

## A MULTIDISCIPLINARY UMBRELLA REVIEW ON FLIPPED LEARNING OUTCOMES

**Dr. Abdullah ARSLAN**

ORCID: 0000-0002-3979-6371  
School of Foreign Languages  
Akdeniz University  
Antalya, TURKIYE

**Dr. Alaettin ISERI**

ORCID: 0000-0003-3621-7220  
Faculty of Arts and Sciences  
Kirkklareli University  
Kirkklareli, TURKIYE

**Received:** 19/06/2025 **Accepted:** 26/09/2025

### ABSTRACT

This umbrella review synthesizes evidence from 41 qualitative, meta-analytic, and mixed-method reviews to reveal the effectiveness of FL on cognitive, affective, and interpersonal domains across STEM, Medical and Health Sciences (M&H), and Social Sciences (SS). The results reveal that while flipped learning (FL) demonstrates positive effects on academic achievement in mathematics, science, and English language teaching, evidence for other domainal outcomes across fields remains inconclusive due to several limitations such as study scope, methodology and regional restrictions. Specifically, evidence across fields on other cognitive outcomes including autonomy, conceptual understanding, and higher-order thinking skills has the potential for fostering autonomy, conceptual understanding, and higher-order thinking skills, but the current evidence base is insufficient to definitively confirm these claims. Similarly, FL practices across fields seem conducive to enhancing affective and interpersonal domain outcomes; however, the predominance of qualitative evidence necessitates more quantitative and mixed method research for robust conclusions. Our synthesis also highlights the lack of in-in reporting and exploration of specific mechanisms by which FL cultivates domainal outcomes such as higher order thinking skills, motivation, and collaboration. Moreover, the evidence in the reviews highlights the absence of detailed reporting on the measurement metrics used to assess affective and interpersonal outcomes across fields in the reviews. Researchers should prioritize discipline-specific investigations, explore outcome-related mechanisms, develop standardized metrics, and bridge the gap between theory and practice through the integration of relevant learning theories into the design of FL strategies.

**Keywords:** Flipped learning, learning outcomes, multidisciplinary umbrella review, thematic analysis.

### INTRODUCTION

Educational institution and thought are a dynamic structure that constantly transforms to meet the conditions and requirements of the age it exists in. This transformation, created by the conditions of the 21st century such as technological advancements, globalization, and a dynamic job market, requires educational systems to focus not just on knowledge transmission but also on fostering a broad range of skills and competencies. In this respect, the essence of the traditional approach, which focuses on the transfer of specific information and the preservation of values in education, cannot meet the needs of our age for individuals who can adapt to change, cope with change, and live together with multiple values. Traditional and passive learning methods are no longer sufficient to equip learners with these crucial skills (Huang et al., 2020; Lee & Tsai, 2011; Martin, 2018; Shih & Huang, 2019; Taguma et al., 2018). Simply transferring knowledge is inadequate for preparing individuals for the challenges they may face in their personal, social, and professional lives (Care,

2017; Fry et al., 2009). These skills, broadly including problem-solving, communication, critical thinking, and collaboration, are essential for navigating the complexities of the 21st century (Ramma et al., 2021). Therefore, educators and institutions must embrace innovative methodologies that promote active learning and skill development (Bialik & Fadel, 2015; Francis et al., 2020; Gaddam et al., 2016).

Flipped learning (FL) has emerged in the first quarter of the 21st century—especially during the pandemic—as a technology-supported, interactive approach that relaxes dependence on time and place and synthesizes experiential, cognitive, and social learning theories. Its theoretical arguments and practical outcomes have gained renewed attention in recent years as a pedagogical approach aligned with the needs of 21st-century learners (Khlaisang & Teo, 2024; Zhao et al., 2021). It allows for personalized learning experiences (Alfonso et al., 2025), fosters critical thinking, and encourages collaboration and communication (Castedo et al., 2018; Rodríguez-Chueca et al., 2019; Sanchez-Gil-Machín et al., 2025; Sevillano-Monje et al., 2022). By integrating FL, educators can create engaging environments that effectively equip students with the necessary competencies to thrive in a dynamic and interconnected world (Bialik & Fadel, 2015; Miller & Ahmad, 2000). In FL pedagogy, students first engage with course content outside of the classroom through videos, podcasts, and reading texts (Bergmann & Sams, 2012; Bishop & Verleger, 2013; Lage et al., 2000). This pre-class engagement supports self-paced learning and accommodates diverse learning styles and preferences (Hsieh et al., 2016; Yu & Wang, 2016). Classroom sessions then transform into collaborative spaces where students apply and expand their knowledge (Hamdan et al., 2013; King, 1993). Teachers act as facilitators, emphasizing practical application and critical analysis of ideas through discussions, problem-solving tasks, and interactive activities (Bergmann & Sams, 2012; Jensen et al., 2015). The adaptability of FL makes it effective not only for reinforcing subject-specific knowledge but also for cultivating broader skills such as problem-solving, communication, and teamwork (Price & Walker, 2019; Shih & Tsai, 2019), with proven applicability across diverse disciplines including nursing, language learning, mathematics, and engineering (Betihavas et al., 2016; Lo & Hew, 2020; Sullivan, 2022).

In terms of effectiveness and outcomes in teaching–learning processes, FL offers a promising model; by integrating FL, educators can create engaging learning environments that equip students with the skills and competencies needed to thrive in a dynamic, interconnected world (Bialik & Fadel, 2015; Miller & Ahmad, 2000). FL, while not a new concept, has seen a resurgence in recent years as a transformative pedagogical approach that prioritizes student-driven learning by reshaping traditional classroom practices (Al-Hammouri et al., 2020; Castedo et al., 2019; Chen et al., 2019; Francis et al., 2020; Rodríguez et al., 2019; Shih & Huang, 2019). In FL pedagogy, students first engage with course content outside of the classroom through videos, podcasts, and reading texts (Bergmann & Sams, 2012; Bishop & Verleger, 2013; Lage et al., 2000). The pre-class session allows for self-paced learning, catering to diverse learning styles and preferences (Hsieh et al., 2016; Yu & Wang, 2016). In-class session transforms into a dynamic, collaborative space where students build upon their independently acquired knowledge (Hamdan et al., 2013; King, 1993). Teachers guide rather than control the learning process, emphasizing practical application and critical analysis of ideas through collaborative discussions, problem-solving tasks, and interactive activities (Bergmann & Sams, 2012; Jensen et al., 2015).

The versatility and adaptability of FL allow it to effectively promote both subject-specific knowledge and critical skills such as problem-solving, communication, and teamwork (Price & Walker, 2019; Shih & Tsai, 2019) across diverse disciplines (Betihavas et al., 2016; Lo & Hew, 2020; Sullivan, 2022). In addition, FL aligns with modern educational goals by accommodating diverse learning styles and preparing students for real-world situations (Xiao et al., 2018). By fostering a dynamic, learner-focused approach, it equips students with essential skills for today's workplaces while demonstrating effectiveness at different educational levels and across multiple subject areas (Birgili et al., 2021), such as nursing (El-Banna et al., 2017), STEM fields (Kostaris et al., 2017), and English language learning (Lee & Wallace, 2018; Lin & Hwang, 2018). The growing recognition of FL's adaptability has also attracted global research interest (Birgili et al., 2021; Cheng et al., 2020; Galindo-Domínguez & Bezanilla, 2019), and its widespread adoption in universities underscores its potential to transform learning and better prepare students for the demands of the 21st century (Birgili et al., 2021; Jiang et al., 2020; Lundin et al., 2018).

Recent years have witnessed a surge in meta-analytical and qualitative systematic reviews exploring FL pedagogy, indicating its growing adoption and prompting investigations into its efficacy (Chen et al.,

2019; Evans et al., 2019). However, the overall effectiveness of FL in enhancing academic performance and skills across different disciplines and educational levels remains inconclusive. For example, research on FL in Medical and Health fields, and STEM fields shows promising but mixed results, largely due to methodological weaknesses like small sample sizes and lack of standardized measures (Evans et al., 2019; Chen et al., 2018). In Medical and Health fields, studies often show no significant or only minor gains in academic performance (Evans et al., 2019; Chen et al., 2018; Gillette et al., 2018), with any positive findings limited by poor generalizability (Hu et al., 2018). FL's effectiveness appears constrained by contextual factors, and existing research overemphasizes academic achievement while neglecting other outcomes (Kozikoglu, 2019; Fung et al., 2021).

Given the rapid growth and popularity of FL across disciplines, there is a need to move beyond fragmented findings toward a comprehensive synthesis of evidence. An umbrella review, defined as a systematic synthesis of existing systematic reviews and meta-analyses, enables the consolidation of diverse strands of research to provide a higher-level overview of what is known and where gaps remain (Aromataris et al., 2015). This approach is particularly suited to FL, where multiple reviews exist but vary in scope, methodological rigor, and outcome measures. By integrating this body of evidence, an umbrella review can offer a more coherent understanding of FL's effectiveness across cognitive, affective, and interpersonal domains, thereby clarifying its pedagogical value and guiding future research and practice.

As an eclectic approach, the theoretical structure of FL can be understood through the arguments of learning and teaching theories—namely, behaviorist, cognitive, and social learning theories. The pedagogy of FL is not merely an inversion of instructional sequences, but a sophisticated educational approach grounded in a powerful confluence of these established theories. This is exemplified by the quality of social interaction, within a social learning context, that is facilitated by FL's technology-based opportunities, which are independent of time and space. Rather than operating in isolation, socio-constructivist theory, self-determination theory (SDT), and Kolb's experiential learning theory collectively provide a comprehensive framework that explains FL's efficacy and directly shapes its three-domain structure of cognitive, affective, and interpersonal engagement.

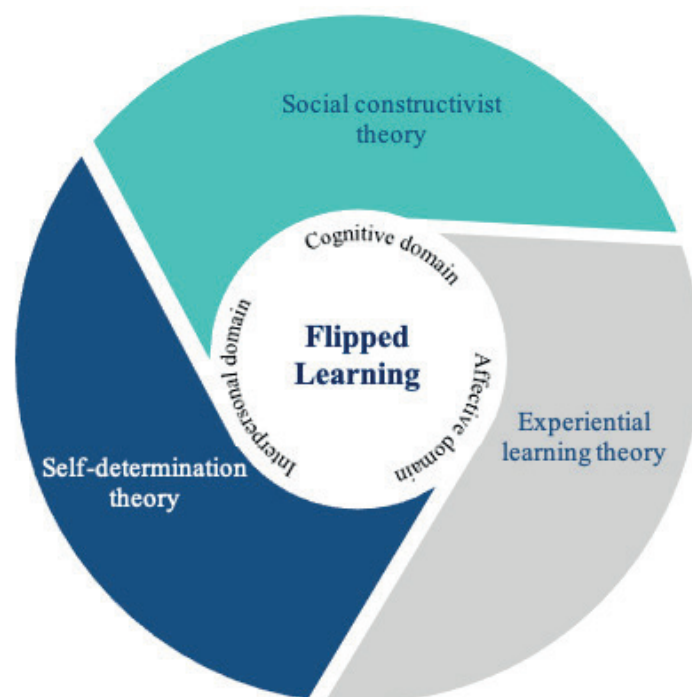
In the context of cognitive theory, which focuses on revealing an individual's cognitive qualities and potential, FL finds a strong theoretical anchor. However, social constructivism forms the foundational pillar of FL, positing that knowledge is not transmitted but co-constructed through collaborative social interactions and authentic tasks (Vygotsky, 1978). FL operationalizes this by transforming the classroom into a dynamic workshop where students, as active agents and knowledge constructors (King, 1993), engage in collaborative activities such as jigsaw techniques, peer instruction, and group problem-solving (Altemueller & Lindquist, 2017; Robertson, 2022). This environment fosters profound cognitive engagement by requiring students to articulate, defend, and refine their understanding. Studies consistently show that this constructivist approach in FL leads to improved conceptual understanding, better learning retention, transfer of learning (Kanjug et al., 2018), enhanced critical thinking, and superior academic performance (Cevikbas & Kaiser, 2020; Noguera Fructuoso et al., 2022; Sarwar et al., 2024). The interactive nature of FL, facilitated by technology, creates a learning procedure that evolves from social to individual and back again, significantly boosting cognitive and behavioral engagement (Lin & Pryor, 2019).

While constructivism explains the cognitive process, self-determination theory (SDT) elucidates the motivational engine that drives it. SDT argues that intrinsic motivation and high-quality engagement are fueled by the satisfaction of three basic psychological needs: autonomy, competence, and relatedness (Ryan & Deci, 2000). FL is uniquely designed to support these needs. The flexibility of pre-class learning materials supports autonomy by allowing students to control the pace, time, and place of their initial learning (Xu & Feng, 2017; Yang & Chen, 2024). The in-class focus on active, collaborative problem-solving directly fosters relatedness by creating a supportive community of peers and facilitators (Khosa & Burch, 2023; Zainuddin & Perera, 2017). Successfully navigating challenges with peer and instructor support then builds a sense of competence (Sergis et al., 2018). The satisfaction of these needs is intrinsically linked to the affective domain, generating increased intrinsic motivation, self-efficacy, learning satisfaction (Alonso et al., 2023; Wu et al., 2025), and a greater feeling of commitment and safety among students (Steen-Utheim & Foldnes, 2017). Research confirms that FL environments satisfying these needs lead to higher autonomous learning ability, more positive learning experiences, and improved outcomes (El-Bassiouny & El-Naggar, 2023; Oraif, 2018).

The interpersonal domain and the bridge from theory to practice are explicitly addressed by Kolb's experiential learning theory (ELT). ELT posits that learning is a cyclical process of concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 1984). FL perfectly aligns with this cycle: pre-class content delivery facilitates abstract conceptualization, while the reclaimed in-class time is dedicated to concrete experiences (e.g., simulations, experiments, real-world projects), reflective observation through group discussions, and active experimentation with new ideas (Joo, 2025; Williams et al., 2018). This process makes learning tangible and transformative, moving beyond rote memorization to develop essential practical and soft skills. Studies show that integrating FL with experiential activities leads to significantly higher final scores, better system skills, and greater learning satisfaction compared to traditional methods (Chen, 2021; Lantu et al., 2023). This approach is particularly effective in fostering interpersonal skills such as teamwork, communication, and intercultural sensitivity, as students engage in profound collaborative exchanges and real-world application (Phan & Ninh, 2024; Wu et al., 2023).

Taken together, the three lenses provide a coherent rationale for FL's impact across cognitive, affective, and interpersonal domains. Constructivist design and ELT-aligned cycles explain cognitive gains in retention, transfer, conceptual understanding, and performance (Chen, 2021; Lantu et al., 2023; Kanjug et al., 2018; Sarwar et al., 2024). SDT accounts for affective outcomes—intrinsic motivation, satisfaction, self-efficacy—by showing how FL's choice, scaffolding, and community satisfy autonomy, competence, and relatedness (Alonso et al., 2023; Chiu, 2021; Khosa & Burch, 2023; Sergis et al., 2018; Yang & Chen, 2024; Wu et al., 2025). Finally, the sociocultural emphasis on co-construction and the need for relatedness clarify interpersonal benefits: peer teaching, collaborative problem-solving, intercultural sensitivity, and communicative competence (Altemueller & Lindquist, 2017; Cevikbas & Kaiser, 2020; Gholami Pasand et al., 2021; Noguera Fructuoso et al., 2022; Zainuddin & Perera, 2017; Zhang & Zhou, 2019; Wu et al., 2023).

As an eclectic approach within the context of learning and teaching theories, the theoretical framework of FL is shown in Figure 1. In conclusion, this theoretical synthesis demonstrates that the three theories work synergistically to support FL pedagogy across all three domains. Social constructivist theory provides the foundational framework for collaborative knowledge construction, SDT supplies the motivational mechanisms necessary for sustained engagement, and experiential learning theory offers the practical application context that transforms theoretical knowledge into meaningful understanding. Together, these theories create a robust theoretical foundation that validates the multidimensional approach to examining FL effectiveness across cognitive, affective, and interpersonal learning outcomes for our umbrella review.



**Figure 1.** Conceptual framework of flipped learning approach

The impact of technology on learning and teaching processes in the first quarter of the 21st century, particularly through devices that eliminate dependence on time and space, has enabled a transformation of the educational phenomenon into a new form. This transformation in the learning and teaching process has necessitated a reconfiguration of its actors, their roles, the content, the forms of interaction and communication, as well as the very concepts of time and space. As one of the approaches encompassing this global, technology-based transformation, the primary aim of this research is to examine the arguments of FL and, through the results of its practical application, to scrutinize the outcomes of FL reflected in global-scale research. This synthesis focuses on both the core cognitive, social, and affective behavioral outcomes of the learning-teaching process, as well as on different disciplinary areas (such as M&H fields, SS, and STEM fields) that influence the process due to their varied content, practices, and methodologies.

Guided by the overarching question-“What are the learning outcomes of globally conducted FL-focused studies based on different methodologies and across different disciplinary areas?”-this research seeks to answer the following specific questions regarding the FL-focused studies:

1. What are the key (e.g., publication date, field, review type, learner type) and methodological characteristics (e.g., inclusion criteria, quality assessment, data analysis methods) of the included reviews?
2. What are the general learning outcomes reported in the included reviews in terms of discipline and review type?
3. What are the disciplinary learning outcomes reported in the included reviews?

## **METHOD**

An umbrella review, defined as the overview of reviews and review synthesis, specifically refers to a review of compelling evidence from a series of other systematic reviews (Becker & Oxman, 2011). This type of review offers a valuable tool for gaining a comprehensive understanding of the existing research landscape within a specific field. From a methodological standpoint, an umbrella review's focus lies in identifying and examining existing reviews, rather than directly searching for and analysing primary research studies (Hartling et al., 2014).

The reasons for conducting an umbrella review are manifold. The primary reason is to provide a concise summary of existing research syntheses pertinent to a specific topic and not any further “meta-analysis” of the results of these publications (Becker and Oxman 2011). In other words, reviews within an umbrella review may encompass analyses of evidence for various interventions addressing the same issue or condition (Whittemore et al., 2014). An umbrella review prioritizes clarity and ease of interpretation, focusing on the results and findings presented in the included reviews. Besides, researchers conducting an umbrella review are advised against retrieving primary studies (original research) from the included systematic reviews for the purpose of accessing additional data (Smith et al., 2011). Both quantitative and qualitative findings should be presented in a tabular format, accompanied by supporting text, including details such as the number of studies reporting each outcome, the total number of included studies, and the heterogeneity of the results.

Considering the increasing adoption of FL pedagogy and the numbers of systematic reviews conducted on FL practices across fields, conducting an umbrella review is particularly timely and well-suited for evidence synthesis regarding FL practices across fields, identifying existing research gaps and cultivating a nuanced understanding of FL practices on domainial and disciplinary outcomes (Grant & Booth, 2009; Ioannidis, 2009). To enhance the clarity of the existing research evidence base concerning FL practices (Lockwood & White, 2012), we employed a systematic approach to extract descriptive and thematic data. Initially, detailed descriptive data were extracted from the included reviews. Subsequently, a summary of these reviews is presented (Figures 3 and 4), followed by a thematic analysis. This comprehensive methodology enables the provision of valuable insights into the effectiveness of FL across diverse contexts while simultaneously identifying areas that warrant further investigation.

### **Search Strategy**

Following the PRISMA guidelines (Moher et al., 2015), we conducted a comprehensive search for relevant studies in English published up to June 1, 2024. We utilized databases such as Web of Science, EBSCO Academic Search Complete, and ERIC, employing search terms like “flipped learning”, “flipped classroom”, “inverted learning”, and “inverted classroom”. Due to the limitations of specific criteria for retrieving review articles (Holly, 2017), we manually screened titles and abstracts for terms like “review”, “systematic review”,

and “meta-analysis”. We also employed reference tracking and free text searches to further enhance our evidence collection on FL pedagogy (Greenhalgh & Peacock, 2005). Our initial search yielded 7,419 FL-related studies. After removing duplicates using EndNote x9, 2,946 studies remained for title and abstract screening. Two reviewers independently evaluated these studies, ultimately selecting 118 reviews for in-depth full-text review. This process ensured the identification of relevant reviews for our review.

**Inclusion and Exclusion Criteria**

This review included exclusively qualitative, meta-analytic, and mixed-method reviews, excluding 2,828 primary studies on FL approach. We further excluded reviews that were integrative, scoping, narrative, literature-based, non-systematic, or critical appraisals (n = 25), overly broad (n = 5), lacked defined outcomes (n = 12), or focused on non-educational contexts (n = 13). During the appraisal phase, to manage overlapping primary studies, we employed a pre-specified protocol using the Corrected Covered Area (CCA) metric to quantitatively assess overlap between review pairs (Pieper et al., 2014). Reviews with CCA values exceeding 10% were considered to have high overlap, and we selected the most comprehensive, recent, and methodologically rigorous review (based on AMSTAR 2 scores) from each overlapping pair (Shea et al., 2017). This process resulted in the exclusion of nine reviews to prevent double-counting of primary studies and ensure the validity and reproducibility of our umbrella review conclusions. Further, reviews not meeting the Joanna Briggs Institute’s appraisal checklist (n = 4) were also excluded (Aromataris et al., 2020). Our selection criteria also required reviews to align with Bishop and Verleger’s (2013) definition of FL, which necessitates pre-class video/audio resources and interactive in-class activities. This led to the exclusion of five qualitative reviews with differing focuses and methods, and four reviews lacking disciplinary outcome reports.

To ensure the robustness of our analysis, we applied two additional criteria for reporting effect sizes in the extant meta-analytic reviews. Thus, we reported only findings of meta-analytic reviews with at least five studies as to FL outcomes, ensuring sufficient statistical power (Jackson & Turner, 2017), together with a minimum threshold of  $d = 0.40$ , indicating a significant educational impact (Hattie, 2012). These criteria led to the exclusion of reporting meta-analytic findings in two mixed-method reviews (Lo et al., 2017; Lo & Hew, 2019). Ultimately, 41 reviews met our rigorous standards and were included in our narrative synthesis (see Fig. 2).

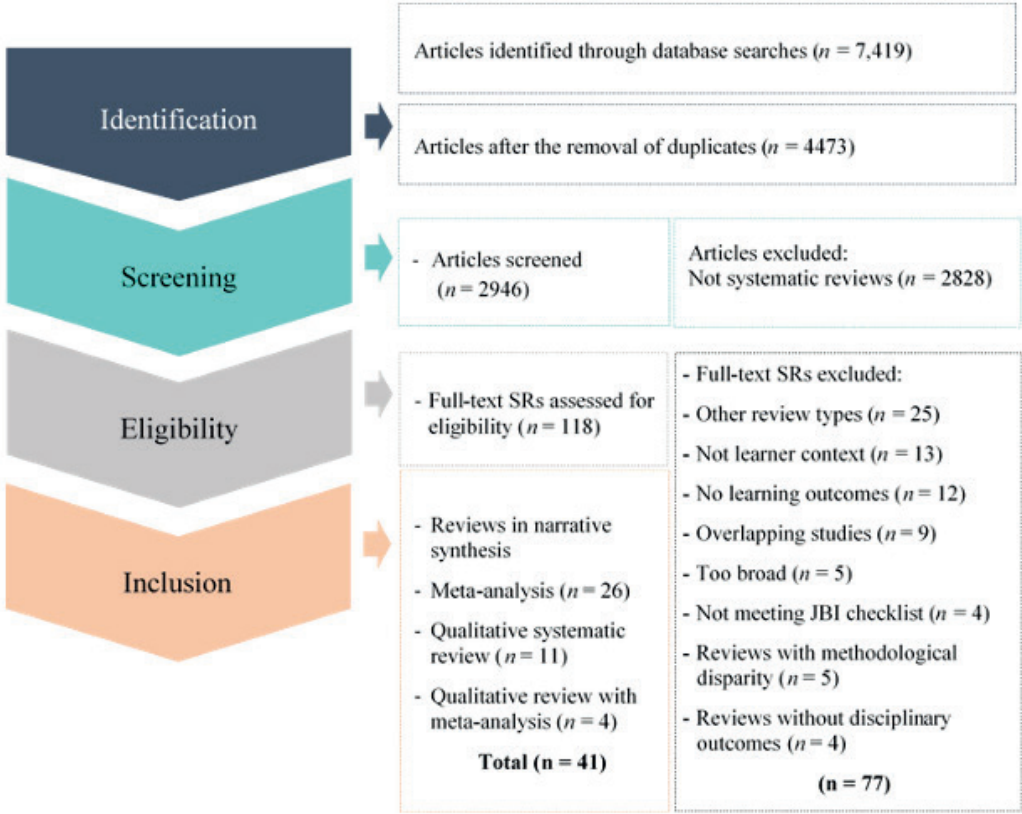


Figure 2. PRISMA search methodology

## Criteria Data Extraction and Synthesis

Following the JBI umbrella review protocol (Aromataris et al., 2015), we created a data extraction template to capture key characteristics of the included studies. We systematically recorded descriptive data such as author/year, review type, coverage period, databases used, number of studies included, field of review, learner type, language, and geographical regions in a Word spreadsheet (available online). Figure 3 provides a summary of the descriptive analysis. To assess the methodological quality of the included reviews, we developed a checklist based on established practices in meta-analysis research, including resources like the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2023) and the R-AMSTAR tool (Kung et al., 2010). This checklist covered the following key areas:

- Sources of primary studies: We examined whether authors searched for unpublished (grey) literature and if their searches were language restricted.
- Quality assessment: We evaluated how authors assessed the quality or risk of bias in the primary studies and documented the tools used for this assessment.
- Analysis methods: We noted the statistical models (fixed-effects or random-effects) used for data aggregation and the strategies employed to analyze heterogeneity in the meta-analyses.
- Publication bias: We recorded the methods authors used to assess publication bias.

Prior to thematic analysis, we immersed ourselves in the data to gain a deep understanding of the content and develop preliminary insights relevant to the research questions (Ayres, 2008), which provided us with a strong foundation for the thematic analysis process. For the 26 meta-analytic reviews, we extracted quantitative findings including effect sizes, which were systematically coded according to the cognitive, affective, and interpersonal learning domains (Borenstein et al., 2009; Higgins et al., 2019). The 11 qualitative systematic reviews were analyzed using thematic analysis. For the four mixed-method reviews included, we adopted a convergent synthesis approach, treating quantitative and qualitative strands according to their methodological nature while maintaining integration throughout the analysis (Hong et al., 2018; Lizarondo et al., 2020). Quantitative strands were analyzed alongside meta-analytic findings, while qualitative strands followed the same thematic approach used for qualitative systematic reviews (Booth et al., 2016; Thomas et al., 2004). We employed a convergent approach where quantitative and qualitative findings were analysed simultaneously and integrated to provide complementary insights, with quantitative data informing the scope of themes while qualitative insights provided contextual understanding, ensuring methodological rigor and analytical coherence across all review types (Creswell & Clark, 2017).

We then carefully examined the 41 reviews to identify initial themes related to the research questions. Initially emerged themes were further analyzed and categorized into appropriate learning domains. To ensure the validity of the themes, we applied them back to our dataset, confirming that they provided a coherent narrative that effectively addressed our research questions (Braun & Clarke, 2006; Neuendorf, 2018; Saldaña, 2016). Through iterative coding and recoding, we categorized themes under relevant learning domains, achieving consensus on code and theme formulation (Kynğäs et al., 2020). This process ensured reliability with a 95% inter-rater agreement, with discrepancies resolved through discussion.

## FINDINGS

Aligned with the research questions and methodology, the findings are presented under the headings of key descriptive and methodological characteristics of the reviews and multidisciplinary learning outcomes.

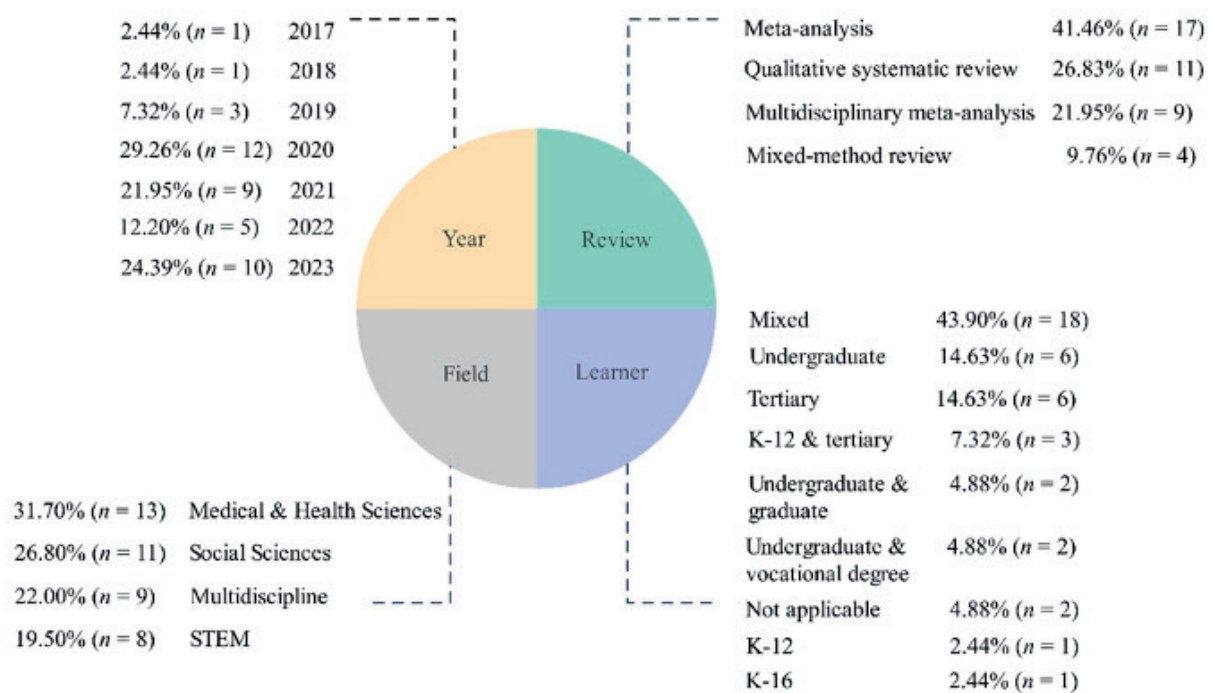
### Key Descriptive and Methodological Characteristics of the Reviews

Examining the general descriptive characteristics of the reviews within our umbrella review, the distribution of publication years indicates a marked rise in FL research in recent years, with most reviews published in 2020. This suggests a growing interest in assessing the effectiveness of FL and its potential to transform learning experiences (see Fig. 3). After 2020, a noticeable decline occurred in 2022 ( $n = 5$ ), possibly reflecting a return to conventional teaching methods as schools started to reopen, or a saturation in the research area following a period of intense investigation. The 2023 resurgence suggests renewed interest in FL research, potentially driven by its long-term benefits or the need for further refinement and understanding of FL practices.

The prevalence of reviews in M&H ( $n = 13$ ), SS (Social Sciences) ( $n = 11$ ), and STEM ( $n = 8$ ) fields indicates the broad applicability of FL across diverse disciplines. This aligns with FL's versatility and adaptability

to different learning contexts and subject areas. Reviews with multidiscipline ( $n = 9$ ) also highlight the adaptability of FL approach in various educational settings across fields. However, the lesser representation of STEM fields might indicate a delay in adopting such methods in areas typically associated with more rigid and sequential teaching approaches. Although the dominance of disciplinary ( $n = 17$ ) and multidisciplinary ( $n = 9$ ) meta-analyses reflects a focus on quantitative synthesis of research findings to assess the overall effectiveness of FL practices, qualitative ( $n = 11$ ) and mixed-method reviews ( $n = 4$ ) highlight the importance of exploring the nuances and contextual factors affecting FL implementation and outcomes.

Most reviews focused on mixed learner groups ( $n = 18$ ) including learners from K-12 to graduate levels, suggesting the potential benefits of FL for learners of various backgrounds and levels. This wide-ranging inclusion could be due to the versatility and relevance of FL across diverse learner demographics (Persky & McLaughlin, 2017). The equal representation of undergraduate ( $n = 6$ ) and tertiary education ( $n = 6$ ) emphasizes the importance of FL in higher education. It is worth noting that 'tertiary level' in this context spans learners from undergraduate, graduate, to vocational degree levels. As the field continues to develop, it would be advantageous for future research to investigate the effects of FL among different learner groups, including those in K-12 ( $n = 1$ ) and K-16 ( $n = 1$ ) settings.



**Figure 3.** The overview of the reviews by year, field, review, and learner type

When considering the methodological characteristics of the reviews in our umbrella review, our analysis of 11 qualitative reviews revealed that only two (Aljaber et al., 2023; Arslan, 2020) searched the grey literature in English. Ten reviews conducted their search in English (e.g., Cevikbas & Kaiser, 2022; Kazeminia et al., 2022), while one searched in Chinese (Li et al., 2020). Similarly, ten reviews included articles written in English (e.g., Ozbay & Cinar, 2021; Senali et al., 2022), and one review included articles in Chinese (Li et al., 2020). Three of the qualitative reviews conducted methodological quality assessments on the included studies. Aljaber et al. (2023) used the Medical Education Research Quality Instrument (MERSQI) and the Joanna Briggs Institute (JBI) checklist, Kazeminia et al. (2022) utilized the Mixed Methods Appraisal Tool (MMAT) checklist, and Ozbay & Cinar (2021) used the Quality Assessment Tool for Quantitative Studies (QATQS).

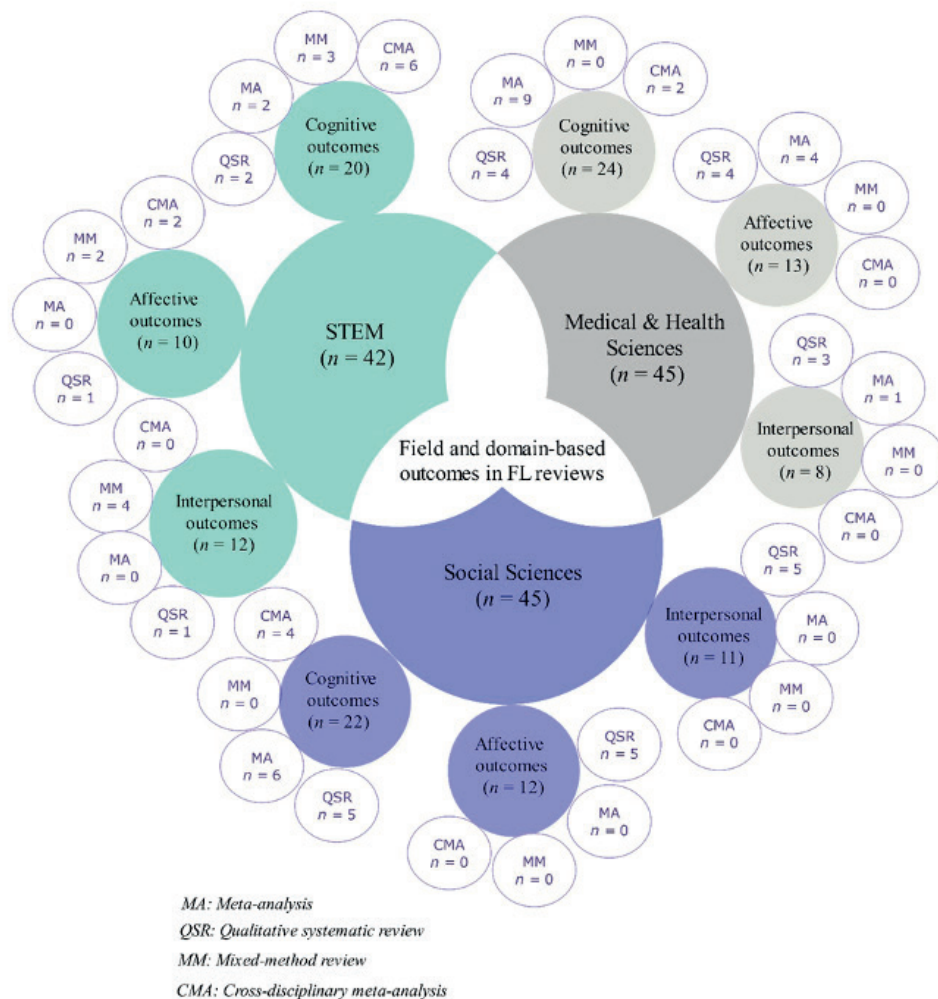
Our analysis of 26 meta-analytic reviews, together with 4 mixed-method reviews, indicated that 12 reviews searched for grey literature in English (e.g., Aybirdi et al., 2023; Bredow et al., 2021). Two reviews searched in both English and Turkish (e.g., Tural & Yazar, 2021), one in English and Chinese (e.g., Cui et al., 2023), and one in English and Korean (e.g., Park & Suh, 2021). Eight reviews included articles written in English (e.g., Naing et al., 2023; Ni et al., 2023), three included articles in both English and Chinese (e.g., Peng et al., 2022), and two included articles written in Chinese (e.g., Li et al., 2020). Two meta-analytic reviews included articles regardless of language (e.g., Xiao et al., 2022). Eleven meta-analyses conducted

methodological quality assessments on included studies. Eight used the Cochrane risk-of-bias tool (e.g., Li et al., 2022; Zhang et al., 2021; Zhu, 2021), while the remaining three used different tools: the EPHPP Quality Assessment Tool (Cui et al., 2023), the Newcastle–Ottawa Scale (Ge et al., 2020), and the What Works Clearinghouse experimental standards (Zhang et al., 2021).

All meta-analyses (including mixed-method reviews) utilized a random-effects model for analysis and assessed heterogeneity, predominantly using the I<sup>2</sup> statistic. Nearly all (90%, n = 27) investigated heterogeneity causes, often through subgroup analyses. All meta-analytic reviews assessed publication bias. Three of them relied solely on visual inspection of funnel plots (Jang & Kim, 2020; Naing et al., 2023; Xiao et al., 2022), while 27 used additional methods like Fail-Safe N, Egger’s test, Begg’s test, and trim-and-fill. The number of moderator variables in the meta-analytic reviews varied considerably, from none (e.g., Li et al., 2022; Sopamena et al., 2023; Turan, 2023) to as many as 12 variables (e.g., Zheng et al., 2020). A wide range of moderator variables employed across existing meta-analytic reviews presents a challenge for synthesis and interpretation of moderator analyses. Thus, our review prioritizes “discipline” as the key moderator due to its consistent presence and acknowledged significance across various fields.

### General Distribution of Multidisciplinary Learning Outcomes in FL Reviews

Our analysis of 41 reviews on FL practices emerged a total of 132 codes across cognitive, affective, and interpersonal learning domains within STEM, M&H, and SS fields. The cognitive domain accounted for half of reported outcomes (50%, n = 66), reflecting a traditional mindset that emphasizes academic success and cognitive skills in educational settings. The affective domain, encompassing emotions, attitudes, and motivation, was less frequently addressed (26.52%, n = 35). Similarly, the interpersonal domain, emphasizing social engagement and collaboration, received the least attention (23.48%, n = 31) across fields (see Fig. 4).



**Figure 4.** Overall distribution of outcomes by field and domain in FL reviews

Analysis of STEM related research revealed a prevalent emphasis on cognitive outcomes, with 20 codes identified primarily through cross-disciplinary meta-analyses ( $n = 6$ ) and mixed-method reviews ( $n = 3$ ). This suggests a growing body of research investigating the impact of FL on cognitive skills development in STEM fields. However, affective and interpersonal outcomes received less attention, with only 10 and 12 codes identified, respectively. Notably, no meta-analytic reviews focused solely on affective or interpersonal outcomes in STEM, indicating a need for further quantitative research in these areas.

Like STEM fields, the reviews in the field of M&H sciences underscored cognitive outcomes, with 24 codes identified mainly through meta-analyses ( $n = 9$ ) and qualitative systematic reviews ( $n = 4$ ). The high number of meta-analyses suggests a mature research trajectory in M&H sciences, aiming to consolidate findings and establish best practices. Affective outcomes were also explored, with 13 codes identified through both meta-analyses ( $n = 4$ ) and qualitative systematic reviews ( $n = 4$ ). However, interpersonal outcomes received less attention, with 8 codes identified through qualitative systematic reviews ( $n = 5$ ) and only one meta-analysis contributing to this domain. This highlights the need for more quantitative and mixed-methods research on FL's impact on interpersonal skills development in M&H fields. Key insights from the analysis of FL research in SS fields highlights a notable imbalance in the investigation of domainial outcomes in terms of review type. While 22 cognitive outcomes were identified across meta-analyses ( $n = 6$ ), qualitative systematic reviews ( $n = 5$ ), and cross-disciplinary meta-analytic reviews ( $n = 4$ ), affective and interpersonal outcomes were explored solely through qualitative systematic reviews, with 12 and 11 codes identified, respectively.

The overall descriptive analysis across fields underscores a pronounced emphasis on cognitive outcomes within FL research, constituting over half of the reported outcomes across fields. This dominant cognitive focus reflects traditional educational paradigms, where knowledge acquisition and cognitive skills development are paramount. In contrast, the affective and interpersonal domains, though integral to holistic education, were relatively underexplored, accounting for a smaller portion of the outcomes. This disparity highlights a potential research gap, especially considering the growing emphasis on social-emotional learning and 21st century skills crucial for modern education and workplaces.

## **Learning Outcomes in STEM Fields**

### **Cognitive Domain Outcomes in STEM**

Across 15 reviews spanning various STEM disciplines, a consistent pattern emerges demonstrating FL's positive impact on academic achievement. Meta-analytic evidence reveals moderate effect sizes consistently ranging from 0.40 to 0.72 across mathematics and science education (Guler et al., 2023; Dogan et al., 2021; Sopamena et al., 2023; Turan, 2023), with qualitative reviews providing convergent evidence across 127 empirical studies (Cevikbas & Kaiser, 2023; Wright & Park, 2021). A notable trend emerges regarding educational level, with elementary education demonstrating consistently larger effect sizes ( $g = 0.97$  to  $1.16$ ) compared to higher education levels (Guler et al., 2023; Tural & Yazar, 2021), suggesting FL's particular effectiveness with younger learners in STEM contexts.

Cross-disciplinary meta-analyses ( $n = 6$ ) reveal FL's broad applicability across STEM fields, with effect sizes consistently falling within the moderate to large range. Technology-related disciplines show particularly strong outcomes ( $g = 0.52$  to  $0.80$ ), while traditional STEM subjects like mathematics and science demonstrate stable moderate effects ( $g = 0.43$  to  $0.68$ ). This pattern suggests FL's universal applicability across diverse STEM contexts while maintaining consistent effectiveness levels.

Beyond academic performance, reviews consistently identify three key cognitive outcomes across STEM disciplines. Autonomous learning emerges as a universal benefit, with convergent evidence from mathematics (Cevikbas & Kaiser, 2023) and engineering education (Lo & Hew, 2019). Conceptual

understanding represents another consistent trend, supported across mathematics and science education reviews, with approximately one-third of studies in mathematics reporting this outcome (33 of 97 studies). Higher-order thinking skills, particularly reflective thinking, constitute a third consistent pattern, though less frequently reported across disciplines. The self-paced learning mechanism emerges as a critical factor underlying these cognitive benefits, with reviews consistently attributing cognitive gains to students' ability to control learning pace through video lectures and review materials as needed (Cevikbas & Kaiser, 2023; Lo & Hew, 2019).

### **Affective Domain Outcomes in STEM**

Several reviews ( $n = 5$ ) highlighted the positive impact of FL on affective domain in STEM disciplines. A qualitative review in mathematics (Cevikbas & Kaiser, 2023) identified student satisfaction and enjoyment as a key outcome, with 17 of 97 studies reporting high satisfaction due to flexibility, accessibility of materials, and active learning opportunities. This finding is echoed in a cross-disciplinary meta-analysis (Bredow et al., 2021) with an effect size of 0.53 across 10 studies, suggesting potential for enhancing satisfaction in flipped mathematics courses. Similarly, a mixed method review in science (Dogan et al., 2023) reported the creation of "amusing" learning environments, despite the unclear number of studies reporting this outcome. A cross-disciplinary meta-analysis (Strelnan et al., 2020b) also reported increased satisfaction ( $g = 0.81$ ,  $k = 7$ ) in the field of ICT.

Beyond satisfaction and enjoyment, positive thought and interest were identified in the qualitative review of Cevikbas and Kaiser (2023) and two mixed-method reviews in science education (Dogan et al., 2023; Turan, 2023). Increased student motivation emerged as another positive outcome in two reviews (Cevikbas & Kaiser, 2023; Turan, 2023), with 17 and 6 primary studies reporting this outcome, respectively. Finally, Cevikbas and Kaiser (2023) identified self-confidence and efficacy as an affective outcome in 14 of 97 studies.

### **Interpersonal Domain Outcomes in the Field of STEM**

Our study highlights that FL practices in STEM disciplines demonstrate a positive impact on various interpersonal outcomes including peer learning and collaboration, active in-class time and feedback, interaction, and engagement. A finding emerged across multiple reviews ( $n = 4$ ) is the enhancement of peer learning and collaboration. Cevikbas and Kaiser (2023) found evidence in 23 out of 97 mathematics education studies, while mixed-method reviews in engineering (Lo & Hew, 2019) and mathematics (Lo et al., 2017) reported similar benefits in 10 out of 31 and 33 out of 61 studies, respectively. Additionally, a mixed method review in science education (Dogan et al., 2023), identified the formation of a social learning environment as a key outcome, without a mention of specific number of studies reporting this outcome. The ability to engage in discussions, share understanding, and participate in small group activities during in-class time were noted as contributing factors in these reviews.

Furthermore, in several reviews ( $n = 4$ ) FL appears to foster active in-class time and feedback. Cevikbas and Kaiser (2023) reported such findings in 12 out of 97 mathematics education studies, with evidence also present in mixed-method reviews across engineering, mathematics, and science education (Dogan et al., 2023; Lo et al., 2017; Lo & Hew, 2019). The freed-up in-class time allows for increased instructor feedback, scaffolding, and one-on-one support in FL environments. Enhanced interaction between students and instructors was another emerged outcome, as noted by Cevikbas and Kaiser (2023) in 8 out of 97 studies in mathematics education and supported by mixed method review in science education (Dogan et al., 2023). Finally, increased student engagement was identified as an interpersonal outcome, with evidence present in 28 out of 97 mathematics education studies (Cevikbas & Kaiser, 2023) and 5 out of 64 science education studies (Turan, 2023), particularly using active learning tools like clickers.

**Table 1.** Disciplinary and learning outcomes in STEM fields

Domain	Theme	Field	QSR	MA	MM	CMA	Sample review	Effect size & number	
Cognitive	Academic achievement (35.74 %, n = 15)	Mathematics	1	2		1	Guler et al., 2023	g = 0.40, k = 37	
		Science			2	2	Dogan et al., 2023	g = 0.72, k = 30	
		Science & mathematics	1				Wright & Park, 2022	n = 10	
		Science & engineering				1	Wagner et al., 2021	d = 2.00, k = 5	
		Engineering				1	Strelan et al., 2020a	g = 0.72, k = 15	
		Engineering & technology				1	Zheng et al., 2020	g = 0.69, k = 16	
		Computer & information technology				2	Total & Yazar, 2021	g = 0.72, k = 26	
		Technology				1	Bredow et al., 2021	g = 52, k = 23	
	Autonomous learning (4.76 %, n = 2)	Mathematics	1				Cevikbas & Kaiser, 2023	n = 16	
		Engineering			1		Lo & Hew, 2019	n = 15	
	Conceptual learning (4.76 %, n = 2)	Science			1		Dogan et al., 2023	n = NA	
		Mathematics	1				Cevikbas & Kaiser, 2023	n = 33	
	Higher order thinking skills (2.38 %, n = 1)	Mathematics	1				Cevikbas & Kaiser, 2023	n = 5	
	Affective	Satisfaction & enjoyment (9.52 %, n = 4)	Science			1		Dogan et al., 2023	n = NA
			Mathematics	1		1		Bredow et al., 2021	g = 0.53, k = 10
ICT					1		Strelan et al., 2020b	g = 0.81, k = 7	
Positive thought & interest (7.14 %, n = 3)	Science			2			Turan, 2023	n = 20	
	Mathematics	1					Cevikbas & Kaiser, 2023	n = 27	
Learner motivation (4.76 %, n = 2)	Science			1			Turan, 2023	n = 6	
	Mathematics	1					Cevikbas & Kaiser, 2023	n = 17	
Self-confidence & efficacy (2.38 %, n = 1)	Mathematics	1					Cevikbas & Kaiser, 2023	n = 14	
Interpersonal	Peer learning & collaboration (9.52 %, n = 4)	Mathematics	1		1			Lo et al., 2017	n = 33
		Engineering			1			Lo & Hew, 2019	n = 10
		Science			1			Dogan et al., 2023	n = NA
	Active in-class time & feedback (9.52 %, n = 4)	Mathematics	1		1				Cevikbas & Kaiser, 2023
Science				1				Dogan et al., 2023	n = NA
Engineering				1				Lo & Hew, 2019	n = 8
Interaction (4.76 %, n = 2)	Mathematics	1						Cevikbas & Kaiser, 2023	n = 8
	Science			1				Dogan et al., 2023	n = NA
	Engagement (4.76 %, n = 2)	Mathematics	1					Cevikbas & Kaiser, 2023	n = 28
		Science			1			Turan, 2023	n = 5
Total	100 %, n = 42		13	2	16	11			

Note: QSR = Qualitative systematic review; MA = Meta analysis; MM = Mixed-method review; CMA = Cross-disciplinary meta-analysis

In QSRs, “n” indicates the number of primary studies reporting outcomes

In MAs, Cohen’s d (standardized mean difference) indicates effect size: 0.2 small, 0.5 medium, 0.8 large.

In MAs, Hedges’ g (small-sample-adjusted standardized mean difference) indicates effect size: 0.2 small, 0.5 medium, 0.8 large.

## Learning Outcomes in M&H Fields

### Cognitive Domain Outcomes in M&H Fields

Across various M&H science fields, the results indicate that FL has the potential for promoting students' academic achievement. Evidence from qualitative systematic reviews ( $n = 4$ ), meta-analyses ( $n = 9$ ), and cross-disciplinary reviews ( $n = 3$ ) highlights this achievement in nursing ( $n = 4$ ), pharmacy ( $n = 3$ ), health professions education ( $n = 2$ ), medical microbiology ( $n = 1$ ), radiology ( $n = 1$ ), Chinese medicine ( $n = 1$ ), and anatomy ( $n = 1$ ).

Reviews in nursing education ( $n = 4$ ) reported significant improvements in theoretical knowledge (SMD = 1.33,  $k = 32$ ) and skill scores (SMD = 1.58,  $k = 23$ ) compared to traditional classrooms (Li et al., 2020). This is further supported by another meta-analysis revealing an overall effect size ( $g = 0.68$ ,  $k = 18$ ) in favour of FL (Park & Suh, 2021). Additionally, FL significantly enhances operational performance (SMD = 1.79,  $k = 22$ ) and practical ability (ES = 1.47,  $k = 6$ ) among nursing students (Xu et al., 2019). Qualitative evidence also points to FL pedagogy effectively supporting learning and improving teaching quality, as evidenced by pre-test and post-test score comparisons (Ozbay & Cinar, 2021). Although less extensively studied, there is also an indication of positive effects ( $g = 0.53$ ,  $k = 8$ ) on the development of clinical competence (Park & Suh, 2021). The reviewed meta-analyses on FL in nursing education ( $n = 3$ ), while indicating potential benefits, present methodological considerations worth noting. Xu et al. (2019) identified potential overrepresentation of positive outcomes in their sample. Study variations in design, populations, and implementation approaches (Li et al., 2020; Park & Suh, 2021; Xu et al., 2019) reflect the diverse contexts in which FL was implemented.

Similar positive outcomes were observed in pharmacy education, with significant improvements in theoretical and experimental test scores (Aljaber et al., 2023; Cui et al., 2023; Peng et al., 2022). A qualitative review of 35 studies highlighted improved exam performance and overall student success, primarily in US pharmacy programs (Aljaber et al., 2023). This is supported by meta-analyses showing substantial increases in theoretical test scores (SMD = 1.08,  $k = 10$ ) compared to traditional lecture-based learning, particularly among Chinese undergraduates (Peng et al., 2022). Furthermore, improvements in experimental test scores (MD = 6.62,  $k = 6$ ) were also observed with FL approach. A meta-analytic review of 22 US-based studies, with historical control research designs, provides further evidence of the positive impact of FL on pharmacy student performance (SMD = 1.30,  $k = 22$ ), predominantly based on exam score comparisons (Cui et al., 2023).

In health professions education, a qualitative review focusing on Chinese undergraduate health professions students revealed that FL significantly improved knowledge and skills acquisition compared to traditional lectures (Li et al., 2020). Similarly, a meta-analysis involving undergraduate health professions students from diverse geographical locations found that FL enhanced academic performance (SMD = 0.54,  $k = 44$ ), though study variations influenced the strength of conclusions (Naing et al., 2023). Cross-disciplinary meta-analyses further support these findings, indicating a moderate to large effect size of FL on academic achievement in health sciences (Bredow et al., 2021,  $g = 0.46$ ,  $k = 46$ ) and a moderate effect on learning in health-care programs (Tutal & Yazar, 2021,  $g = .0.97$ ,  $k = 12$ ).

The positive impact of FL on academic achievement extends to various M&H disciplines. For example, meta-analyses in medical microbiology and radiology education demonstrated the positive impact of FL on student learning outcomes. Li et al. (2022) reported that FL significantly improved both theoretical (SMD = 0.56,  $k = 11$ ) and practical (SMD = 0.70,  $k = 5$ ) examination scores in medical microbiology compared to traditional classrooms. Similarly, Ge et al. (2020) found that FL enhanced theoretical (SMD = 1.12,  $k = 11$ ) and practical (SMD = 2.59,  $k = 10$ ) examination scores in radiology education. Notably, both meta-analyses included studies with diverse student populations and research designs, and neither indicated significant publication bias, strengthening the reliability of the findings. Similarly, a qualitative review of anatomy education studies indicated improved exam scores among students exposed to FL practices (Kazeminiya et al., 2022). In Chinese medicine education, FL demonstrably enhanced theoretical knowledge acquisition (MD = 6.61,  $k = 23$ ) and practical skills (MD = 8.8,  $k = 6$ ), with notable variation in implementation approaches (Xiao et al., 2020). Furthermore, a cross-disciplinary meta-analysis (Tutal & Yazar, 2021) suggests a large effect ( $g = 1.47$ ,  $k = 6$ ) of FL on student learning in physical education.

Moving beyond mere academic achievement, FL demonstrates a positive impact on various cognitive skills within M&H sciences. Extant reviews highlighted significant improvements in areas such as autonomous learning, conceptual understanding, and higher-order thinking. For instance, a qualitative systematic review (Li et al., 2020) focusing on health professions education found that FL enhanced self-learning abilities, conceptual understanding, and cognitive thinking skills like clinical reasoning and problem-solving. This finding is supported by several meta-analyses. (Ge et al., 2020) reported that FL in radiology promotes autonomous learning (OR = 1.70, k = 5) and consolidation of knowledge (OR = 1.38, k = 5), while (Xiao et al., 2022) found similar results for self-learning ability (RR = 1.52, k = 5) in Chinese medicine education. Additionally, two meta-analytic reviews exploring the impact of FL on critical thinking skills in nursing education show promising results. For example, Xu et al. (2019) found a significant improvement in nursing students' ability to think and analyse (ES = 1.62, k = 5) with FL compared to traditional methods. Similarly, Park and Suh (2021) reported a positive effect of FL on developing critical thinking skills in nursing students (g = 0.87, k = 18).

### **Affective Domain Outcomes in M&H Sciences**

Drawing on qualitative (n = 4) and meta-analytic reviews (n = 4), our research highlights several key themes. Satisfaction and enjoyment emerged as prominent affective outcomes associated with FL practices in qualitative reviews conducted in nursing education (Ozbay & Cinar, 2021) health professions education (Li et al., 2020), anatomy (Kazeminia et al., 2022), and pharmacy education (Aljaber et al., 2023). This is supported by meta-analytic reviews across diverse M&H disciplines, including nursing (Park & Suh, 2021; Xu et al., 2019), health professions education (Naing et al., 2023), and radiology (Ge et al., 2020). Meta-analyses exploring student satisfaction with FL pedagogy report positive outcomes across varying contexts. Park and Suh (2021) found enhanced satisfaction among nursing students (g = 0.79, k = 6), while Xu et al. (2019) reported an even stronger effect (ES = 1.39, k = 22). Naing et al. (2023) identified more modest improvements in health professional programs (SMD = 0.48, k = 8), suggesting context-dependent effectiveness. Finally, Ge et al. (2020) reported high student satisfaction and appraisal for FL courses compared to traditional lectures in radiology education (OR = 1.70, k = 7).

Furthermore, evidence suggests that flipped courses enhance students' positive thought and interest towards learning in pharmacy, anatomy, and nursing education (Aljaber et al., 2023; Kazeminia et al., 2022; Xu et al., 2019). While limited in scope, existing research suggests that FL can positively influence student interest and preference towards learning within M&H sciences. A qualitative review in anatomy found that FL increased student interest towards learning, although this was only reported in 5 studies (Kazeminia et al., 2022). Similarly, a meta-analysis of nursing education revealed a significant positive effect of FL on student participation interest (ES = 1.58, k = 15), based on available evidence (Xu et al., 2019). In pharmacy education, Aljaber et al. (2023) analysed 35 studies, primarily from the US, and found that 26 of them reported a clear preference for FL through students' surveys and written comments, further supporting the potential of FL to enhance student engagement and motivation.

While evidence is limited, qualitative reviews (n = 2) in pharmacy and anatomy education suggest a potential positive influence of FL on student self-confidence and self-efficacy. Aljaber et al. (2023) found that 5 out of 35 reviewed pharmacy education studies reported a positive student perception of improved self-confidence with FL approach. Similarly, Kazeminia et al. (2022) noted that 3 out of 26 anatomy education studies observed an increase in self-confidence levels among students exposed to FL, attributing this to enhanced interactions between learners, teachers, and learning materials. Despite these positive findings related to affective domain outcomes, a notable gap exists in understanding the impact of FL on learner motivation, as this area remains unexplored in the extant M&H reviews.

**Table 2.** Disciplinary and learning outcomes in M&H Sciences

Domain	Theme	Discipline	QSR	MA	MM	CMA	Sample review	Effect size & number
Cognitive	Academic achievement (35.56%, n = 16)	Nursing	1	3			Xu et al., 2019	SMD = 1.79, k = 22
		Pharmacy	1	2			Cui et al., 2023	SMD = 1.30, k = 22
		Health professional	1	1	2		Naing et al., 2023	SMD = 0.54, k = 44
		Medical microbiology		1			Li et al., 2023	SMD = 0.56, k = 13
		Radiology		1			Ge et al., 2020	SMD = 2.59, k = 19
		Anatomy	1				Kazeminia et al., 2022	n = 11
	Autonomous learning (6.67%, n = 3)	Chinese medicine		1			Xiao et al., 2020	MD = 6.61, k = 25
		Physical education			1		Tutal & Yazar, 2021	g = 1.47, k = 6
		Health professions	1				Li et al., 2020	n = 27
		Radiology		1			Ge et al., 2020	OR = 1.98, k = 5
		Chinese medicine		1			Xiao et al., 2022	RR = 1.52, k = 5
		Higher order thinking skills (6.67%, n = 3)	Nursing		2			Park & Suh, 2021
	Health professions		1				Li et al., 2020	n = 12
	Conceptual learning (4.44%, n = 2)		Health professions	1			Li et al., 2020	n = 17
	Affective	Satisfaction & enjoyment (17.78%, n = 8)	Radiology		1			Ge et al., 2020
Nursing			1	2			Xu et al., 2019	ES = 1.39, k = 6
Health professional			1	1			Naing et al., 2023	SMD = 0.48, k = 8
Radiology				1			Ge et al., 2020	OR = 1.70, k = 7
Positive thought & interest (6.67%, n = 3)		Anatomy	1				Kazeminia et al., 2022	n = 3
		Pharmacy	1				Aljaber et al., 2023	n = 26
		Anatomy	1				Kazeminia et al., 2022	n = 3
Self-confidence & efficacy (4.44%, n = 2)		Nursing		1			Xu et al., 2019	ES = 1.58, k = 15
		Pharmacy	1				Aljaber et al., 2023	n = 5
		Anatomy	1				Kazeminia et al., 2022	n = 3
Interpersonal	Interaction (6.67%, n = 3)	Pharmacy		1			Aljaber et al., 2023	n = 5
		Anatomy	1				Kazeminia et al., 2022	n = 3
	Peer learning & collaboration (4.44%, n = 2)	Nursing		1			Xu et al., 2019	ES = 1.41, k = 8
		Nursing		1			Xu et al., 2019	ES = 1.60, k = 10
	Engagement (4.44%, n = 2)	Health Professional	1				Li et al., 2020	n = 18
		Health professional	1				Li et al., 2020	n = 2
		Pharmacy	1				Aljaber et al., 2023	n = 5
Active in-class time & feedback (2.22%, n = 1)	Anatomy	1				Kazeminia et al., 2022	n = 5	
Total	100 %, n = 45		21	21	0	3		

Note: QSR = Qualitative systematic review; MA = Meta analysis; MM = Mixed-method review; CMA = Cross-disciplinary meta-analysis

In MAs, ES indicates effect size, SMD indicates standardized mean difference, MD indicates mean difference, OR indicates odds ratio, RR indicates risk ratio

## Affective Domain Outcomes in M&H Sciences

The integration of qualitative ( $n = 3$ ) and quantitative ( $n = 1$ ) evidence from 4 reviews underscores several key themes regarding the association between FL and interpersonal outcomes in M&H sciences, including anatomy, pharmacy, nursing, and health professions education (Li et al., 2020). Qualitative reviews in anatomy and pharmacy education suggest increased student interaction in FL environments compared to traditional settings, although this observation is based on a small number of studies (Aljaber et al., 2023; Kazeminia et al., 2022). A meta-analysis in nursing education (Xu et al., 2019) reports a significant positive effect of FL on communication and interaction opportunities ( $ES = 1.41$ ,  $k = 8$ ), though interpretive caution is warranted given the study characteristics.

FL demonstrates potential as a pedagogical approach for cultivating teamwork and communication skills, particularly in China, albeit with some caveats. Li et al.'s (2020) qualitative review found that most of Chinese undergraduate students in health professions programs reported enhanced communication and teamwork, likely due to increased group learning opportunities in flipped courses. Similarly, Xu et al.'s (2019) meta-analysis demonstrated a significant positive effect of FL on nursing students' collaborative spirit and teamwork ( $ES = 1.60$ ,  $k = 10$ ). However, limitations in research methodology and sample diversity warrant further investigation to solidify these findings.

Engagement also emerged as a crucial interpersonal outcome in our synthesis. Qualitative reviews in health professions and pharmacy education suggest a potential positive association between FL and increased student engagement. Li et al. (2020) found that 6 out of 56 studies in their review of Chinese undergraduate health professions programs reported improved student engagement with faculty and peers in FL settings. Similarly, Aljaber et al. (2023) noted that 5 out of 35 pharmacy education studies indicated increased student engagement resulting from greater interaction with instructors and peers in FL settings. Furthermore, Kazeminia et al. (2022) reported that 5 out of 26 anatomy education studies observed higher levels of student activity in FL groups. Despite a limited scope, the findings suggest FL's interactive and collaborative approach fosters a more engaging learning environment for health professions students.

## Learning Outcomes in SS Fields

### Cognitive Domain Outcomes in SS Fields

Reviews exploring FL practices across SS fields found positive impact of FL academic achievement. The extensive FL research in English language teaching (ELT) overwhelmingly supports its efficacy in enhancing language learning. Qualitative reviews ( $n = 3$ ) consistently highlight improvements in various language skills, including speaking, reading, writing, and grammar, among students exposed to FL approach compared to traditional teaching methods (Arslan, 2020; Linling & Abdullah, 2023; Turan & Akdag-Cimen, 2020). Several meta-analyses ( $n = 5$ ) provide further quantitative evidence, with effect sizes ranging from 0.58 to 1.30 demonstrating the significant positive impact of FL on language learning achievement (Aybirdi et al., 2023; Chen et al., 2023; Ni et al., 2023; Shahnama et al., 2021; Vitta & Al-Hoorie, 2023). For example, Chen et al. (2023), in their analysis of 26 studies, found a statistically significant overall effect ( $d_{\text{overall}} = 1.09$ ) favouring FL pedagogy, with a particular benefit observed for writing and speaking skills, although the number of studies focusing on each individual skill was limited. Similarly, Ni et al. (2023) reported a moderate effect ( $ES = 0.68$ ,  $k = 24$ ) on academic achievement in ELT. Aybirdi et al. (2023) further supported this with a statistically significant effect of FL on L2 learners' academic achievements ( $g = 1.30$ ,  $k = 40$ ). Vitta and Al-Hoorie (2023) revealed a moderate effect size ( $g = 0.58$ ,  $k = 56$ ) for L2 learning, with increased effectiveness for intermediate to advanced learners ( $g = 0.89$ ,  $k = 16$ ) and skill-based outcomes like writing ( $g = 1.50$ ,  $k = 13$ ).

studies). Shahnama et al. (2021) also found a large positive impact of FL on student achievement ( $d = 1.26$ ,  $k = 69$ ), with enhanced performance when pre-class materials were paired with additional activities. Moreover, cross-disciplinary meta-analyses ( $n = 2$ ) confirm the advantage of FL in language courses compared to other disciplines, with effect sizes indicating its substantial impact on language learning (Bredow et al., 2021,  $g = 0.57$ ,  $k = 35$ ; Tatal & Yazar, 2021,  $g = 0.95$ ,  $k = 63$ ).

The effectiveness of FL is not limited to ELT, with reviews demonstrating its potential effectiveness in broader educational contexts (Akçayır & Akçayır, 2018), educational sciences (Jang & Kim, 2020; Strelan et al., 2020a; Tatal & Yazar, 2021), business education (Senali et al., 2022), and introductory statistics (Farmus et al., 2020). The cross-disciplinary meta-analysis of Jang and Kim (2020) reported a moderate effect size ( $ES = 0.72$ ,  $k = 26$ ) for educational sciences, while Tatal and Yazar (2021), in their cross-disciplinary meta-analysis, found a large size ( $g = 0.84$ ,  $k = 18$ ) in the same field. Furthermore, the cross-disciplinary meta-analysis of Strelan et al. (2020a) indicated a moderate effect size ( $g = 0.75$ ,  $k = 7$ ) in favour of teaching field.

While limited in study scope, reviews on FL in business and statistics education suggest potential for fostering academic performance. For example, a qualitative review of business and entrepreneurship studies by Senali et al. (2022) found that 11 out of 30 studies reported higher academic performance with FL compared to traditional approaches. Similarly, a meta-analysis of introductory statistics courses by Farmus et al. (2020) revealed a statistically significant improvement in final performance outcomes for students in flipped courses, with an average difference of 6.9% and an effect size of  $g = 0.43$  based on 11 studies. The moderating effect of weekly quizzes suggests implementation factors influence outcomes.

In addition to well-documented academic achievement regarding FL implementation in SS reviews, our research also reveals its significant influence on other cognitive domains, particularly within ELT ( $n = 3$ ) and broader educational contexts ( $n = 1$ ). Another cognitive outcome emerged from qualitative reviews ( $n = 4$ ) is autonomous learning FL empowers students to take ownership of their learning process. In ELT, Arslan (2020) identified 18 studies highlighting the support FL provides for self-paced language acquisition, while Linling and Abdullah (2023) found evidence in 9 studies that FL enhances learner autonomy during the COVID-19 pandemic. Similarly, Turan and Akdag-Cimen (2020) reported in 5 studies that FL enables language learners to progress at their own pace. This emphasis on learner autonomy is also evident in broader educational research, with Akçayır and Akçayır (2018) identifying 6 studies that showcase the role of FL in fostering student autonomy across diverse disciplines.

Beyond autonomous learning, FL practices appear to contribute to deeper conceptual understanding and the development of higher order thinking skills. Akçayır and Akçayır (2018) identified 7 studies demonstrating the positive impact of FL on conceptual understanding, allowing students to grasp complex concepts more effectively. Furthermore, their analysis revealed 6 studies that highlighted the improvement of critical thinking skills through FL approach, equipping students with essential analytical and problem-solving capabilities.

The current evidence supporting cognitive domain outcomes, particularly including conceptual understanding and higher order thinking skills, while suggestive of positive outcomes, suffers from limitations that hinder definitive conclusions. A heavy reliance on single review for reporting outcomes, such as improved conceptual understanding and critical thinking, raises concerns about the consistency and generalizability of these findings. Additionally, even within larger reviews, the number of studies reporting specific outcomes is often low, limiting the ability to draw strong conclusions about the impact of FL on particular aspects of domainial outcomes. Furthermore, the predominance of qualitative evidence, valuable for understanding student and teacher experiences, may introduce biases and challenges in generalizing findings to broader populations and diverse educational contexts.

**Table 3.** Disciplinary and learning outcomes in the field of SS

Domain	Theme	Discipline	QSR	MA	MM	CMA	Sample review	Effect size	
Cognitive	Academic achievement (35.54%, n = 16)	EFL & ESL	3	5		2	Ni et al., 2023	ES = 0.68, k = 24	
		General education	1			3	Jang & Kim, 2020	ES = 0.80, k = 26	
		Business	1				Senali et al., 2022	n = 11	
		Introductory statistics		1			Farmus et al., 2020	g = 0.43, k = 10	
	Autonomous learning (8.89%, n = 4)	EFL/ESL	3				Linling & Abdullah, 2023	n = 9	
		General education	1				Akcayır & Akcayır, 2018	n = 6	
	Conceptual learning (2.22%, n = 1)	General education	1				Akcayır & Akcayır, 2018	n = 7	
	Higher order thinking skills (2.22%, n = 1)	General education	1				Akcayır & Akcayır, 2018	n = 6	
	Affective	Satisfaction & enjoyment (6.67%, n = 3)	EFL/ESL	1				Arslan, 2020	n = 7
			General education	1				Akcayır & Akcayır, 2018	n = 13
Business			1				Senali et al., 2022	n = 8	
Positive thought & interest (6.67%, n = 3)		EFL/ESL	1				Arslan, 2020	n = 23	
		Business	1				Senali et al., 2022	n = 7	
		General education	1				Akcayır & Akcayır, 2018	n = 10	
Self-confidence & efficacy (6.67%, n = 3)		EFL/ESL	2				Linling & Abdullah, 2023	n = 6	
		General education	1				Akcayır & Akcayır, 2018	n = 5	
Learner motivation (6.67%, n = 3)		EFL/ESL	2				Turan & Akdag-Cimen, 2020	n = 5	
		General education	1				Akcayır & Akcayır, 2018	n = 7	
Interpersonal	Engagement (8.89%, n = 4)	EFL/ESL	3				Arslan, 2020	n = 19	
		General education	1				Akcayır & Akcayır, 2018	n = 10	
	Interaction (6.67%, n = 3)	EFL	2				Linling & Abdullah, 2023	n = 8	
		General education	1				Akcayır & Akcayır, 2018	n = 7	
	Active in-class time & feedback (6.67%, n = 3)	EFL/ESL	1				Arslan, 2020	n = 17	
		Business	1				Senali et al., 2022	n = 6	
		General education	1				Akcayır & Akcayır, 2018	n = 4	
	Peer learning & collaboration (2.22%, n = 1)	EFL/ESL	1				Arslan, 2020	n = 14	
	Total	100 %, n = 45		34	6	0	5		

Note: QSR = Qualitative systematic review; MA = Meta analysis; MM = Mixed-method review; CMA = Cross-disciplinary meta-analysis

### Affective Domain Outcomes in SS Fields

FL use in SS fields has the potential for improving affective domain outcomes, as it revealed through a synthesis of several qualitative reviews (n = 5). An emerged theme in the reviews is the enhancement of student satisfaction and enjoyment (n = 3). For example, 7 studies out of 78 studies reported increased

satisfaction with flipped language courses (Arslan, 2020). Similarly, in broader educational fields, 13 articles out of 71 indicated that FL led to heightened learning satisfaction (Akçayır & Akçayır, 2018). This extends to the business field, where 8 studies out of 30 studies highlighted FL as a source of enjoyable learning experiences for students (Senali et al., 2022).

Another key finding is the cultivation of positive thoughts and heightened interest among students engaged in FL environments. In the field of ELT, 23 studies reported students' positive perceptions about FL practices (Arslan, 2020). In the same vein, 13 studies in broader educational disciplines revealed positive student feedback regarding FL experiences (Akçayır & Akçayır, 2018). Furthermore, 7 studies out of 30 studies in a qualitative review in business education demonstrated that FL effectively enhanced student interest in the subject matter (Senali et al., 2022).

Further, a few qualitative reviews (n = 3) highlight that FL is also conducive to fostering self-confidence and self-efficacy among learners. For example, in ELT, 10 studies out of 78 in a review noted that FL contributed to building self-confidence in language learners (Arslan, 2020), while 6 studies out of 22 in EFL contexts highlighted the approach's efficacy in enhancing self-confidence during the COVID-19 pandemic (Linling & Abdullah, 2023). Additionally, 5 articles out of 71 in broader educational fields reported increased student confidence through FL practices (Akçayır & Akçayır, 2018).

Several qualitative reviews (n = 3) suggest a positive relationship between FL and student motivation in ELT and general educational disciplines. Arslan (2020) found out in 35 studies out of 78 that FL pedagogy increases language learners' motivation. Similarly, Turan and Akdag-Cimen (2020) noted that 5 studies out of 43 on FL in ELT contexts indicated an increase in student motivation. Additionally, Akçayır and Akçayır (2018) reported that 7 articles out of 71 in general education disciplines observed increased student motivation with FL pedagogy. Despite the promising outcomes reported in these studies, the reliance on qualitative data and the scarcity of reviews directly examining motivation as a primary outcome necessitate further exploration to confirm the connection between FL and affective domain outcomes.

### Interpersonal Domain Outcomes in SS Fields

Reviewing several qualitative studies (n = 5) revealed that incorporating FL practices in SS fields can potentially improve interpersonal outcomes. Engagement emerges as a key outcome, highlighted in a few qualitative reviews (n = 3) on ELT (Arslan, 2020; Turan & Akdag-Cimen, 2020; Linling & Abdullah, 2023) and in a qualitative review in broader educational contexts (Akçayır & Akçayır, 2018). Studies from various locations around the world in these reviews showcase the effectiveness of FL practices in boosting student engagement, as evidenced by 19 studies out of 78 studies (Arslan, 2020), 11 out of 43 (Turan & Akdag-Cimen, 2020), 11 out of 22 studies (Linling & Abdullah, 2023), in ELT contexts and 10 out of 71 studies (Akçayır & Akçayır, 2018) in educational fields.

Beyond engagement benefit of FL, it was also reported to foster peer interaction and student-teacher interaction. A few qualitative reviews in ELT field (n = 2) report that 8 out of 22 studies found FL boosted peer interactions (Linling & Abdullah, 2023), while 6 out of 43 studies showed enhanced peer interactions in the same field (Turan & Akdag-Cimen, 2020). Additionally, 7 out of 71 studies in various educational disciplines indicated that FL boosted student-teacher interaction levels (Akçayır & Akçayır, 2018). Furthermore, FL was also shown to facilitate active learning, with 28 out of 78 studies reporting increased student activity during language learning (Arslan, 2020) and 7 out of 30 studies in business education showing support for active learning process (Senali et al., 2022). This active learning environment is facilitated by more efficient use of class time, as reported in 9 out of 71 studies across educational disciplines (Akçayır & Akçayır, 2018). FL approach also fosters collaborative learning, as evidenced by Arslan's (2020) analysis of 78 ELT studies, where 14 of them highlighted increased opportunities for collaboration in flipped language courses.

Qualitative evidence on FL's interpersonal impact is complex and varied due to limitations. Reliance on single-study observations with varying positive outcome reports hinders a clear understanding of domain outcomes. The lack of reviews specifically addressing interpersonal outcomes like interaction hinders conclusive assessments of FL's effectiveness in these areas. Additionally, the predominance of qualitative evidence, while insightful, raises concerns about generalizability due to potential biases and limitations inherent in qualitative methods.

## DISCUSSIONS AND SUGGESTIONS

In our research, 41 reviews retrieved from Web of Science, EBSCO Academic Search Complete, and ERIC, were synthesized to reveal the effectiveness of FL on students' cognitive, affective, and interpersonal development across STEM, M&H, and SS fields. Although the current research landscape—shaped by traditional education—often prioritizes cognitive skills and knowledge acquisition (see Fig. 4), the 21st century demands a broader skill set that also includes social-emotional competencies and interpersonal abilities. Less extensively studied areas as affective and interpersonal domains mirror traditional mindset of teachers and researchers that goes back to traditional educational philosophies placing a heavy emphasis on knowledge acquisition and measurable cognitive skills (Bloom et al., 1956; Skinner, 1953). This focus on cognitive outcomes has had a lasting impact on educational practices and assessment methods, often neglecting affective and interpersonal dimensions of learning. The underemphasis on affective and interpersonal learning in FL research creates a critical skills gap, leaving students ill-equipped for the demands of the 21st-century workforce, which increasingly seeks individuals with strong communication, collaboration, and problem-solving skills (Casner-Lotto & Barrington, 2006; Trilling & Fadel, 2009), while also neglecting the significant contributions of social-emotional competencies (SEL) and positive relationships to students' overall well-being and mental health (Durlak et al., 2011).

Specifically, considering cognitive domain outcomes, our synthesis of qualitative and quantitative evidence on FL practices in STEM, M&H sciences, and SS fields reveals that FL is effective for promoting academic achievement in mathematics, science and English language teaching, whereas evidence regarding academic achievement remains tentative and inconclusive in other STEM-related disciplines such as ICT and engineering, in M & H sciences, and SS disciplines like educational sciences, business, and statistics education, excluding ELT field. In terms of academic achievement in ICT and engineering, several limitations necessitate cautious interpretation and highlight avenues for further research. A primary concern lies in the broad categorization of STEM disciplines, obscuring the unique pedagogical needs and potential variations in effectiveness across specific domains. For instance, grouping diverse fields such as mechanical engineering and biochemistry under “science and engineering” limits the applicability of findings for educators in particular areas. Similarly, the broad label of “engineering” masks potential differences in FL efficacy across sub-disciplines like chemical or civil engineering. Additionally, concerns regarding generalizability and robustness arise due to the limited number of studies reporting effect sizes in several meta-analyses conducted in ICT and engineering fields. For example, the seemingly large effect size for computer science is based on a mere seven studies, emphasizing the need for a larger pool of research to provide more reliable estimations.

Similarly, while FL demonstrates the potential for improving academic achievement in nursing, pharmacy education, and health professions education in China and internationally, interpreting these results requires caution. This caution stems from the concentrated evidence base in certain geographic regions (e.g., Ge et al., 2020; Xiao et al., 2020). With a smaller pool of studies and smaller sample sizes within those studies, it becomes challenging to establish the true effectiveness of FL pedagogy in M&H fields due to limited generalizability and decreased statistical power (Cohen, 1988). In the same vein, the presence of publication bias makes it challenging to accurately identify and compare the effectiveness of FL interventions. While cultural, educational, and technological contexts significantly influence the effectiveness of FL (Berry et al., 2011), overreliance on region-specific research can lead to inaccurate assumptions; therefore, conducting cross-cultural and comparative research is crucial for understanding contextual influences and ensuring the generalizability of findings (Berry et al., 2011; Creswell & Poth, 2018).

Despite encouraging academic achievement indicator of FL for educational sciences, business, and statistics education, the generalizability of evidence in these fields is tempered by the limited breadth of primary studies and reviews reporting this outcome. Thus, including multiple primary studies in reviews increases the generalizability of findings across diverse educational settings and reduces the risk of drawing conclusions based on isolated results (Cooper et al., 2019). Another concern is about the variety of educational contexts encompassed and a lack of discipline-specific focus in educational sciences. This raises questions about the applicability of these results in specific educational fields. To gain a more nuanced understanding of FL effectiveness in SS fields, research must move beyond broad generalizations and delve into the unique characteristics of specific contexts (Anderson-Cook, 2005), since understanding the unique characteristics of

each academic discipline, including their specific epistemologies, methods, and ways of thinking, is crucial for effective teaching and learning (National Research Council, 2005). This understanding is particularly relevant for adapting FL approach to align with the diverse learning goals and pedagogical approaches within various SS disciplines.

Our synthesis across STEM, M&H sciences, and SS fields suggests potential benefits of FL pedagogy in fostering autonomy, conceptual understanding, and higher order thinking skills (HOTS), the evidence remains limited and inconclusive. Several factors within the reviewed studies may contribute to this uncertainty. Firstly, the limited number of primary studies reporting and calculating effect sizes for these cognitive outcomes across fields hinders drawing reliable and generalizable conclusions (Higgins et al., 2019). Secondly, methodological variations across studies may influence the observed effect sizes for these outcomes. Thirdly, there is a widespread lack of reporting on the specific mechanisms through which FL pedagogy cultivates autonomy, conceptual understanding, and HOTS across fields. This is crucial because there is no “one size fits all” strategy for fostering these cognitive outcomes. A strategy effective for one student might not be as effective for another (Kupers et al., 2013; Reeve, 2009; Stroet et al., 2013). Additionally, higher order thinking skills (HOTS) do not develop automatically (Williams, 2015) and require active engagement in educational activities that strengthen them. Although classroom discussions, practice assignments, and targeted remedial help can promote HOTS, there is a need to explore and identify teaching approaches that specifically foster their development (Huang et al., 2022). Integrating proper learning strategies into FL is also crucial for enhancing students’ HOTS (Chang & Hwang, 2018; Ye et al., 2019). Considering these multifaceted constructs within FL practices across disciplines, it is essential to establish a set of design principles and strategies that cultivate self-directed learning, conceptual understanding, and HOTS in students.

Our synthesis also provides empirical validation for the theoretical framework underpinning FL pedagogy. The robust evidence for academic achievement outcomes across fields (with effect sizes ranging from 0.40 to 1.30) directly substantiates social constructivist principles, as these gains emerge from the collaborative knowledge construction that occurs when students engage in peer instruction, group problem-solving, and discussion-based activities during transformed in-class sessions. Self-determination theory’s explanatory power is particularly evident in the affective domain findings, where reviews consistently report enhanced satisfaction, motivation, and self-efficacy—outcomes that align precisely with SDT’s postulation that FL satisfies students’ basic psychological needs. Specifically, the flexibility of pre-class learning supports autonomy (as evidenced by improved autonomous learning outcomes reported across all three fields), collaborative in-class activities foster relatedness (reflected in enhanced peer learning and interaction outcomes), and scaffolded problem-solving builds competence (manifested in improved self-confidence and efficacy). The interpersonal outcomes identified across our synthesis—including enhanced collaboration, peer learning, and engagement—validate both the social dimension of constructivism and Kolb’s experiential learning cycle, as these benefits emerge when students move from abstract conceptualization during pre-class preparation to concrete experience and active experimentation during collaborative in-class activities. However, our findings also reveal a critical theory-practice gap: while these theories predict strong outcomes across all domains, the evidence remains predominantly focused on cognitive achievements, with affective and interpersonal outcomes being underexplored despite their theoretical importance. This disparity suggests that current FL implementations may not fully operationalize the theoretical mechanisms, particularly those supporting affective and interpersonal development, indicating the need for more intentional design of FL environments that explicitly leverage SDT’s motivational principles and Kolb’s complete experiential cycle to realize the full potential of this pedagogical approach.

Regarding affective domain outcomes, our synthesis primarily involves qualitative reviews across STEM, M&H sciences, and SS fields. It reveals inconclusive evidence for enhancing these outcomes. This is due to the low number of primary studies and the scarcity of meta-analytic reviews that report on affective outcomes. Qualitative reviews often rely on narrative descriptions and thematic analysis, which can be subjective and prone to interpretation bias (Cooper et al., 2019). Given that FL implementations vary widely across disciplines and educational levels, qualitative reviews may struggle to synthesize this heterogeneity and identify generalizable patterns. In this respect, quantitative and mixed-method research could offer a more comprehensive understanding of research topics. Meta-analyses can statistically account for variations in study characteristics and explore potential moderators of FL’s effects on affective domain (Borenstein et al., 2009).

By analysing quantitative data, researchers can identify specific FL strategies that are most effective for different student populations, subject areas, and learning outcomes. Merging quantitative and qualitative strands in a mixed-method research design may provide a robust and holistic understanding of the effectiveness of FL across various fields (Johnson & Onwuegbuzie, 2004). Thus, the scarcity of meta-analytic reviews and mixed-method reviews addressing affective aspects of FL across fields hinders a comprehensive understanding of FL's impact on affective domain outcomes. Another consideration for M&H sciences is the geographic concentration of research and potential underrepresentation of null findings (Franco et al., 2014). The diversity in FL implementations—spanning instructional designs, student populations, and outcome measures—reflects the complexity of synthesizing evidence across varied educational contexts (Borenstein et al., 2009). Thus, these issues can compromise the reliability and generalizability of findings, limiting the understanding of FL's effectiveness on affective domain outcomes across diverse cultural and educational contexts in this field.

Our analysis of the current reviews respecting what supports affective domain outcomes like satisfaction, motivation or self-confidence/efficacy in FL environments across fields reveals the lack of sufficient reporting and examination on the mechanisms of FL's impact on the affective domain outcomes in the extant reviews. The absence of detailed reporting and examination into how FL enhances affective outcomes hinders our understanding of both the effectiveness of FL affective outcomes and underlying pedagogical mechanisms involved. Investigating the specific components of FL that contribute to changes in the affective domain, such as increased motivation or improved self-confidence/efficacy, could provide valuable insights for designing and implementing effective FL strategies that foster positive affective outcomes and enhance student learning experiences. A final point of concern is the lack of detailed account for the metrics reported by the extant reviews across fields that were employed to assess affective outcomes in the included studies in these reviews. This might stem from two potential sources. First, the variety of measurement tools and methods used across primary studies introduces inconsistency and heterogeneity in primary studies examining the affective domain outcomes and makes it difficult for reviewer authors to synthesize findings and report them. Second, the field may lack standardized and validated instruments specifically designed to assess affective domain constructs, further contributing to the observed insufficient reporting in the extant reviews. The use of diverse and often unclear measurement approaches in affective domain research presents a dual challenge. It hinders the verification and validation of findings across studies, impeding the development of a robust evidence base. Additionally, it obscures the specific impacts of FL interventions on learners' affective experiences, limiting our understanding of how FL promotes various aspects of the affective domain across fields.

As for interpersonal domain outcomes, evidence in the extant reviews across STEM, M&H sciences, and SS fields suggests potential benefits of FL for promoting students' interpersonal outcomes, the predominantly qualitative evidence base, with its inherent subjectivity and susceptibility to interpretation bias (Cooper et al., 2019), limits our ability to draw definitive conclusions. The scarcity of quantitative and mixed method research further restricts our understanding of FL's impact on students' interpersonal development (Borenstein et al., 2009; Johnson & Onwuegbuzie, 2004). Similar methodological considerations apply to interpersonal outcomes in M&H fields. Another point of concern, like the affective domain, is that most reviews provide limited exploration of the mechanisms or pedagogical strategies through which FL fosters interpersonal skills. For instance, while qualitative reviews in STEM (e.g., Cevikbas & Kaiser, 2023; Lo & Hew, 2019) and M&H sciences (e.g., Kazeminia et al., 2022; Li et al., 2020) report positive interpersonal outcomes such as enhanced collaboration and peer learning, they typically describe these as observed outcomes without detailing the specific FL design elements or pedagogical processes that produced them. Similarly, reviews in social sciences (e.g., Arslan, 2020; Akcayir & Akcayir, 2018) document improvements in engagement and interaction but offer minimal analysis of which particular FL strategies—such as structured group roles, peer assessment protocols, or specific collaborative task designs—contribute to these interpersonal gains.

Across the included reviews, a specific evidence gap concerns interpersonal outcomes (e.g., collaboration quality, communication, peer support): these are inconsistently defined, infrequently treated as primary endpoints, and commonly operationalized with ad-hoc or single-item indicators lacking reported reliability/validity, which limits comparability and meta-synthesis. Moreover, while FL commonly employs collaborative in-class practices (team-based problem solving, jigsaw, peer instruction, peer review, simulations), reviews seldom report these practices with sufficient granularity (e.g., group composition, task interdependence, role allocation, assessment criteria) or pair them with standardized, multi-item measures, making it difficult to attribute changes in interpersonal skills to specific FL designs. We therefore interpret interpersonal findings cautiously and recommend future studies use validated interpersonal metrics and detailed reporting of collaborative task design to enable cross-study comparison

Our study involving the synthesis of 41 reviews regarding FL practices across fields reveals several areas requiring further investigation to solidify our understanding of its effectiveness and optimize its implementation across fields. At review level, it is recommended that STEM fields prioritize meta-analytic reviews, while M&H sciences and SS fields focus on conducting more mixed-method reviews to advance research synthesis. More reviews encompassing research from diverse geographical contexts are needed to ensure the global applicability and cultural sensitivity of findings. Reviews focusing on specific domains, such as motivation within M&H sciences, can provide more nuanced insights into FL's impact in the fields.

At domain level, further reviews and research on cognitive skills such as conceptual understanding and higher order thinking skills (HOTS) are required to provide a more holistic picture of FL's impact on students' cognitive skills. Deeper exploration of affective and interpersonal outcomes across disciplines is also warranted. At disciplinary level, more discipline-specific reviews and research are needed to understand the effectiveness of FL within their unique contexts. Moreover, given the potential of FL for teacher education and professional development, more reviews and research specific to educational sciences in SS domain is needed.

At outcome level, reviews and research are needed to reveal how FL impacts motivation within M&H sciences. To optimize the understanding of FL's effectiveness, further reviews and research across fields should prioritize the exploration of outcome-related mechanisms, specifically investigating the processes through which FL practices influence various learning outcomes, such as academic achievement, student engagement, and motivation. Respecting development of common metrics, establishing standardized and validated instruments for measuring various cognitive, affective, and interpersonal outcomes is crucial for ensuring consistency and comparability across research studies and contexts. At conceptual framework level, although FL pedagogy seems to align with social constructivism, self-determination, and experiential learning theories in principle, there is a gap between theory and practice. Current research suggests that FL pedagogy may not fully leverage the potential of these theories to promote cognitive, affective, and interpersonal learning outcomes. Further research is needed to explore, implement theory-based FL practices, and establish a set of design principles and strategies that are explicitly grounded in these theories. This will help bridge the gap between theory and practice, creating engaging and effective FL environments that foster holistic student development.

Several limitations of this study should be mentioned. Firstly, our umbrella review is limited to the reports of what researchers investigated, published and systematically reviewed or meta-analysed. Secondly, our umbrella review is limited to certain review types including qualitative systematic reviews, meta-analytic reviews, and mixed-method reviews. Inclusion of other review such as scoping reviews, integrative reviews, literature reviews would yield different results in terms of FL practices. Thirdly, this study is limited to reviews published in English language, coupled with specific inclusion criteria for reviews such as the Joanna Briggs Institute's appraisal checklist and specific researcher-decision based criteria for the inclusion of meta-analytic reviews. Finally, the review is limited to narrative synthesis of included reviews. Though not practically possible to quantify effect sizes reported in meta-analytic reviews by conducting moderator analysis, a meta-meta-analysis would yield different results as to FL practices across fields.

## BIODATA AND CONTACT ADDRESSES OF AUTHORS



**Dr. Abdullah ARSLAN** is an Associate Professor of English Language Teaching (ELT) at the School of Foreign Languages, Akdeniz University, Antalya, Turkiye. He earned his Ph.D. in Curriculum and Instruction from Abant Izzet Baysal University in 2011. Prior to his current position, he served as a faculty member at Shanghai International Studies University, China, for five years, where he delivered both undergraduate and graduate courses. His primary academic interests include EFL teaching, flipped learning, CEFR-based curriculum design, technology-integrated language instruction, systematic review methodologies, and professional development in ELT. His current research primarily focuses on Artificial Intelligence (AI) in ELT. He has published over 15 peer-reviewed journal articles covering a wide range of topics in ELT, including CEFR, technology-integrated language teaching, curriculum design, review methodology, and professional development. He is also the author of a book on English pronunciation and has contributed several chapters to edited academic volumes. In addition, he has presented papers at national and international conferences.

Abdullah ARSLAN  
School of Foreign Languages  
Address: Akdeniz University, 07200, Antalya, Turkiye  
Phone: +90 242 227 44 00/4727  
E-mail: [abdullaharslan@akdeniz.edu.tr](mailto:abdullaharslan@akdeniz.edu.tr)



**Dr. Alaettin ISERI** an Assistant Professor in the Department of Educational Sciences at the Faculty of Arts and Sciences, Kırklareli University, in Kırklareli, Turkiye. He earned his Ph.D. in Curriculum and Instruction from Abant Izzet Baysal University in 2011. He has been serving as a faculty member at Kırklareli University since 2015. His primary academic interests include educational and curriculum theory, education ideology and epistemology, intercultural education, learning and learning outcomes, curriculum evaluation, improving education at the higher education level, and artificial intelligence (AI). He is also currently actively involved in various European Union projects in the field of adult education. His publications include national and international articles, book chapters, and conference papers on curriculum development theory, ideology, epistemology, learning outcomes, the relationship between PISA competencies and Turkish curriculum learning outcomes, the epistemological nature of Turkish curriculum learning outcomes, 21st-century skills and textbook content, and internationalization in higher education.

Alaettin ISERI  
Department of Educational Sciences  
Address: Kırklareli University, Faculty of Arts and Sciences,  
Kayali Campus, 39000, Kırklareli, Turkiye.  
Phone: +90 288 246 17 34/3523  
mail: [alaettiniseri@klu.edu.tr](mailto:alaettiniseri@klu.edu.tr)

## REFERENCES

- Abeyssekera, L., & Dawson, P. (2014). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research & Development*, 34(1), 1–14. <https://doi.org/10.1080/07294360.2014.934336>
- Afonso, A., Morgado, L., Noguera, I., Sepúlveda-Parrini, P., Hernandez-Leo, D., Alkhasawneh, S. N., Spilker, M. J., & Carvalho, I. C. (2025). Flexible learning by design: Enhancing faculty digital competence and engagement through the FLeD project. *Education Sciences*, 15(7), 934. <https://doi.org/10.3390/educsci15070934>

- Al-Hammouri, M. M., Rababah, J. A., Rowland, M. L., Tetreault, A. S., & Aldalaykeh, M. (2020). Does novel teaching approach work? A students' perspective. *Nurse Education Today*, 85, 104–229. <https://doi.org/10.1016/j.nedt.2019.104229>
- Alonso, R. K., Vélez, A., Martínez-Monteaudo, M. C., & Rico-González, M. (2023). Flipped learning in higher education for the development of intrinsic motivation: A systematic review. *Education Sciences*, 13(12), 1226. <https://doi.org/10.3390/educsci13121226>
- Altemueller, L., & Lindquist, C. (2017). Flipped classroom instruction for inclusive learning. *British Journal of Special Education*, 44(3), 341–358. <https://doi.org/10.1111/1467-8578.12177>
- Anderson-Cook, C. M. (2005). Experimental and quasi-experimental designs for generalized causal inference. *Journal of the American Statistical Association*, 100(470), 708–708. <https://doi.org/10.1198/jasa.2005.s22>
- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-based Healthcare*, 13(3), 132–140. <https://doi.org/10.1097/xeb.0000000000000055>
- Aromataris, E., Godfrey, C., Holly, C., Khalil, H. & Tungpunkom, P. (Eds.). (2020). Chapter 10: Umbrella reviews. In *JBIM manual for evidence synthesis*. Adelaide: Joanna Briggs Institute (JBI). <https://doi.org/10.46658/jbimes-20-11>
- Ayres, L. (2008). Thematic coding and analysis. In: L.M. Given (Ed.), *The Sage encyclopaedia of qualitative research methods* (867–868). Sage Publications. <http://dx.doi.org/10.4135/9781412963909.n451>
- Becker, L. A., & Oxman, A. D. (2008). Overviews of reviews. *Cochrane handbook for systematic reviews of interventions* (607–631). <https://doi.org/10.1002/9780470712184.ch22>
- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. ASCD.
- Berry, J. W., Poortinga, Y. H., Breugelmans, S. M., Chasiotis, A., & Sam, D. L. (2011). *Cross-cultural psychology: Research and applications* (3rd ed.). Cambridge University Press.
- Betihavas, V., Bridgman, H., Kornhaber, R., & Cross, M. (2016). The evidence for 'flipping out': A systematic review of the flipped classroom in nursing education. *Nurse Education Today*, 38, 15–21. <https://doi.org/10.1016/j.nedt.2015.12.010>
- Bialik, M., & Fadel, C. (2015). Skills for the 21st century: What should students learn?. Retrieved July 10, 2023, from [https://curriculumredesign.org/wp-content/uploads/CCR-Skills\\_FINAL\\_June2015.pdf](https://curriculumredesign.org/wp-content/uploads/CCR-Skills_FINAL_June2015.pdf)
- Birgili, B., Seggie, F. N., & Oguz, E. (2021). The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis. *Journal of Computers in Education*, 8(3), 365–394. <https://doi.org/10.1007/s40692-021-00183-y>
- Bishop, J., & Verleger, M. (2013). Testing the flipped classroom with model-eliciting activities and video lectures in a mid-level undergraduate engineering course. In 2013 IEEE Frontiers in Education Conference (FIE) (pp. 161–163). <https://doi.org/10.1109/fie.2013.6684807>
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. David McKay Publisher.
- Booth A., Noyes J., Flemming K., Gerhardus A., Wahlster P., van der Wilt G. J., Mozygemba K., Refolo P., Sacchini D., Tummers M., Rehfuss E. (2016). Guidance on choosing qualitative evidence synthesis methods for use in health technology assessments of complex interventions. *Integrate-HTA*. <http://www.integrate-hta.eu/downloads/>
- Bonwell, C., & Eison, J. (1991). *Active learning: Creating excitement in the classroom*. Washington, DC: ASHE-ERIC higher education reports. Retrieved June 25, 2022, from <https://files.eric.ed.gov/fulltext/ED336049.pdf>

- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons. <https://doi.org/10.1002/9780470743386>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Care, E. (2017). 21st century skills: From theory to action. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and teaching of 21st century skills: Research and applications* (pp. 3–17). Springer.
- Casner-Lotto, J., & Barrington, L. (2006). Are they really ready to work?: Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce. *Partnership for 21st Century Skills*.
- Castedo, R., López, L. M., Chiquito, M., Navarro, J., Cabrera, J. D., & Ortega, M. F. (2018). Flipped classroom—comparative case study in engineering higher education. *Computer Applications in Engineering Education*, 27(1), 206–216. <https://doi.org/10.1002/cae.22069>
- Castedo, R., López, L. M., Chiquito, M., & Martín, J. D. C. (2019). To flip or not to flip? A case study on university engineering students. In *Innovative Trends in Flipped Teaching and Adaptive Learning* (pp. 17-37). IGI Global. <https://doi.org/10.4018/978-1-5225-8142-0.ch002>
- Cevikbas, M., & Kaiser, G. (2020). Flipped classroom as a reform-oriented approach to teaching mathematics. *ZDM Mathematics Education*, 52, 1291–1305. <https://doi.org/10.1007/s11858-020-01191-5>
- Chang, S.-C., & Hwang, G.-J. (2018). Impacts of an augmented reality-based flipped learning guiding approach on students' scientific project performance and perceptions. *Computers & Education*, 125, 226–239. <https://doi.org/10.1016/j.compedu.2018.06.007>
- Chapman, E. (2003). Alternative approaches to assessing student engagement rates. *Practical Assessment, Research and Evaluation*, 8(13), 1–10. <https://doi.org/10.7275/3e6e-8353>
- Chen, K.-S., Monrouxe, L., Lu, Y.-H., Jenq, C.-C., Chang, Y.-J., Chang, Y.-C., & Chai, P. Y.-C. (2018). Academic outcomes of flipped classroom learning: A meta-analysis. *Medical Education*, 52(9), 910–924. <https://doi.org/10.1111/medu.13616>
- Chen, C.-K., Huang, N.-T. N., & Hwang, G.-J. (2019). Findings and implications of flipped science learning research: A review of journal publications. *Interactive Learning Environments*, 30(5), 949–966. <https://doi.org/10.1080/10494820.2019.1690528>
- Cheng, S.-C., Hwang, G.-J., & Lai, C.-L. (2020). Critical research advancements of flipped learning: a review of the top 100 highly cited papers. *Interactive Learning Environments*, 30(9), 1751–1767. <https://doi.org/10.1080/10494820.2020.1765395>
- Chen, C.-C. (2021). Effects of flipped classroom on learning outcomes and satisfaction: An experiential learning perspective. *Sustainability*, 13(16), 9298. <https://doi.org/10.3390/su13169298>
- Chiu, T. K. F. (2021). Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic. *Journal of Research on Technology in Education*, 54(sup1), S14–S30. <https://doi.org/10.1080/15391523.2021.1891998>
- Cohen, J., (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2019). *The handbook of research synthesis and meta-analysis* (3rd ed.). Russell Sage Foundation. <https://doi.org/10.7758/9781610448864>
- Coyle, V. C., Newman, D. L., & Connor, K. A. (2017). Innovative instruction in STEM education: The role of student feedback in the development of a flipped classroom. In *Blended learning: Concepts, methodologies, tools, and applications* (pp. 155-180). IGI Global. <https://doi.org/10.4018/978-1-5225-0783-3.ch008>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Sage Publications.

- Deci, E. L., Ryan, R. M., & Guay, F. (2013). Self-determination theory and actualization of human potentials. In D. M. McInerney, H. W. Marsh, R. G. Craven, F. Guay (Eds.), *Theory driving research: New wave perspectives on self-processes and human development* (pp. 109–133). Information Age Publishing.
- Dewey, J. (1938). *Experience and education*. Macmillan.
- Drew, V., & Mackie, L. (2011). Extending the constructs of active learning: Implications for teachers' pedagogy and practice. *The Curriculum Journal*, 22(4), 451–467. <https://doi.org/10.1080/09585176.2011.627204>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- El-Banna, M. M., Whitlow, M., & McNelis, A. M. (2017). Flipping around the classroom: Accelerated bachelor of science in nursing students' satisfaction and achievement. *Nurse Education Today*, 56, 41–46. <https://doi.org/10.1016/j.nedt.2017.06.003>
- El-Bassiouny, N., & El-Naggar, D. H. (2023). Flipped classroom application in an advanced research methods course: Insights from social cognitive theory and self-determination theory. *The International Journal of Management Education*, 21(3), 100881. <https://doi.org/10.1016/j.ijme.2023.100881>
- Evans, L., Vanden Bosch, M. L., Harrington, S., Schoofs, N., & Coviak, C. (2019). Flipping the classroom in health care higher education: A systematic review. *Nurse Educator*, 44(2), 74–78. <https://doi.org/10.1097/NNE.0000000000000554>
- Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses*. John Wiley & Sons.
- Fox, W. H., & Docherty, P. D. (2019). Student perspectives of independent and collaborative learning in a flipped foundational engineering course. *Australasian Journal of Educational Technology*, 35(5), 79–94. <https://doi.org/10.14742/ajet.3804>
- Francis, N., Morgan, A., Holm, S., Davey, R., Bodger, O., & Dudley, E. (2020). Adopting a flipped classroom approach for teaching molar calculations to biochemistry and genetics students. *Biochemistry and Molecular Biology Education*, 48(3), 220–226. <https://doi.org/10.1002/bmb.21328>
- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203), 1502–1505. <https://doi.org/10.1126/science.1255484>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <https://doi.org/10.3102/00346543074001059>
- Fry, H., Ketteridge, S., & Marshall, S. (2009). *A handbook for teaching and learning in higher education: Enhancing academic practice*. Routledge. <https://doi.org/10.4324/9781315763088>
- Fung, C., Besser, M., & Poon, K. (2021). Systematic literature review of flipped classroom in mathematics. *Eurasia Journal of Mathematics Science and Technology Education*, 17(6), em1974. <https://doi.org/10.29333/ejmste/10900>
- Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95(1), 148–162. <https://doi.org/10.1037//0022-0663.95.1.148>
- Gaddam, V. L., Gowda, V., & Vaidyanathan, K. (2016). Comparison of the effectiveness of horizontal integration with traditional teaching approach in first-year MBBS students. *International Journal of Medical Science and Public Health*, 5(7), 1360–364. <https://doi.org/10.5455/ijmsph.2016.07102015202>

- Galindo-Domínguez, H., & Bezanilla, M. J. (2019). A systematic review of flipped classroom methodology at university level in Spain. *International Journal of Technology and Educational Innovation*, 5(1), 81–90. <https://doi.org/10.24310/innoeduca.2019.v5i1.4470>
- Gholami Pasand, P., Amerian, M., Dowlatabadi, H., & Mohammadi, A. M. (2021). Developing EFL learners' intercultural sensitivity through computer-mediated peer interaction. *Journal of Intercultural Communication Research*, 1–17. <https://doi.org/10.1080/17475759.2021.1943496>
- Gianoni-Capenakas, S., Lagravere, M., Pacheco-Pereira, C., & Yacyshyn, J. (2019). Effectiveness and perceptions of flipped learning model in dental education: A systematic review. *Journal of Dental Education*, 83(8), 935–945. <https://doi.org/10.21815/JDE.019.109>
- Gillette, C., Rudolph, M., Kimble, C., Rockich-Winston, N., Smith, L., & Broedel-Zaugg, K. (2018). A meta-analysis of outcomes comparing flipped classroom and lecture. *American Journal of Pharmaceutical Education*, 82(5), 6898. <https://doi.org/10.5688/ajpe6898>
- Gomez-Lanier, L. (2018). Building collaboration in the flipped classroom: A case study. *International Journal for the Scholarship of Teaching and Learning*, 12(2). <https://doi.org/10.20429/ijstl.2018.120207>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Gredler, M. E. (2009). *Learning and instruction: Theory into practice*. (6th ed.). Pearson Education Limited.
- Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *British Medical Journal*, 331(7524), 1064–1065. <https://doi.org/10.1136/bmj.38636.593461.68>
- Hamdan, N., McKnight, P., McKnight, K., & Arfstrom, K. M. (2013). A review of flipped learning. Retrieved January 10, 2023, from <http://www.flippedlearning.org>
- Hartling, L., Chisholm, A., Thomson, D., & Dryden, D. M. (2012). A descriptive analysis of overviews of reviews published between 2000 and 2011. *PLOS ONE*, 7(11), e49667. <https://doi.org/10.1371/journal.pone.0049667>
- Hartling, L., Vandermeer, B., & Fernandes, R. M. (2014). Systematic reviews, overviews of reviews and comparative effectiveness reviews: a discussion of approaches to knowledge synthesis. *Evidence-Based Child Health: A Cochrane Review Journal*, 9(2), 486–494. <https://doi.org/10.1002/ebch.1968>
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Higgins, J. P. T., López-López, J. A., Becker, B. J., Davies, S., Dawson, S., Grimshaw, J., McGuinness, L. A., Moore, T. H., Rehfuess, E., Thomas, J., & Caldwell, D. M. (2019). Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health*, 4, e000858. <https://doi.org/10.1136/bmjgh-2018-000858>
- Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (Eds.). (2023) *Cochrane handbook for systematic reviews of interventions version 6.4*. Retrieved September 10, 2023, from [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook).
- Holly, C., Salmond, S., & Saimbert, M. (Eds.). (2017). *Comprehensive systematic review for advanced practice nursing*. <https://doi.org/10.1891/9780826131867>
- Hong, Q. N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice*, 24(3), 459–467. <https://doi.org/10.1111/jep.12884>
- Hsieh, J. C., Wu, W. V., & Marek, M. W. (2016). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, 30(1–2), 1–21. <https://doi.org/10.1080/09588221.2015.1111910>

- Hu, R., Gao, H., Ye, Y., Ni, Z., Jiang, N., & Jiang, X. (2018). Effectiveness of flipped classrooms in Chinese baccalaureate nursing education: A meta-analysis of randomized controlled trials. *International Journal of Nursing Studies*, 79, 94–103. <https://doi.org/10.1016/j.ijnurstu.2017.11.012>
- Huang, H. L., Chou, C. P., Leu, S., You, H. L., Tiao, M. M., & Chen, C. H. (2020). Effects of a quasi-experimental study of using flipped classroom approach to teach evidence-based medicine to medical technology students. *BMC Medical Education*, 20(1), <https://doi.org/10.1186/s12909-020-1946-7>
- Huang, Y., Silitonga, L. M., & Wu, T. (2022). Applying a business simulation game in a flipped classroom to enhance engagement, learning achievement, and higher-order thinking skills. *Computers & Education*, 183, 104494. <https://doi.org/10.1016/j.compedu.2022.104494>
- Ioannidis, J. P. (2009). Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *Canadian Medical Association Journal*, 181(8), 488-493. <https://doi.org/10.1503/cmaj.081086>
- Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research Synthesis Methods*, 8(3), 290-302. <https://doi.org/10.1002/jrsm.1240>
- Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE-Life Sciences Education*, 14, 1–12. <https://doi.org/10.1187/cbe.14-08-0129>
- Jiang, M. Y., Jong, M. S., Lau, W. W., Chai, C., Liu, K. S., & Park, M. (2020). A scoping review on flipped classroom approach in language education: challenges, implications and an interaction model. *Computer Assisted Language Learning*, 35(5–6), 1218–1249. <https://doi.org/10.1080/09588221.2020.1789171>
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26. <https://doi.org/10.3102/0013189x033007014>
- Joo, S. (2025). Flipped classroom and experiential learning for international marketing courses. In *Flipping the classroom with experiential and community-engaged learning in business*. Edward Elgar Publishing. <https://doi.org/10.4337/9781035321629.00019>
- Kanjug, I., Srisawasdi, N., Chaijaroen, S., & Kanjug, P. (2018). Using constructivist instructional design for flipped classroom to enhancing cognitive learning performance. In T.-T. Wu et al. (Eds.), *Innovative technologies and learning (LNCS, Vol. 11003, pp. 128-137)*. Springer. [https://doi.org/10.1007/978-3-319-99737-7\\_13](https://doi.org/10.1007/978-3-319-99737-7_13)
- Khlaisang, J., & Teo, T. (2024). An innovation-based virtual flipped learning system in a ubiquitous learning environment the 21st century skills of higher education learners. *Educational Technology & Society*, 27(1), 100–116. <https://www.jstor.org/stable/48754844>
- Khosa, A., & Burch, S. (2023). Flipped classroom: Students' cognitive needs of relatedness, competence, and autonomy in a fully-flipped program. *Journal of University Teaching and Learning Practice*, 20(5), 1-27. <https://doi.org/10.53761/1.20.5.04>
- King, A. (1993). From sage on the stage to guide on the side. *College Teaching*, 41(1), 30-35. <https://doi.org/10.1080/87567555.1993.9926781>
- Kolb, D. (1984) *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
- Kostaris, C., Stylianos, S., Sampson, D. G., Giannakos, M., & Pelliccione, L. (2017). Investigating the potential of the flipped classroom model in K-12 ICT teaching and learning: An action research study. *Journal of Educational Technology & Society*, 20(1). [https://doi.org/10.1007/978-3-319-33808-8\\_12](https://doi.org/10.1007/978-3-319-33808-8_12)
- Kozikoğlu, I. (2019). Analysis of the studies concerning flipped learning model: A comparative meta-synthesis study. *International Journal of Instruction*, 12(1), 851–868. <https://doi.org/10.29333/iji.2019.12155a>

- Ku, H.-Y., Tseng, H. W., & Akarasriworn, C. (2013). Collaboration factors, teamwork satisfaction, and student attitudes toward online collaborative learning. *Computers in Human Behavior*, 29(3), 922–929. <https://doi.org/10.1016/j.chb.2012.12.019>
- Kung, J., Chiappelli, F., Cajulis, O. O., Avezova, R., Kossan, G., Chew, L., & Maida, C. A. (2010). From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *The Open Dentistry Journal*, 4(1), 84–91. <https://doi.org/10.2174/1874210601004010084>
- Kupers, E., Van Dijk, M., Van Geert, P., & McPherson, G. E. (2013). A mixed-methods approach to studying co-regulation of student autonomy through teacher–student interactions in music lessons. *Psychology of Music*, 43(3), 333–358. <https://doi.org/10.1177/0305735613503180>
- Kyngäs, H., Kääriäinen, M., & Elo, S. (2020). The trustworthiness of content analysis. In Kyngäs, H., Mikkonen, K., Kääriäinen, M. (Eds.), *The application of content analysis in nursing science research* (pp. 41–48). Springer. [https://doi.org/10.1007/978-3-030-30199-6\\_5](https://doi.org/10.1007/978-3-030-30199-6_5)
- Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 3(1), 30 – 43. <https://doi.org/10.2307/1183338>
- Lantu, D. C., Labdhagati, H., Bangun, Y. R., & Sumarli, F. D. (2023). The use of a flipped classroom and experiential learning in an undergraduate management course. *International Journal of Educational Management*, 37(1), 103-116. <https://doi.org/10.1108/ijem-05-2021-0189>
- Lapitan, L. DS., Chan, A. L. A., Sabarillo, N. S., Sumalinog, D. A. G., & Diaz, J. M. S. (2023). Design, implementation, and evaluation of an online flipped classroom with collaborative learning model in an undergraduate chemical engineering course. *Education for Chemical Engineers*, 43, 58–72. <https://doi.org/10.1016/j.ece.2023.01.007>
- Lee, G., & Wallace, A. (2018). Flipped learning in the English as a foreign language classroom: Outcomes and perceptions. *TESOL Quarterly*, 52(1), 62–84. <https://doi.org/10.1002/tesq.372>
- Lee, S. W. Y., & Tsai, C. C. (2011). Students’ perceptions of collaboration, self-regulated learning, and information seeking in the context of Internet-based learning and traditional learning. *Computers in Human Behavior*, 27(2), 905–914. <https://doi.org/10.1016/j.chb.2010.11.016>
- Lin, C., & Hwang, G. (2018). A learning analytics approach to investigating factors affecting EFL students’ oral performance in a flipped classroom. *Educational Technology & Society*, 21(2), 205–219. Retrieved January 5, 2023, from [http://www.j-ets.net/ETS/journals/21\\_2/18.pdf](http://www.j-ets.net/ETS/journals/21_2/18.pdf)
- Lin, H., & Pryor, M. (2019). Technology-facilitated socialised learning in a flipped classroom: Towards a conceptual framework. 2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET), 1-6. <https://doi.org/10.1109/ithet46829.2019.8937343>
- Lizarondo, L., Stern, C., Carrier, J., Godfrey, C., Rieger, K., Salmond, S., Apostolo, J., Kirkpatrick, P., & Loveday, H. (2019). Chapter 8: Mixed methods systematic reviews. *JBIR Reviewer’s Manual*. <https://doi.org/10.46658/jbirm-19-02>
- Lo, C. K., & Hew, K. F. (2020). Developing a flipped learning approach to support student engagement: A design-based research of secondary school mathematics teaching. *Journal of Computer Assisted Learning*, 37(1), 142–157. <https://doi.org/10.1111/jcal.12474>
- Lockwood, C., & White, S. (2012). *Synthesizing descriptive evidence*. Lippincott Williams & Wilkins.
- Lundin, M., Rensfeldt, A. B., Hillman, T., Lantz-Andersson, A., & Peterson, L. (2018). Higher education dominance and siloed knowledge: A systematic review of flipped classroom research. *International Journal of Educational Technology in Higher Education*, 15(1), <https://doi.org/10.1186/s41239-018-0101-6>
- Martin, J. P. (2018). *Skills for the 21st century: Findings and policy lessons from the OECD survey of adult skills*. Retrieved January 20, 2023, from <https://docs.iza.org/pp138.pdf>

- Mayer, R. E. (2012). Information processing. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook*, Vol. 1. Theories, constructs, and critical issues (pp. 85–99). American Psychological Association. <https://doi.org/10.1037/13273-004>
- Miller, C., & Ahmad, Y. (2000). Collaboration and partnership: an effective response to complexity and fragmentation or solution built on sand? *International Journal of Sociology and Social Policy*, 20(5/6), 1-38. <https://doi.org/10.1108/01443330010789151>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Prisma-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1). <https://doi.org/10.1186/2046-4053-4-1>
- Motameni, R. (2018). The combined impact of the flipped classroom, collaborative learning, on students' learning of key marketing concepts. *Journal of University Teaching and Learning Practice*, 15(3), 47–65. <https://doi.org/10.53761/1.15.3.4>
- Munir, M. T., Baroutian, S., Young, B. R., & Carter, S. (2018). Flipped classroom with cooperative learning as a cornerstone. *Education for Chemical Engineers*, 23, 25–33. <https://doi.org/10.1016/j.ece.2018.05.001>
- National Research Council. (2005). *How students learn: History, mathematics, and science in the classroom*. National Academies Press.
- Neuendorf, K. A. (2018). Content analysis and thematic analysis. *Advanced Research Methods for Applied Psychology*, 211–223. <https://doi.org/10.4324/9781315517971-21>
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom. *Theory and Research in Education*, 7(2), 133–144. <https://doi.org/10.1177/1477878509104318>
- Noguera Fructuoso, I., Albó, L., & Beardsley, M. (2022). University students' preference for flexible teaching models that foster constructivist learning practices. *Australasian Journal of Educational Technology*, 38(4), 22–39. <https://doi.org/10.14742/ajet.7968>
- Oraif, I. M. K. (2018). An investigation into the impact of the flipped classroom on intrinsic motivation (IM) and learning outcomes on an EFL writing course at a university in Saudi Arabia based on self-determination theory (SDT) [Doctoral dissertation, University of Leicester].
- Park, J. H., Han, W. S., Kim, J., & Lee, H. (2020). Strategies for flipped learning in the health professions education in South Korea and their effects: A systematic review. *Education Sciences*, 11(1), 9. <https://doi.org/10.3390/educsci11010009>
- Persky, A. M., & McLaughlin, J. E. (2017). The flipped classroom – from theory to practice in health professional education. *American Journal of Pharmaceutical Education*, 81(6), 118. <https://doi.org/10.5688/ajpe816118>
- Phan, T. A., & Ninh, T. T. D. (2024). The impact of company field trips and representative image on students' CSR knowledge sharing intentions. *Business and Professional Communication Quarterly*. <https://doi.org/10.1177/23294906241237721>
- Pieper, D., Antoine, S.-L., Mathes, T., Neugebauer, E. A., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology*, 67(4), 368–375. <https://doi.org/10.1016/j.jclinepi.2013.11.007>
- Prawat, R. S., & Floden, R. E. (1994). Philosophical perspectives on constructivist views of learning. *Educational Psychologist*, 29(1), 37–48. [https://doi.org/10.1207/s15326985ep2901\\_4](https://doi.org/10.1207/s15326985ep2901_4)
- Price, C. L., & Walker, M. D. (2019). Improving the accessibility of foundation statistics for undergraduate business and management students using a flipped classroom. *Studies in Higher Education*, 46(2), 245–257. <https://doi.org/10.1080/03075079.2019.1628204>

- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809.x>
- Pritchard, A. (2008). Ways of learning. <https://doi.org/10.4324/9780203887240>
- Ramma, Y., Bhoola, A., Jawaheer, S., Gunness, S., Wah, H. T. Y. L. K., Gopee, A. K., & Authelsingh, D. (2021). Teaching and learning science in the 21st century: A study of critical thinking of learners and associated challenges. In J. Naidoo (Ed.), *Teaching and learning in the 21st century* (pp.139–156). [https://doi.org/10.1163/9789004460386\\_009](https://doi.org/10.1163/9789004460386_009)
- Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, 44(3), 159–175. <https://doi.org/10.1080/00461520903028990>
- Ripoll, V., Godino-Ojer, M., & Calzada, J. (2021). Teaching chemical engineering to biotechnology students in the time of COVID-19: Assessment of the adaptation to digitalization. *Education for Chemical Engineers*, 34, 21–32. <https://doi.org/10.1016/j.ece.2020.11.001>.
- Robertson, W. H. (2022). The constructivist flipped classroom. *Journal of College Science Teaching*, 52(2), 17–22. <https://doi.org/10.1080/0047231x.2022.12290644>
- Rodríguez, G., Díez, J., Pérez, N., Baños, J. E., & Carrió, M. (2019). Flipped classroom: Fostering creative skills in undergraduate students of health sciences. *Thinking Skills and Creativity*, 33, 100575. <https://doi.org/10.1016/j.tsc.2019.100575>
- Rodríguez-Chueca, J., Molina-García, A., García-Aranda, C., Pérez, J., & Rodríguez, E. (2019). Understanding sustainability and the circular economy through flipped classroom and challenge-based learning: An innovative experience in engineering education in Spain. *Environmental Education Research*, 26(2), 238–252. <https://doi.org/10.1080/13504622.2019.1705965>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage Publications.
- Sanchez-Gil-Machín, R., Baena Morales, S., Molina-García, N., & Ferriz-Valero, A. (2025). Impact of feedback in flipped learning on the development of soft skills of university students. *Education Sciences*, 15(1), 63. <https://doi.org/10.3390/educsci15010063>
- Sarwar, M. N., Maqbool, M. A., Ullah, S., Rana, A. S., Khan, S. U., Ibrahim, A. A., Alam, K., Zafar, S., Ullah, Z., & Nazar, M. F. (2024). Fostering conceptual understanding of photocatalysis for sustainable development: A social constructivism flipped-classroom model. *Sustainability*, 16(23), 10324. <https://doi.org/10.3390/su162310324>
- Sergis, S., Sampson, D. G., & Pelliccione, L. (2018). Investigating the impact of Flipped Classroom on students' learning experiences: A Self-Determination Theory approach. *Computers in Human Behavior*, 78, 368-378. <https://doi.org/10.1016/j.chb.2017.08.011>
- Sevillano-Monje, V., Martín-Gutiérrez, Á., & Hervás-Gómez, C. (2022). The flipped classroom and the development of competences: A teaching innovation experience in higher education. *Education Sciences*, 12(4), 248. <https://doi.org/10.3390/educsci12040248>
- Schank, R. C. (2006). Epilogue: The fundamental issue in the learning sciences. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences*, (pp.587–592). <https://doi.org/10.1017/cbo9780511816833.037>
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., ... & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358, j4008. <https://doi.org/10.1136/bmj.j4008>
- Sheldon, K. M., Ryan, R. M., & Reis, H. T. (1996). What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin*, 22(12), 1270–1279. <https://doi.org/10.1177/01461672962212007>

- Shih, H. J., & Huang, S. C. (2019). College students' metacognitive strategy use in an EFL flipped classroom. *Computer Assisted Language Learning*, 33(7), 755–784. <https://doi.org/10.1080/09588221.2019.1590420>
- Shih, W., & Tsai, C. (2019). Effect of flipped classroom with BOPPPS model on learners' learning outcomes and perceptions in a business etiquette course. *The Asia-Pacific Education Researcher*, 29(3), 257–268. <https://doi.org/10.1007/s40299-019-00475-z>
- Skinner, B. F. (1953). *Science and human behavior*. Macmillan Company.  
<https://doi.org/10.1002/sce.37303805120>
- Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection. *Educational and Psychological Measurement*, 69(3), 493–525. <https://doi.org/10.1177/0013164408323233>
- Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, 100(4), 765–781. <https://doi.org/10.1037/a0012840>
- Steen-Utheim, A. T., & Foldnes, N. (2017). A qualitative investigation of student engagement in a flipped classroom. *Teaching in Higher Education*, 23(3), 307–324. <https://doi.org/10.1080/13562517.2017.1379481>
- Stroet, K., Opdenakker, M., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review*, 9, 65–87. <https://doi.org/10.1016/j.edurev.2012.11.003>
- Sullivan, J. M. (2022). Flipping the classroom: An innovative approach to graduate nursing education. *Journal of Professional Nursing*, 38, 40–44. <https://doi.org/10.1016/j.profnurs.2021.11.005>
- Taguma, M., Feron, E., & Lim, M. H. (2018). Future of education and skills 2030: Conceptual learning framework. OECD Publishing. Retrieved February 10, 2023, from <https://www.oecd.org/education/2030-project/about/documents/>
- Thomas, J., Harden, A., Oakley, A., Oliver, S., Sutcliffe, K., Rees, R., Brunton, G., & Kavanagh, J. (2004). Integrating qualitative research with trials in systematic reviews. *BMJ*, 328(7446), 1010–1012. <https://doi.org/10.1136/bmj.328.7446.1010>
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. John Wiley & Sons.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Whittemore, R., Chao, A., Jang, M., Minges, K. E., & Park, C. (2014). Methods for knowledge synthesis: An overview. *Heart & Lung*, 43(5), 453–461. <https://doi.org/10.1016/j.hrtlng.2014.05.014>
- Williams, R. (2015). *Higher-order thinking skills: Challenging all students to achieve*. Skyhorse Publishing.
- Williams, J., Ayadi, O. F., Perkins, C., & Hyman, L. M. (2018). The flipped class: Experiential learning manifested. *Southwestern Business Administration Journal*, 17(1), 1.
- Winarti, A., Yuanita, L., & Nur, Moh. (2019). The effectiveness of multiple intelligences-based teaching Strategy in enhancing the multiple intelligences and science process skills of junior high school students. *Journal of Technology and Science Education*, 9(2), 122. <https://doi.org/10.3926/jotse.404>
- Wu, W.-C. V., Hsieh, J. C., & Yang, J. C. (2023). Effects of flipped reading–writing constructivist instruction on EFL learners' writing performance and intercultural sensitivity. *The Asia-Pacific Education Researcher*, 32(1), 123–139. <https://doi.org/10.1007/s40299-021-00639-w>
- Wu, Q., Han, Z., Hao, Z., Chen, J., & Pan, Y. (2025). Impact of flipped classroom method in physical education on the intrinsic motivation, self-efficacy, and learning satisfaction: A meta-analysis. *PLoS One*, 20(5), e0324609. <https://doi.org/10.1371/journal.pone.0324609>

- Xiao, L., Larkins, R., & Li-Jun, M. (2018). Track effect: Unraveling the enhancement of college students' autonomous learning by using a flipped classroom approach. *Innovations in Education and Teaching International*, 1–12. <https://doi.org/10.1080/14703297.2017.1415815>
- Xu, S., & Feng, W. (2017). Comparative research on the application of traditional classroom teaching model and flipped classroom teaching model—taking modern educational technology course as an example. In *Proceedings of the 2nd International Seminar on Education Innovation and Economic Management (SEIEM 2017)*. <https://doi.org/10.2991/seiem-17.2018.4>
- Yang, Q., & Chen, C. (2024). The effectiveness of flipped classroom on autonomous learning ability: a meta-analysis. *Asia Pacific Education Review*. <https://doi.org/10.1007/s12564-024-10013-2>
- Ye, X. D., Chang, Y. H., & Lai, C. L. (2019). An interactive problem-posing guiding approach to bridging and facilitating pre-and in-class learning for flipped classrooms. *Interactive Learning Environments*, 27(8), 1075–1092. <https://doi.org/10.1080/10494820.2018.1495651>
- Yu, Z., & Wang, G. (2016). Academic achievements and satisfaction of the clicker-aided flipped business English writing class. *Journal of Educational Technology & Society*, 19(2), 298–312. <http://www.jstor.org/stable/jeductechsoci.19.2.298>
- Zainuddin, Z., & Perera, C. J. (2017). Exploring students' competence, autonomy and relatedness in the flipped classroom pedagogical model. *Journal of Further and Higher Education*, 43(1), 115–126. <https://doi.org/10.1080/0309877x.2017.1356916>
- Zhang, X., & Zhou, M. (2019). Interventions to promote learners' intercultural competence: A meta-analysis. *International Journal of Intercultural Relations*, 71, 31–47. <https://doi.org/10.1016/j.ijintrel.2019.04.006>
- Zhao, L., He, W., & Su, Y. S. (2021). Innovative pedagogy and design-based research on flipped learning in higher education. *Frontiers in Psychology*, 12, 577002. <https://doi.org/10.3389/fpsyg.2021.577002>

## Supplementary Data of the Included Reviews

No	Author/Year	Review type	Periods covered	Database	Review details				
					Study number	Field	Learner type	Language	Region
1	Akcayır & Akcayır, 2018	QSR	1980–2016	1	71 articles	Education	NA	English	NA
2	Aljaber et al., 2023	QSR	inception to April 2021	8	35 articles	Pharmacy	undergraduate	English	United states, UK, Malasia, Saudi Arabia, China, Australia, Lebanon, Thailand
3	Arslan, 2020	QSR	published after 2012	6	78 studies	EFL/ESL	NA	English	NA
4	Aybirdi et al., 2023	MA	2015-2021	6	40 studies	EFL/ESL	Mixed	English	Asia, USA, China, Middle East, Turkiye
5	Bredow et al., 2021	CMA	2018- May 2019	5	317 studies	Cross-disciplinary	Undergraduate and graduate	English	Asia, Europe, Latin America, Middle East, North America/ Australia
6	Cevikbas & Kaiser, 2022	QSR	inception to 16 June 2021	4	97 studies	Mathematics	K-12, undergraduate, and graduate	English	NA
7	Chen et al., 2023	MA	published after 2000	7	26 studies	EFL/ESL	undergraduate	English	Asia, Europe, North America
8	Cui et al., 2023	MA	inception to October 10th, 2022	7	22 articles	Pharmacy	Mixed	English and Chinese	USA, China, Thailand, Malaysia, Qatar
9	Dogan et al., 2021	MM	2012-2019	4	40 studies	Science	K-12 and undergraduate	English and Turkish	NA
10	Farmus et al., 2020	MA	inception to March 2020	3	10 articles	Introductory statistics	Undergraduate	English	NA
11	Ge et al., 2020	MA	inception to February, 2020	6	19 articles	Radiology	Undergraduate and graduate	English	NA
12	Guler et al., 2023	MA	January, 2012 – March, 2020	7	37 articles	Mathematics	K-12, undergraduate, and graduate	English	Asia, Europe, USA, China
13	Jang & Kim, 2020	CMA	2012-2017	5	43 studies	Cross-disciplinary	Undergraduate	English	NA
14	Kazeminia et al., 2022	QSR	inception to April 2021	7	26 articles	Anatomy	undergraduate	English	Asia, Europe, USA, Canada, China, Australia, Middle East
15	Li et al., 2020	QSR	inception to June 2019	2	59 studies	Health professions	undergraduate	Chinese	China
16	Li et al., 2020	MA	inception to January, 2020	6	32 articles	Nursing	Undergraduate, higher vocational school	Chinese	China
17	Li et al., 2022	MA	no date restriction	6	13 articles	Medical microbiology	undergraduate	English and Chinese	NA

18	Linling & Abdullah, 2023	QSR	inception to October 8, 2022	2	22 articles	EFL	Mixed	English	Iran, China, Indonesia, Turkiye
19	Lo & Hew, 2019	MA	2008-2017	5	31 articles	Engineering	K-12 and undergraduate	English	USA, Japan, Taiwan, Lebanon Singapore
20	Lo et al., 2017	MM	2012-2016	7	82 articles	Mathematics	K-12 and undergraduate	English	USA, Taiwan, Canada, Australia, Germany, New Zealand
21	Naing et al., 2023	MA	inception to April 2022	10	45	Health professions	undergraduate	English	USA
22	Ni et al., 2023	MA	2000 -December 2020	5	24 articles	EFL	undergraduate	English	China, Other regions
23	Ozbay & Cinar, 2021	QSR	2013-2020	3	24 articles	Nursing	Undergraduate and graduate	English	USA, Korea, Australia, Spain, China, Norway, Egypt
24	Park & Suh, 2021	MA	January, 2000 – December, 2020	9	18 studies	Nursing	undergraduate	English and Korean	NA
25	Peng et al., 2022	MA	inception to 30 June 2021	7	11 studies	Pharmacy	undergraduate and junior college	English and Chinese	China
26	Senali et al., 2022	QSR	inception to July-2021	2	30 articles	Business	NA	English	Asia, Europe, USA, China, Australia
27	Shahnama et al., 2021	MA	published after 2018	7	69 studies	EFL/ESL	Mixed	English	NA
28	Sopamena et al., 2023	MA	2015 to 2021	7	13 studies	Mathematics	undergraduate	English	NA
29	Strelan et al., 2020a	CMA	published before 2018	7	53 articles	Cross-disciplinary	K-12 and undergraduate	English	NA
30	Strelan et al., 2020b	CMA	published before 2018	7	198 articles	Cross-disciplinary	K-12 and undergraduate	English	NA
31	Turan & Akdag-Cimen, 2019	QSR	published before 2018	4	43 articles	EFL	NA	English	Taiwan, China, Turkiye, Arabia, U.S, Korea, Indonesia
32	Turan, 2023	MM	Inception to October, 2019	4	82 studies	Science	K-12, K-16, and undergraduate	English	Asia, Europe, USA, Canada, China, Australia, Middle East, Turkiye, South Africa
33	Tutal & Yazar, 2021	CMA	July 2017 -February 2020	11	203 studies	Cross-disciplinary	K-12 and undergraduate	English and Turkish	Turkiye and NA
34	Vitta & Al-Hoorie, 2020	MA	inception to August 2019	6	56 studies	ESL	NA	English	NA

35	Wagner et al., 2020	CMA	2012-2018	3	25 studies	Cross-disciplinary	K-12	English	USA, Hong Kong, Greece
36	Wright & Park, 2021	QSR	no date restriction	2	30 studies	Science and mathematics	K-16	English	NA
37	Xiao et al., 2022	MA	inception to May 2021	3	25 articles	Traditional Chinese medicine	undergraduate	English and Chinese	China
38	Xu et al., 2019	MA	published before 2018	5	22 articles	Nursing	K-12 and undergraduate	English and Chinese	China
39	Zhang et al., 2021	CMA	2000-2020	3	20 articles	Cross-disciplinary	undergraduate	English and Chinese	NA
40	Zheng et al., 2020	CMA	2013-2019	3	95 articles	Cross-disciplinary	K-12 and undergraduate	English	Africa, Asia, Europe, USA, China
41	Zhu, 2021	CMA	October 2017 -March 2020	7	27 studies	Cross-disciplinary	K-12	English	NA

*Note: QSR = Qualitative systematic review; MA = Meta analysis; MM = Mixed-method review; CMA = Cross-disciplinary meta-analysis*