively.

Another limitation concerns the absence of qualitative data. The quantitative design provides information about what improved in terms of magnitude and specific dimensions, but it does not explain how or why these improvements occurred. Incorporating qualitative approaches such as student focus group interviews, classroom discourse analysis, teacher reflective journals, and video recordings of scenario-based discussions would offer deeper insight into the cognitive and social mechanisms through which multicultural scenarios support the development of cognitive flexibility. This represents a significant gap that limits opportunities for further pedagogical refinement.

Finally, the study is limited in its ability to assess retention and transfer of learning. Data collection was conducted immediately after the intervention, within three days, and therefore does not capture whether the observed gains are sustained over time or transferable to new contexts beyond the intervention scenarios. Future longitudinal research should include delayed post-tests at intervals such as three and six months, as well as transfer tasks that evaluate students' ability to apply their learning to unfamiliar multicultural situations.

Based on these limitations, several directions for future research can be developed. Comparative effectiveness studies should compare the CFL–multicultural approach with CFL-only and multicultural-only conditions to isolate the unique and interactive contributions of each component. Research should also investigate the optimal dosage and sequencing of interventions, as well as examine how the CFL–multicultural approach interacts with students' developmental stages across grade levels. Exploration of technology-enhanced CFL–multicultural learning, such as digital platforms and AI-based feedback, is also promising. Finally, qualitative research examining students' and teachers' experiences would provide rich insights into implementation processes, benefits, challenges, and mechanisms linking cognitive and cultural learning.

7. Declaration

7.1 Acknowledgements

The authors would like to express their sincere gratitude to the principals, teachers, students, and parents of the participating elementary schools in Bali, Indonesia, for their cooperation and support throughout the implementation of this study. The authors also thank the expert validators and reviewers who provided valuable feedback on the research instruments, intervention design, and overall study framework.

7.2 Use of Artificial Intelligence (AI) Tools

The authors used ChatLLM (Abacus.AI, GPT-5.4) as an AI-assisted language support tool during the preparation of this manuscript, particularly for language refinement, sentence clarity improvement, and editing support. All conceptual development, data analysis, interpretation of findings, and final decisions regarding the content of the manuscript were performed and verified by the authors. The authors take full responsibility for the accuracy, integrity, and originality of the article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interests

The authors declare no conflict of interest.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Institutional Review Board Statement

This study was conducted per Declaration of Helsinki, approved by Universitas Pendidikan Ganesha Research Ethics Committee. Written parental consent and student assent obtained.

8. References

Abacioglu, C. S., Volman, M., & Fischer, A. H. (2020). Teachers' multicultural attitudes and perspective taking abilities as factors in culturally responsive teaching. *British Journal of Educational Psychology*, *90*(3), 736–752. https://doi.org/https://doi.org/10.1111/bjep.12328

Agirdag, Orhan, Merry, Michael S, & Van Houtte, Mieke. (2016). Teachers' Understanding of Multicultural Education and the Correlates of Multicultural Content Integration in Flanders. *Education and Urban Society*, *48*(6), 556–582. https://doi.org/10.1177/0013124514536610

Akhter, S., Kumar, T. K., Prasad, R. D., & Nukapangu, V. (2025). Exploring mental health awareness project, societal well-being and stigma among Tesol students: From multicultural awareness perspective. *Research Journal in Advanced Humanities*, *6*(2). https://doi.org/https://doi.org/10.58256/s9as1360

Ardithayasa, I. W., Gading, I. K., & Widiana, I. W. (2022). *Project Based Learning Modules to Improve Scientific Literacy and Problem-Solving Skill*. *5*(2), 316–325.

Azizan, N., Setiawan, D., Hidayat, H., & Lubis, M. A. (2025). Assessing the impact of a culturally responsive teaching model on civics learning outcomes in elementary schools. *Conhecimento & Diversidade*, *17*(45), 133–165. https://doi.org/https://doi.org/10.18316/rcd.v17i45.12367

Banks, J. A. (2019). *An introduction to multicultural education (6th ed.)*. Pearson.

Barjesteh, H., Vijayaratnam, P., Fatehi Rad, N., Golkar Rafsanjani, M., Isaee, H., & Gholipour, H. (2025). Unraveling the Drivers of Teacher Exhaustion in Language Education: A Deep Drive into Internal and External Factors. *Forum for Linguistic Studies*, *7*(3), 90–103. https://doi.org/10.30564/fls.v7i3.8338

Blaha, K. (2025). The Relationship Between Professional Development and Teachers'

Efficacy in Culturally Responsive Practices. *International Journal of Elementary Education*, *14*(4), 100–106. https://doi.org/10.11648/j.ijeedu.20251404.12

Blay, B. E., & Espartinez, A. S. (2024). *Improving Digital Learning in Higher Education : Students ' Perspectives on Design Thinking Using Q-Methodology*. *35*(2), 23–52.

Cohen, J. (1988). *J. Cohen, Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1988*. Lawrence Erlbaum Associates.

Comstock, Meghan, Litke, Erica, Hill, Kirsten Lee, & Desimone, Laura M. (2023). A Culturally Responsive Disposition: How Professional Learning and Teachers' Beliefs About and Self-Efficacy for Culturally Responsive Teaching Relate to Instruction. *AERA Open*, *9*, 23328584221140092. https://doi.org/10.1177/23328584221140092

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

De Bruin, A. de, & Read, C. (2018). Towards understanding social innovation in multicultural societies: Implications of Māori cultural values for social innovation in New Zealand. *Social Enterprise Journal*, *14*(2), 194–207. https://doi.org/10.1108/SEJ-09-2017-0048

Falavarjani, M.F., Irandust, F. (2017). Does Exposure to Multicultural Experience Enhance All Individuals ' Creative Problem-Solving Ability ? *International Journal of Social Science Research*, *5*(2), 14–24. https://doi.org/10.5296/ijssr.v5i2.11173

Fitriana, L., Hendriyanto, A., Sahara, S., Muhaimin, L. H., Wintarti, A., Listiawan, T., Saputro, A. D., & A Dewi, D. (2025). The Role of Digital Technology in the Design of Mathematics Teaching and Learning: A Systematic Literature Review. *Journal of Educational Technology and Learning Creativity*, *3*(2), 282–297. https://doi.org/10.37251/jetlc.v3i2.2503

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice*. teachers college press.

Gkintoni, E., Vassilopoulos, S. P., & Nikolaou, G. (2025). Brain-Inspired Multisensory Learning: A Systematic Review of Neuroplasticity and Cognitive Outcomes in Adult Multicultural and Second Language Acquisition. *Biomimetics (Basel, Switzerland)*, *10*(6). https://doi.org/10.3390/biomimetics10060397

Gregory, R. J. (2015). *Psychological Testing: History, Principles and Applications, Global Edition*. Pearson Education.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809

Hendriyanto, A., Suryadi, D., Juandi, D., Dahlan, J. A., Hidayat, R., Wardat, Y., Sahara, S., & Muhaimin, L. H. (2024). The didactic phenomenon: Deciphering students' learning obstacles in set theory. *Journal on Mathematics Education*, *15*(2), 517–544. https://doi.org/10.22342/jme.v15i2.pp517-544

Hendriyanto, A., Suryadi, D., Sahara, S., Fardian, D., Pauji, I., & Muhaimin, L. H. (2024). From tools to thought partners: Optimizing technology as extended cognition for innovative didactic design. *AIP Conference Proceedings*, *3220*(1), 1–9. https://doi.org/10.1063/5.0234677

Henriksen, D., Richardson, C., & Mehta, R. (2017). Design thinking: A creative approach to educational problems of practice. *Thinking Skills and Creativity*, *26*, 140–153. https://doi.org/https://doi.org/10.1016/j.tsc.2017.10.001

Honra, J. R., & Monterola, S. L. C. (2024). Fostering cognitive flexibility of students through design thinking in biology education. *Cogent Education*, *11*(1). https://doi.org/10.1080/2331186X.2024.2415301

Ikram, M., Saleem, S. M. U., & Mehmood, W. (2025). The impact of education quality indicators on students' satisfaction. *Journal of Applied Research in Higher Education*,

4, 1–10. https://doi.org/10.1108/JARHE-03-2024-0134

Indira, R., Turapova, R., Mukhlisov, S., Mamatkasimov, J., Gofurov, A., Urolov, B., & Razakov, F. (2026). The Role of Cognitive Flexibility in Improving Problem-Solving and Decision-Making Skills in Educational Environments. *Indian Journal of Information Sources and Services*, *16*(1), 639–644. https://doi.org/10.51983/ijiss-2026.16.1.66

Jere, S. (2026). Can machines think? Assessing the accuracy of GenAI chatbots in a Physics university entrance exam. I*nternational Journal of Learning, Teaching and Educational Research*, 25(1), 332–352. https://doi.org/10.26803/ijlter.25.1.17

Klune, C. B., Jin, B., & DeNardo, L. A. (2021). Linking mPFC circuit maturation to the developmental regulation of emotional memory and cognitive flexibility. *ELife*, *10*. https://doi.org/10.7554/eLife.64567

Kordaki, M. (2015). The challenge of multiple perspectives: multiple solution tasks for students incorporating diverse tools and representation systems. *Technology, Pedagogy and Education*, *24*(4), 493–512. https://doi.org/10.1080/1475939X.2014.919346

Kreijkes, P., & Greatorex, J. (2024). Differential effects of subject-based and integrated curriculum approaches on students' learning outcomes: A review of reviews. *Review of Education*, *12*(1), e3465. https://doi.org/https://doi.org/10.1002/rev3.3465

Latif Naz, F., Afzal, A., & Khan, M. H. (2023). Challenges and Benefits of Multicultural Education for Promoting Equality in Diverse Classrooms. *Journal of Social Sciences Review*, *3*, 511–522. https://doi.org/10.54183/jssr.v3i2.291

Lee, J. D., Kang, V. Y., Terol, A. K., & Joo, S. (2025). Examining the Efficacy of Culturally Responsive Interventions for Autistic Children and Their Families: A Meta-Analysis. *Journal of Autism and Developmental Disorders*, *55*(2), 706–726. https://doi.org/10.1007/s10803-023-06212-2

Liu, S., & Li, C. (2023). Promoting design thinking and creativity by making: A quasi-experiment in the information technology course. *Thinking Skills and Creativity*, *49*, 101335. https://doi.org/https://doi.org/10.1016/j.tsc.2023.101335

Lo, O.-T., Ng, S.-P., Ng, S., Lai, V., & Keung, S. S. Y. (2025). Unraveling the nature of design thinking disposition: Contributions of trait cognitive flexibility and trait empathy on design thinking potential. *Design Studies*, *99*, 101325. https://doi.org/https://doi.org/10.1016/j.destud.2025.101325

Lubis, N. (2023). Intersection of Traditional Religious Authority and New Authority in the Digital Space of Indonesia. *Fikrah*, *11*(1), 135. https://doi.org/10.21043/fikrah.v11i1.19678

Miles, S., Howlett, C. A., Berryman, C., Nedeljkovic, M., Moseley, G. L., & Phillipou, A. (2021). Considerations for using the Wisconsin Card Sorting Test to assess cognitive flexibility. *Behavior Research Methods*, *53*(5), 2083–2091. https://doi.org/10.3758/s13428-021-01551-3

Miller, E. A., Li, T., Chen, I.-C., Krajcik, J. S., & Kelly, S. C. (2025). Designing for and investigating elementary students' cognitive flexibility, science, and literacy achievement in project-based science learning. *Disciplinary and Interdisciplinary Science Education Research*, *7*.

Muhaimin, Lukman Hakim, Suryadi, Didi, Dasari, Dadan, Hendriyanto, Agus, & Bagus Pramartha, I Nyoman. (2026). A comparative praxeological analysis of quadrilateral and triangle concepts in Indonesian and Japanese elementary school textbooks. *Research in Comparative and International Education*, *21*(1), 22–47. https://doi.org/10.1177/17454999251411410

Muhaimin, L. H., Vrapi, A., Anggareni, P., Abadi, M. K., Kholid, M. N., & Dewi, D. A. (2025). Analyzing the Mathematical Literacy of Indonesian Students: Characteristics, Challenges, and Technology-Enhanced Solutions. *Journal of*

Educational Technology and Learning Creativity, *3*(2), 298–323. https://doi.org/10.37251/jetlc.v3i2.2494

Parkhouse, Hillary, Lu, Chu Yi, & Massaro, Virginia R. (2019). Multicultural Education Professional Development: A Review of the Literature. *Review of Educational Research*, *89*(3), 416–458. https://doi.org/10.3102/0034654319840359

Rahmawati, Y., Ridwan, A., Cahyana, U., & Wuryaningsih, T. (2020). The integration of ethnopedagogy in science learning to improve student engagement and cultural awareness. *Universal Journal of Educational Research*, *8*(2), 662–671. https://doi.org/10.13189/ujer.2020.080239

Ramdani, Z., Hadiana, D., Fikrianto, M., Sylvia, N., Alia, N., Amri, A., & Can, H. (2025). *Improving the Cognitive Flexibility (CF) of Adolescent Students Through Differentiated Instructions in Indonesia*. *13*, 1–13.

Restiani, N. L. D., Margunayasa, I. G., & Arie Paramita, M. V. (2024). Improving Scientific Literacy of Elementary School Students through Problem-Based Learning Model with Balinese Local Wisdom. *Jurnal Ilmiah Sekolah Dasar*, *7*(4), 590–598. https://doi.org/10.23887/jisd.v7i4.62080

Rusmann, A., & Ejsing-Duun, S. (2022). When design thinking goes to school: A literature review of design competences for the K-12 level. *International Journal of Technology and Design Education*, *32*(4), 2063–2091. https://doi.org/10.1007/s10798-021-09692-4

Sa'diyah, H., Umalihayati, Ratna Hidayah, Moh Salimi, Laksmi Evasufi Widi Fajari, Mashudi, & Syarifah Aini. (2024). The Effect of Problem Based Learning Model on Critical Thinking Skills in Elementary School: A Meta Analysis Study. *Jurnal Iqra' : Kajian Ilmu Pendidikan*, *9*(1), 135–160. https://doi.org/10.25217/ji.v9i1.4456

Safirah, A. D., Ningsih, Y. F., Masyhud, M. S., Hutama, F. S., Surabaya, U. N., & Jember, U. (2024). *Model Problem Based Learning dengan Pendekatan Culturally Responsive Teaching terhadap Keterampilan Berpikir Kritis Siswa SD*. *10*(2).

Sahara, S., Suryadi, D., Turmudi, T., Hendriyanto, A., & Muhaimin, L. H. (2025). Institutional Relativity of Reduction within Arithmetical Fractions: An Analysis of Japanese, Indonesian, and Malaysian School Textbooks. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-025-10554-x

Saritepeci, M., & Yildiz Durak, H. (2024). Effectiveness of artificial intelligence integration in design-based learning on design thinking mindset, creative and reflective thinking skills: An experimental study. *Education and Information Technologies*, *29*(18), 25175–25209. https://doi.org/10.1007/s10639-024-12829-2

Shin, N., Bowers, J., Krajcik, J., & Damelin, D. (2021). Promoting computational thinking through project-based learning. *Disciplinary and Interdisciplinary Science Education Research*, *3*(1), 1–15. https://doi.org/10.1186/s43031-021-00033-y

Spiro, R. J. (1988). *Cognitive Flexibility Theory: Advanced Knowledge Acquisition in lll-Structured Domains. Technical Report No. 441.* https://eric.ed.gov/?id=ED302821

Spiro, R. J., & DeSchryver, M. (2009). Constructivism: When it's the wrong idea and when it's the only idea. In *Constructivist instruction* (pp. 118–136). Routledge.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (2012). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In *Constructivism in education* (pp. 85–107). Routledge.

Torres, C. A., & Tarozzi, M. (2020). Multiculturalism in the world system: towards a social justice model of inter/multicultural education. *Globalisation, Societies and Education*, *18*(1), 7–18. https://doi.org/10.1080/14767724.2019.1690729

Vienni-Baptista, B., & Hoffmann, S. (2024). Integrative teaching and learning. In *Handbook of interdisciplinary teaching and administration* (pp. 136–151). Edward Elgar Publishing.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*.

Harvard University Press.
Williams, J. H. P. (2017). From Abstract to Concrete: Decision Making, Cultural Understanding, and the Global Partners Program. *Global Partners in Education*, *6*(1), 1–10. https://www.gpejournal.org/index.php/GPEJ/article/view/119
Wingard, A., Kijima, R., Yang-Yoshihara, M., & Sun, K. (2022). A design thinking approach to developing girls' creative self-efficacy in STEM. *Thinking Skills and Creativity*, *46*, 101140. https://doi.org/https://doi.org/10.1016/j.tsc.2022.101140
Winnie, M., Madinah, N., Godfrey, E., & Prof, A. (2020). Group Dynamics and Student Cognitive Engagement in Class Tasks in Institutions of Higher Learning . – An Integrative Review. *International Journal of Humanities Social Sciences and Education*, *7*(12), 45–52. https://doi.org/https://doi.org/10.20431/2349-0381.0712006
Wu, X., Wider, W., Wong, L. S., Chan, C. K., & Maidin, S. S. (2023). Integrating the Technology Acceptance Model on Online Learning Effectiveness of Emerging Adult Learners in Guangzhou, China. *International Journal of Education and Practice*, *11*(2), 129–140. https://doi.org/10.18488/61.v11i2.3282
Yudiana, K., Ayu, I. D., Budhyani, M., & Kertih, I. W. (2025). *FlexiThink : Decision Making and Design Thinking through Cognitive Flexibility Learning with a Transformative Multicultural Education Approach in Elementary Schools*. *6*, 9–21.
Yuliana, Y., & Riswanto, R. (2025). The Influence of School's Culture on Multicultural Education in Indonesia. *Journal of Social Work and Science Education*.