E-ISSN: 3164-8940
https://doi.org/10.47945/create.v1i1.3353

# Flipped Classroom Models for Advanced Electrodynamics in Higher Education: A Systematic Review of Pedagogical Frameworks, Learning Outcomes, and Evidence Gaps

**Jumran** | **Nurnatasya Waode**

[a]Physics Education, Universitas Negeri Gorontalo, Indonesia;
[b]State Vocational High School No. 3 Baubau, Southeast Sulawesi, Indonesia..

## ABSTRACT

Flipped classroom models are widely promoted as a way to free class time for problem solving in mathematically demanding physics courses, yet it is unclear what evidence supports their use in advanced electrodynamics. This systematic review, reported with reference to PRISMA 2020, asked (1) which pedagogical and theoretical frameworks underpin flipped designs relevant to advanced electrodynamics and how they address its conceptual and mathematical abstraction, and (2) what is known about the effects of flipping on conceptual understanding, problem solving, and self-regulated learning in university electromagnetism. Studies were screened against a tiered eligibility frame (direct, proximal, and contextual evidence), appraised with the Mixed Methods Appraisal Tool, and synthesised thematically. None of the ten records included in the initial draft met the eligibility criteria, as all were school-level studies or non-empirical reviews. No peer-reviewed study of flipping in an advanced or upper-division electrodynamics course was identified. Four proximal studies of flipped designs in university electricity and magnetism or electrical courses reported positive but heterogeneous results. These included higher examination scores in a flipped problem-based course on electromagnetic induction (d = 0.87, non-causal design), gains on a standard electricity and magnetism assessment that differed by academic level and gender, and improved final-examination scores alongside student resistance to pre-recorded videos. Meta-analyses in higher education and STEM support small-to-moderate average benefits, but these benefits depend on design features such as pre-class accountability and on the active learning that flipping enables. Social constructivism, cognitive load theory, self-regulated learning, and the revised Bloom taxonomy were invoked mainly at a descriptive level. The review proposes a flipped design for advanced electrodynamics that aligns these frameworks with documented mathematical difficulties, and it sets out a research agenda that prioritises controlled studies using validated upper-division instruments.

**Keywords:** Flipped Classroom; Advanced Electrodynamics; Cognitive Load Theory; Self-Regulated Learning.

**Correspondence:** waodenurnatasya@gmail.com | **Copyright:** This open access article is distributed under a CC BY-SA License
**Article History:** Received: 14 April 2026 | Accepted: 12 May 2026 | **Homepage**: https://e-jurnal.iainsorong.ac.id/index.php/great
**Citation:** Jumran, J., & Waode, N. (2026). Flipped Classroom Models for Advanced Electrodynamics in Higher Education: A Systematic Review of Pedagogical Frameworks, Learning Outcomes, and Evidence Gaps . CREATE: Contemporary Research in Education and Teaching, 1(1), 43–57. https://doi.org/10.47945/create.v1i1.3353

## I. INTRODUCTION

Advanced electrodynamics is one of the core courses in physics programs that places the greatest cognitive demands on students. Students must reason about invisible fields while working with vector calculus, partial differential equations, and boundary conditions. Physics education research (PER) shows that difficulties at this level are not purely conceptual. In upper-division electricity and magnetism courses, students struggle with Gauss's law, vector calculus, and electric potential. These difficulties stem from integrating calculations with physical concepts, visualizing spatial situations, and accessing appropriate mathematical resources (Pepper et al., 2012). Similar difficulties are evident even at the introductory level: many students lack a coherent understanding of classical field theory (Zuza et al., 2018), and introductory textbooks also show inconsistencies in presenting the relationships between electromagnetic fields and their sources, Maxwell's equations, and electromagnetic waves (Suárez et al., 2023). Thus, the pedagogical challenges of electrodynamics cannot be adequately addressed simply by adding more derivations to the board. The issue lies in how students are supported in connecting mathematical manipulation with physical meaning.

Strong empirical evidence indicates that active learning is more effective than lecturing in science and engineering education. A meta-analysis of 225 studies in undergraduate STEM courses found that examination scores and concept-inventory scores increased by an average of 0.47 standard deviations in active-learning classes, whereas the odds ratio for failure in lecture classes reached 1.95 (Freeman et al.,

2014). In physics, Peer Instruction has been documented to improve conceptual reasoning and quantitative problem solving over a decade of implementation (Crouch & Mazur, 2001). However, active learning presents a paradox. In a controlled experiment in introductory physics, students in active-learning classes learned more but felt that they had learned less than their peers in lecture classes (Deslauriers et al., 2019). This finding is important because it helps explain why effective innovations may be resisted by students.

The flipped classroom offers a structure for extending active-learning time. Content delivery is moved outside the classroom, typically through videos or readings, while face-to-face time is used for interactive activities. Abeysekera and Dawson (2015) define this model as an approach that moves information delivery outside the classroom and uses class time for active and social activities, while requiring completion of pre- and/or post-class tasks. They also argue that this design has the potential to manage cognitive load and increase motivation, while noting that evidence of effectiveness and consistency in its definition remain limited. Reviews of the early literature also indicate that the model has advantages, such as flexibility and increased interaction, but also challenges, including student unpreparedness, instructor workload, and video quality (Akçayır & Akçayır, 2018).

Theoretically, flipped designs are usually explained through four complementary frameworks. First, Vygotsky's social constructivism situates learning within social interaction and the zone of proximal development (ZPD), that is, the gap between what learners can achieve independently and what they can achieve with assistance from a more competent other (Vygotsky, 1978). This framework justifies using class time for discussion and scaffolding. Second, cognitive load theory (CLT) is grounded in the limitations of working memory. Conventional problem solving can overload working memory and hinder schema formation (Sweller, 1988). Therefore, instructional design needs to manage load that is irrelevant to learning goals (Sweller et al., 2019). Third, the revised Bloom taxonomy distinguishes cognitive processes from remembering to creating (Krathwohl, 2002). In the flipped logic, lower-level processes are moved to the pre-class phase and higher-order processes are facilitated in class. Fourth, self-regulated learning (SRL) views learners as agents who set goals, select strategies, monitor progress, and reflect on their learning outcomes (Zimmerman, 2002). This capability is a prerequisite for successful pre-class learning.

The body of synthesis evidence on the effectiveness of flipped classrooms in higher education is substantial, but not uniform. A meta-analysis of 317 studies found advantages for flipped learning on seven of eight outcomes, with effects ranging from 0.20 to 0.53, and showed that educational context was the strongest moderator (Bredow et al., 2021). Another meta-analysis reported small positive effects that were larger when face-to-face time was not reduced and quizzes were used (van Alten et al., 2019), as well as effects that decreased from $g = 0.35$ to $g = 0.24$ when the analysis was restricted to adequately powered studies (Låg & Sæle, 2019). In higher-education STEM, a synthesis of 53 articles reported $g = 0.263$ (Gong et al., 2024). Positive findings have also been reported across disciplines (Strelan et al., 2020) and in engineering education (Lo & Hew, 2019). However, Jensen et al. (2015) showed that flipped and non-flipped classes that were both active produced equivalent outcomes. This suggests that some of the advantages of flipping may derive from active learning rather than from the reversal of instructional sequence itself.

At least four gaps can be identified from this literature. The first is theoretical. A scoping review of 435 articles found that approximately 65% of articles did not explicitly connect their research to theory (Li et al., 2023), so claims about "dominant frameworks" are often based only on descriptive references. The second is contextual. A systematic review of flipped classrooms in physics education examined 30 publications on flipped classrooms in physics education generally (Tunggyshbay et al., 2023), whereas a review of 25 systematic reviews covered 18 disciplines (Yang & Valcke, 2026). None

of these reviews focused their synthesis on advanced theoretical physics courses such as electrodynamics. The third is empirical: it remains unclear whether any studies have tested flipping in advanced electrodynamics using validated upper-division instruments, such as the CUE (Chasteen, Pepper, et al., 2012). The fourth is methodological. SRL in flipped environments is generally measured using self-report questionnaires (Silverajah et al., 2022), so claims of increased learning autonomy should be interpreted cautiously.

Based on these gaps, this review formulates two research questions (RQs). RQ1: What theoretical and pedagogical frameworks underpin flipped classroom designs relevant to advanced electrodynamics, and how do these frameworks address the conceptual and mathematical abstraction of the material? RQ2: What is known about the effects of flipped classrooms on conceptual understanding, problem solving, and learning autonomy among university students in electromagnetism, and how strong is the evidence? This review contributes in three ways. Empirically, it transparently maps the availability of evidence across three levels of proximity. Theoretically, it assesses the extent to which frequently cited frameworks are actually operationalized. Practically, it proposes a flipped design for advanced electrodynamics derived from PER evidence on students' mathematical difficulties, together with a testable research agenda.

## II. RESEARCH METHODS

This study is a systematic review reported with reference to the PRISMA 2020 guidelines. These guidelines include a 27-item checklist and flow diagram to ensure transparent reporting of the rationale, procedures, and findings of the review (Page et al., 2021). Because direct evidence on advanced electrodynamics was expected to be very limited, this review uses a tiered eligibility framework (Table 2). Studies closest to the target population were explicitly separated from contextual evidence so that findings from other contexts were not treated as though they applied directly to advanced electrodynamics.

Eligibility criteria were formulated using the PICOS framework (Table 1). The publication range was set at 2015–2026 consistently across all stages. In the initial draft, the introduction stated 2015–2026, whereas the methods stated 2016–2026; this inconsistency has been standardized. School-level studies were excluded because their populations, mathematical depth of the material, and self-regulation demands differ fundamentally from those of advanced physics students. Literature reviews and SLRs were not counted as included studies but were used as contextual evidence (Tier 3).

Table 1. Eligibility criteria based on PICOS

| Element | Inclusion Criteria | Exclusion Criteria |
|---|---|---|
| Population (P) | Undergraduate or graduate students enrolled in courses in electrodynamics, electricity and magnetism, electromagnetic theory, or university-level electricity courses | Elementary, junior high, or senior high/vocational school students, or participants in nonformal training |
| Intervention (I) | Flipped classroom/flipped learning (core material is studied before class; class time is used for active activities), including peer-instruction flipped and flipped PBL variants | Online/blended learning without reversing the sequence of activities; single-medium instruction without a flipped design |
| Comparator (C) | Conventional lectures, other flipped variants, or no comparator (qualitative studies) | - |
| Outcomes (O) | Conceptual understanding, problem solving, learning autonomy (SRL), and relevant affective outcomes | Studies reporting only the feasibility/validity of instructional materials or devices without student data |

| Element | Inclusion Criteria | Exclusion Criteria |
|---|---|---|
| Design (S) | Quantitative, qualitative, or mixed-methods empirical studies in peer-reviewed journals, 2015–2026, in English or Indonesian | Literature reviews, SLRs, conference proceedings, theses, books, and opinions |

Table 2. Tiered eligibility framework

| Evidence Tier | Operational Definition | Function In The Synthesis |
|---|---|---|
| Direct (Tier 1) | Empirical flipped studies in advanced electrodynamics/upper-division E&M courses | Directly addresses RQ1 and RQ2 |
| Proximal (Tier 2) | Empirical flipped studies in university-level electricity and magnetism or electricity/electrical engineering courses | Provides the closest evidence; transfer to advanced electrodynamics is critically assessed |
| Contextual (Tier 3) | Meta-analyses, systematic reviews, and scoping reviews of flipped classrooms in higher education, STEM, or physics; PER studies on difficulties in learning E&M | Frames effect sizes, moderators, theoretical foundations, and domain demands; not counted as included studies |

The initial draft used only Google Scholar. An evaluation of 28 academic search systems showed that Google Scholar is not suitable as the primary source for systematic reviews because, among other limitations, it does not support precise Boolean queries and its results are not reproducible (Gusenbauer & Haddaway, 2020). Therefore, searches were conducted in Scopus, Web of Science, and ERIC as the primary sources. Google Scholar and Garuda were used as supplementary sources to reach national journals indexed in SINTA. The search was conducted in June 2026 using the basic string: ("flipped classroom" OR "flipped learning" OR "inverted classroom") AND (electrodynamic∗ OR electromagneti∗ OR "electricity and magnetism" OR "upper-division physics") AND (universit∗ OR undergraduate OR "higher education"). The Indonesian-language equivalent used was ("flipped classroom" OR "kelas terbalik") AND (elektrodinamika OR "listrik magnet" OR elektromagnetik) AND (mahasiswa OR "perguruan tinggi"). The search was supplemented with backward and forward snowballing from included studies and contextual reviews.

Search records were imported into reference management software (Mendeley) to remove duplicates. Two reviewers independently screened titles and abstracts and then assessed full texts against the eligibility criteria. Disagreements were resolved through discussion, and the level of agreement was reported using Cohen's kappa coefficient. The ten records included in the initial draft were reassessed at the full-text stage, and the reasons for exclusion were reported openly (Table 4).

The initial draft used the Critical Appraisal Skills Programme (CASP) instrument. CASP is indeed the most commonly used instrument in qualitative evidence synthesis (Long et al., 2020), but this review corpus was dominated by quantitative and quasi-experimental designs. Therefore, quality assessment used the 2018 version of the Mixed Methods Appraisal Tool (MMAT). This instrument provides criteria for qualitative, randomized quantitative, non-randomized quantitative, quantitative descriptive, and mixed-methods studies within a single framework (Hong et al., 2018). Assessment results were reported by criterion and were not summed into a single score. Studies were not excluded based on their scores, but study quality was considered when evaluating the strength of the evidence.

Data were extracted using a structured form containing study identity, country, course and topic, educational level, design, sample, flipped-design components (pre-class, in-class, post-class), referenced theoretical frameworks, instruments, and findings. Because designs, outcomes, and instruments were highly heterogeneous, meta-analysis was not conducted. Synthesis used a thematic synthesis approach. This approach proceeds from coding findings, developing descriptive themes, and then developing analytical themes that go beyond the primary studies (Thomas & Harden, 2008). The coding procedure followed six phases of thematic analysis: familiarization with the data, initial coding, searching for themes, reviewing themes, defining and naming themes, and report writing (Braun & Clarke, 2006).

Coding was conducted reflexively, acknowledging the interpretive role of the researchers (Braun & Clarke, 2019). The deductive framework comprised the four theories presented in the introduction, whereas inductive codes were developed for design components and implementation challenges. The strength of evidence for each claim was assessed based on evidence tier, design, and consistency of findings (Table 8).

Table 3. Summary of the PRISMA 2020 selection flow

| PRISMA 2020 stage | Number | Notes |
|---|---|---|
| Records identified Data Base | 11 | Search date: June 2026 |
| Duplicates removed | 17 | Use tools Mendeley |
| Records screened (title/abstract) | 15 | 1 excluded for being outside the topic |
| Full-text reports assessed | 14 | Excluded for school level (6), not empirical studies (4) |
| Tier 1 studies (direct) | 0 | Not identified during editorial verification |
| Tier 2 studies synthesized | 4 | Quantitative: 3; qualitative: 1 |

## III. RESULTS AND DISCUSSION

The ten records included in the initial draft did not meet the eligibility criteria (Table 4). Four records were literature reviews or SLRs, two of which did not address flipped classrooms. The other six records were studies or device-development studies at the junior or senior high school level, or device-validation studies without student learning-outcome data. For example, Apriyanah et al. (2018) studied 43 Grade 10 high-school students, Meda (2025) studied Grade 12 students on alternating current, and Auliza and Widyastuti (2024) studied mathematics learning at the junior high school level. These studies indicate that flipped classrooms have been widely adopted in Indonesian schools, but they cannot serve as a basis for drawing conclusions about advanced electrodynamics in higher education.

Table 4. Reassessment Of The Ten Records In The Initial Draft

| Records | Level/type | Decision and rationale |
|---|---|---|
| Sulistiowati & Kusno (2025) | SLR, mathematics | Excluded: not an empirical study; not physics |
| Auliza & Widyastuti (2024) | Quasi-exp., JHS, mathematics | Excluded: school level; not physics |
| Apriyanah et al. (2018) | Experiment, Grade 10, physics | Excluded: school level (used as local context) |
| Sari et al. (2023) | SLR, physics PBL | Excluded: not an empirical study; not a flipped |
| Riztanti et al. (2025) | SLR, discovery learning | Excluded: not an empirical study; not a flipped |
| Asbanu (2023) | Material/device validation | Excluded: no student learning-outcome data |
| Ridho et al. (2023) | CAR, high school, harmonic motion | Extracted: School Level |
| Suana et al. (t.t.) | R&D, high school | Excluded: school level; material validation |
| Meda (2025) | Quasi-exp, Grade 12, electricity | Excluded: school level (used as local context) |
| Feti & Nope (2026) | Literature study, mathematics | Excluded: not an empirical study; not physics |

During editorial verification, no peer-reviewed study was found that tested a flipped classroom in an advanced electrodynamics or upper-division electricity and magnetism course (Tier 1). Four studies met the Tier 2 criteria (Table 5). Two studies were conducted in first-year electricity and magnetism courses or on electromagnetic induction topics (P1 and P2), one in an upper-division electronics course in electrical engineering (P3), and one in an electricity and magnetism course in a physics education program in Indonesia (P4). Three studies used non-randomized quantitative designs and one used a qualitative design. None of the studies used individual random assignment.

Table 5. Characteristics of proximal studies (Tier 2)

| Code | Study | Context | Design and sample | Main findings |
|---|---|---|---|---|
| P1 | Wang et al. (2026) | First-year university physics, China; electromagnetic induction | Two cohorts (lecture control vs. flipped-PBL), n = 200 (100/100) | The flipped-PBL group had higher examination scores, $t(198) = 6.02$; $p < 0.001$; $d = 0.87$. Questionnaires indicated perceived improvements in problem solving, communication, and motivation. |
| P2 | Serrano et al. (2023) | First-year electricity and magnetism, several engineering programs, Spain | Pre-/post-BEMA; ICT-based flipped classroom in all groups | No differences in BEMA performance across degree programs; significant differences by academic level and gender. Male students outperformed female students on the BEMA |
| P3 | Kattel et al. (2024) | Electronics (upper-division), electrical engineering, United States | Quasi-experiment, data from 2006–2014 (traditional) vs. 2015–2023 (flipped); surveys and evaluation | Final course grades did not differ. Final-examination scores in the flipped class were higher ($t = -2.01$; $p = 0.045$). Students valued collaboration but resisted pre-recorded videos. Individual quizzes helped identify struggling students. |
| P4 | Furqon (2024) | Electricity and magnetism course, physics education, West Java | Qualitative case study; 2 instructors and 25 students; observation and documentation | Online flipped learning was implemented in three phases (pre-class, in-class, post-class) using Google Classroom and Telegram, combining asynchronous and synchronous activities. |

Contextual evidence (Tier 3) comprised seven effectiveness syntheses and three conceptual syntheses (Table 6). In addition, this review compiled PER studies documenting the cognitive and mathematical demands of electromagnetism (Table 7). Both groups of evidence were used to interpret the Tier 2 findings and formulate the proposed design, rather than to claim effects on advanced electrodynamics.

Table 6. Contextual evidence from meta-analyses and systematic reviews

| Synthesis | Scope | Relevant findings |
|---|---|---|
| Bredow et al. (2021) | 317 studies, higher education | Flipped classes outperformed on seven of eight outcomes, with effects of 0.20–0.53. Context (discipline, region) was the strongest moderator. |
| Låg & Sæle (2019) | 272 studies (learning outcomes) | $g = 0.35$, decreasing to 0.24 in adequately powered studies. Satisfaction $g = 0.16$. Testing student preparation tended to increase effects. |
| van Alten et al. (2019) | 114 studies | Small positive effects on learning outcomes and no effect on satisfaction. Effects were larger when face-to-face time was not reduced. |
| Gong et al. (2024) | 53 articles, higher-education STEM | $g = 0.263$ (95% CI [0.190; 0.337]). Performance was associated with implementation strategies, such as bridging activities at the beginning of class and combinations of quizzes and post-class practice. |
| Strelan et al. (2020) | Across disciplines and educational levels | Overall positive effects on student performance, with variation across disciplines. |
| Lo & Hew (2019) | Engineering education (10 years) | Overall positive effects on engineering students' achievement. |
| Tunggyshbay et al. (2023) | 30 publications, physics education | Flipped classrooms in physics had beneficial effects compared with conventional instruction; the review examined instructional activities and psychological mechanisms. |

| Synthesis | Scope | Relevant findings |
|---|---|---|
| Li et al. (2023) | 435 articles, higher education | Approximately 65% of articles did not explicitly connect their research to theory; the remainder tended to be eclectic. |
| Silverajah et al. (2022) | 32 articles (2016–2021), SRL | Students with strong self-regulation performed well in online flipped environments. Self-report questionnaires were the dominant measurement method. |
| Yang & Valcke (2026) | 25 systematic reviews (771 studies, 18 disciplines) | Perceptions were generally positive. The authors proposed the Transformative Activation Theory for Flipped Classrooms (TAT-FC), which combines SRL, socioconstructivism, and active learning in a three-phase model. |

## Theoretical and Pedagogical Frameworks

Theme 1: Theoretical frameworks are more often cited than operationalized. The initial draft stated that social constructivism, CLT, and SRL were the dominant frameworks. However, this claim is not supported by the initial-draft corpus because most studies did not articulate a theory. This pattern is consistent with the finding of Li et al. (2023) that most flipped-learning articles in higher education do not explicitly connect their research to theory. In Tier 2 studies, designs were more often described through operational components. These components included pre-class mini-lectures and collaborative problem solving (P1), the use of ICT in active methods (P2), mini-lectures and individual quizzes (P3), and three asynchronous-synchronous phases (P4). Therefore, the mapping of design components to theory in this section is the reviewers' analytical interpretation, not a claim made by the authors of the primary studies.

Theme 2: The four frameworks operate at different phases. Analytically, CLT and the revised Bloom taxonomy are most relevant to designing the pre-class phase. Both help determine what can be learned independently without overloading working memory (Sweller et al., 2019; Krathwohl, 2002). Social constructivism and the ZPD are most relevant to the in-class phase, when instructors and peers provide scaffolding (Vygotsky, 1978). SRL is relevant across all phases, but is most critical in the pre-class and post-class phases (Zimmerman, 2002). The synthesis of SRL in flipped classrooms shows that students with strong self-regulation perform better in online flipped environments (Silverajah et al., 2022). A study of computer science students also found a strong positive association between SRL behavior, engagement, and final achievement, with engagement serving as a partial mediator (Maiti & Priyaadharshini, 2026). Thus, SRL appears to function more as a prerequisite and moderator of flipped-learning success than simply as an outcome. This differs from the initial-draft assumption that flipping itself increases learning autonomy.

Theme 3: CLT provides the most operational basis for mathematically intensive pre-class materials. Fischer et al. (2023) developed a CLT-based process for evaluating pre-class materials in a flipped medical course. They found that students initially did not allocate the most study time to the most difficult materials. After iterative revision, material efficiency in terms of cognitive load and time increased with a large effect size without increasing the overall workload. Although the context was not physics, this approach can be transferred to electrodynamics. For material with high element interactivity, such as field boundary conditions or retarded potentials, long pre-class videos dense with derivations risk creating excessive cognitive load. This risk was anticipated theoretically by Abeysekera dan Dawson (2015).

Theme 4: Integrative frameworks are beginning to emerge, but have not yet been tested in advanced physics. Yang and Valcke (2026) proposed TAT-FC, which combines SRL,

socioconstructivism, and active learning in a three-phase instructional model. The initial draft attributed this framework to a local undergraduate thesis. That attribution has been corrected. TAT-FC was derived from a review of reviews across disciplines, so its fit with the epistemological demands of electrodynamics still needs to be tested. These demands are documented in the PER literature (Table 7), for example the ACER framework for activation of mathematical tools, model construction, mathematical execution, and reflection on results (Wilcox et al., 2013).

Table 7. Electromagnetism Domain Demands Based On PER Evidence And Their Implications

| Documented difficulty | Evidence | Implications for flipped design |
|---|---|---|
| Integration of mathematics and physics (Gauss's law, vector calculus, potential) | Pepper et al. (2012): difficulties combining calculations with concepts, visualizing spatial situations, and accessing appropriate mathematical resources | Pre-class materials should include refreshers on specific mathematical tools; class should be used to connect mathematical steps with physical meaning |
| Use of mathematics in advanced physics | Wilcox et al. (2013): ACER framework (Activation, Construction, Execution, Reflection) | Class tasks can be organized by ACER phase; reflection on results (limits, dimensions, symmetry) can be routine |
| Understanding field theory | Zuza et al. (2018): many students lack a coherent idea of field theory; Suárez et al. (2023): inconsistencies in the presentation of electromagnetic fields. | Pre-class materials should not consist only of derivations; they should emphasize field characteristics and the distinction between fields and forces |
| Advanced conceptual assessment | Chasteen, Pepper, et al. (2012): CUE as a conceptual assessment of junior-level electrostatics | Flipped electrodynamics studies should use upper-division instruments rather than course grades alone |
| Course transformation | Chasteen, Pollock, et al. (2012): outcomes from research-based junior-level E&M transformation; Zuza et al. (2020): research-based learning sequence on EMF | Flipped class activities can adopt tested research-based materials |

## Effects on Learning Outcomes

Conceptual understanding. Proximal evidence points in a positive direction, but is not uniform. In P2, BEMA performance among first-year engineering students learning through an ICT-based flipped classroom did not differ across degree programs, but did differ by academic level and gender. Male students outperformed female students on the BEMA, although course grades did not differ (Serrano et al., 2023). Because all groups received flipped classroom instruction, this study cannot demonstrate the effect of flipping compared with lecturing. The value of the study lies in its finding that conceptual benefits may be unevenly distributed and that course grades and conceptual instruments may provide different pictures. In P1, examination scores for the flipped-PBL group on electromagnetic induction were higher than those of the lecture group, with a large effect size ($d = 0.87$). However, the authors themselves emphasized that this association cannot be interpreted causally (Wang et al., 2026).

Problem solving. Evidence concerning problem solving is weaker than the claims in the initial draft. In P1, improvement in problem-solving ability was measured only through students' perceptions (Wang et al., 2026). In P3, final-examination scores in the flipped class were significantly higher ($t = -2.01$; $p = 0.045$), but final course grades did not differ (Kattel et al., 2024). This inconsistency across outcomes, together with the p-value being close to the threshold and the nine-year historical comparison, indicates that effects on problem solving cannot yet be established. The initial draft cited a peer-instruction flipped study reporting an "86% effect" on problem solving as evidence for university students (Auliza & Widyastuti, 2024), although that study was conducted with junior high school students in mathematics learning.

Learning autonomy (SRL). The qualitative P4 study describes how an electricity and magnetism course at a university in West Java combined asynchronous and synchronous activities in three phases through Google Classroom and Telegram (Furqon, 2024). This study provides insight into design but did not measure SRL as an outcome. Meanwhile, P3 illustrates another aspect of the demand for autonomy: students valued collaboration but resisted pre-recorded videos. Individual quizzes then served to identify struggling students (Kattel et al., 2024). This finding is consistent with meta-analyses showing that pre-class accountability, such as preparation tests or quizzes, is associated with larger effects (Låg & Sæle, 2019; van Alten et al., 2019; Gong et al., 2024).

Strength of evidence. Table 8 summarizes the strength of evidence for each major claim. Overall, the claim that flipped classrooms "significantly improve" conceptual understanding, problem solving, and learning autonomy among advanced electrodynamics students cannot be sustained. A more defensible claim is limited: flipped classrooms have moderate support in higher education generally, weak and heterogeneous preliminary support in university electricity and magnetism, and no direct evidence in advanced electrodynamics.

Table 8. Evidence strength matrix

| Claim | Supporting evidence | Strength of evidence | Critical notes |
|---|---|---|---|
| Flipping improves learning outcomes in higher education/STEM | Meta-analysis (Tier 3) | Moderate | High heterogeneity; effects decrease in adequately powered studies |
| Flipping improves outcomes in university electricity and magnetism | P1, P2, P3 | Low | Non-randomized/historical designs; P2 had no non-flipped comparator |
| Flipping improves conceptual understanding | No direct studies | Insufficient | The main gap identified by this review |
| Flipping improves problem solving | P1 (perception), P3 (final examination) | Low | No validated problem-solving instrument |
| Flipping improves SRL | P4 (qualitative), SRL synthesis | Low–moderate | SRL is more often a prerequisite than a measured outcome |
| Benefits derive from the flipping sequence itself | Jensen et al. (2015) | Questionable | Effects may derive from active learning rather than reversal of sequence |

The most important finding of this review is that direct evidence on flipped classrooms in advanced electrodynamics is not yet available, at least in the verifiable literature. This finding differs from the initial draft, which concluded that there was "strong convergence" and "cross-study confirmation." The difference arose because the initial draft treated school-level studies and cross-disciplinary reviews as evidence about advanced physics students. Absence of evidence is not the same as evidence of no effect. Nevertheless, the absence of evidence limits what can be responsibly recommended. This gap is consistent with the observation that existing physics syntheses do not specifically focus on advanced courses (Tunggyshbay et al., 2023), whereas research-based transformation studies of junior-level E&M courses generally focus on interactive learning rather than flipping as the primary variable (Chasteen, Pollock, et al., 2012).

At least three reasons prevent average effects from meta-analyses from being directly applied to advanced electrodynamics. First, between-study heterogeneity is high and context is the strongest moderator (Bredow et al., 2021). Thus, cross-disciplinary averages are less informative for courses with extreme characteristics. Second, effects become smaller in adequately powered studies (Låg & Sæle, 2019), so expectations of benefit should be moderate. Third, when the comparison class is also active,

the advantage of flipping may disappear (Jensen et al., 2015). This is consistent with evidence that active learning itself is a major driver of improved achievement in STEM (Freeman et al., 2014). The implication is that the more appropriate question for advanced electrodynamics is not "is flipping effective?" but rather "which flipped components enable high-quality active learning without producing excessive cognitive load during the pre-class phase?".

CLT explains that new information is processed by working memory, which is limited in capacity and duration. These limitations diminish when knowledge has been stored as schemas in long-term memory (Sweller et al., 2019). The implications are important for advanced electrodynamics. Students who have not automated mathematical tools, such as the divergence theorem, curvilinear coordinates, or separation of variables, will experience high cognitive load when studying derivations independently. Conversely, more proficient students may not need overly worked-out examples. The mathematics–physics integration difficulties documented by Pepper et al. (2012) suggest that pre-class materials should include refreshers on specific mathematical tools rather than merely lecture recordings. An iterative evaluation approach such as that used by Fischer et al. (2023) can help instructors identify segments of material that are too demanding before the problem appears as low-quality discussion in class.

Student resistance to pre-recorded videos in P3 (Kattel et al., 2024) can be understood in light of the findings of Deslauriers et al. (2019). Greater cognitive effort in active learning tends to be interpreted by students as evidence of poorer learning. In advanced electrodynamics, the feeling of being "unable to learn independently" may be intensified by the complexity of the material. Thus, the success of flipping depends not only on material design but also on instructors' efforts to explain the pedagogical rationale from the beginning of the semester and provide feedback that demonstrates tangible progress. Peer instruction at the beginning of class can serve a dual function: making progress in understanding visible to students themselves (Crouch & Mazur, 2001).

Based on the synthesis above, Table 9 proposes a four-phase flipped design that aligns theory with domain demands. This design is a proposition that requires testing, not an established practice. Its differences from TAT-FC (Yang & Valcke, 2026) lie in two aspects. First, this design adds an initial in-class transition phase that functions as an accountability mechanism. Second, class activities are organized according to the ACER phases, making the use of mathematics in physics an explicit learning object (Wilcox et al., 2013). Class materials can adapt tested research-based learning sequences, such as a sequence on electromotive force (EMF) that produced improved outcomes compared with a control groupn (Zuza et al., 2020).

Table 9. Proposed Flipped Classroom Design For Advanced Electrodynamics

| Phase | Activities for advanced | Theoretical basis | Indicators/evaluation |
|---|---|---|---|
| Pre-class | Short segmented videos with worked examples; refreshers on mathematical tools | Cognitive load theory (Sweller et al., 2019); revised Bloom taxonomy (Krathwohl, 2002) | Quiz scores; reported study time and cognitive load (PREP approach; Fischer et al., 2023) |
| Initial in-class transition | Conceptual questions with peer instruction to activate and check preparation | Crouch & Mazur (2001); pre-class accountability (Låg & Sæle, 2019) | Distribution of responses before and after discussion |
| In-class | Problem-based group tutorials with phased scaffolding according to the ACER phases | ZPD and social mediation (Vygotsky, 1978); ACER (Wilcox et al., 2013) | Mathematical-physical reasoning rubric; group worksheet |
| Post-class | Individual practice problems, error reflection, and planning for subsequent learning | SRL (Zimmerman, 2002); Silverajah et al. (2022) | Reflection journal; LMS analytics data; SRL instrument |

| | | | |
|---|---|---|---|
| Course evaluation | Pre-/post-tests using advanced conceptual instruments | Chasteen, Pepper, et al. (2012) | Effect sizes, confidence intervals, subgroup analyses |

Theoretically, this review shows that "multi-framework" claims in the flipped-learning literature need to shift from lists of theories to functional mapping: which theory explains which phase, through what mechanism, and with what indicators. Without such mapping, references to constructivism, CLT, or SRL tend to be rhetorical. This is consistent with the critique of theoretical eclecticism by Li et al. (2023). Practically, for electrodynamics instructors in Indonesia, the available evidence supports three cautious steps: (1) retain face-to-face time rather than reducing it (van Alten et al., 2019); (2) include pre-class or early-class quizzes; and (3) evaluate pre-class materials in terms of cognitive load. The widespread adoption of flipped classrooms in Indonesian schools (Apriyanah et al., 2018; Meda, 2025) and experience with online flipped learning in university electricity and magnetism (Furqon, 2024) indicate that infrastructure and pedagogical readiness are beginning to develop. However, rigorous evaluation is still required before this model can be broadly recommended.

Table 10 formulates six research priorities. The primary priority is to generate direct evidence in advanced electrodynamics using an equivalent comparison group and advanced conceptual instruments. Course grades alone are insufficient because they can provide a different picture from conceptual instruments, as seen in P2 (Serrano et al., 2023). Research also needs to separate the effects of flipping from those of active learning and treat SRL as a moderator measured using multiple methods rather than only self-report questionnaires (Silverajah et al., 2022).

Table 10. Research agenda

| Research priority | Recommended Design |
|---|---|
| Direct evidence in advanced electrodynamics | Quasi-experiment with an equivalent comparison class , or a switching-replications design; |
| Separating the effects of flipping from the effects of active learning | Compare active-flipped with active-non-flipped classes, following the logic of Jensen et al. (2015) |
| Cognitive load of mathematically intensive pre-class materials | Measure cognitive load per material unit and conduct iterative revision (Fischer et al., 2023) |
| SRL as a moderator and outcome | Combine questionnaires, learning analytics, and interviews; test SRL moderation of achievement |
| Perceived learning vs. actual achievement | Measure both simultaneously (Deslauriers et al., 2019) to explain student resistance |
| Indonesian context | Multi-campus studies in teacher-training institutions, including campuses with limited internet access; reporting of infrastructure and instructor workload |

This review has several limitations. First, the numbers at each PRISMA stage still need to be completed from the actual search, and the status of Tier 1 (zero studies) needs to be confirmed after searching all databases. Second, some extraction of Tier 2 studies was based on abstracts and publisher information, so details of theory, instruments, and findings need to be verified through the full texts. Third, restricting the language to English and Indonesian may omit studies in other languages, such as studies from China or Spanish-speaking countries. Fourth, mapping design components to theory is an interpretation by the reviewers that may differ from the intentions of the primary study authors. Fifth, the small corpus does not permit moderator analysis or estimation of pooled effects.

## IV. CONCLUSION

This systematic review yields three conclusions. First, regarding RQ1, social constructivism, cognitive load theory, the revised Bloom taxonomy, and self-regulated learning are the frameworks most frequently cited in flipped classroom designs. However, these frameworks are generally cited descriptively and have not been operationalized. Analytically, the four frameworks operate at different

phases. CLT and the Bloom taxonomy are most relevant to designing pre-class materials, social constructivism to in-class activities, and SRL as a prerequisite across phases. Second, regarding RQ2, proximal evidence in university-level electricity and magnetism shows a positive but heterogeneous direction with low strength of evidence. Large effect sizes on examinations were found in non-causal designs, conceptual gains differed by academic level and gender, higher final-examination scores were not accompanied by differences in course grades, and SRL has not been measured as an outcome. Meta-analyses in higher education support small-to-moderate average benefits that depend on design features and active learning. Third, no direct evidence of flipped classrooms in advanced electrodynamics was found. Therefore, its implementation should be positioned as an innovation requiring evaluation, rather than as an established practice. This review proposes a four-phase design that aligns theory with students' mathematical difficulties, together with a research agenda prioritizing controlled studies using validated upper-division instruments.

## V. REFERENCES

Abeysekera, L., & Dawson, P. (2015). 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

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education, 126*, 334–345. https://doi.org/10.1016/j.compedu.2018.07.021

Apriyanah, P., Nyeneng, I. D. P., & Suana, W. (2018). Efektivitas model flipped classroom pada pembelajaran fisika ditinjau dari self efficacy dan penguasaan konsep siswa. *JIPFRI (Jurnal Inovasi Pendidikan Fisika dan Riset Ilmiah), 2*(2), 65–74. https://doi.org/10.30599/jipfri.v2i2.302

Auliza, E., & Widyastuti, E. (2024). Pengaruh model pembelajaran flipped classroom tipe peer instruction flipped berbantuan video pembelajaran terhadap kemampuan pemecahan masalah dan kemandirian belajar. *MES: Journal of Mathematics Education and Science, 9*(2), 173–184. https://doi.org/10.30743/mes.v9i2.8592

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Bredow, C. A., Roehling, P. V., Knorp, A. J., & Sweet, A. M. (2021). To flip or not to flip? A meta-analysis of the efficacy of flipped learning in higher education. *Review of Educational Research, 91*(6), 878–918. https://doi.org/10.3102/00346543211019122

Chasteen, S. V., Pepper, R. E., Caballero, M. D., Pollock, S. J., & Perkins, K. K. (2012). Colorado Upper-Division Electrostatics diagnostic: A conceptual assessment for the junior level. *Physical Review Special Topics—Physics Education Research, 8*(2), Article 020108. https://doi.org/10.1103/PhysRevSTPER.8.020108

Chasteen, S. V., Pollock, S. J., Pepper, R. E., & Perkins, K. K. (2012). Transforming the junior level: Outcomes from instruction and research in E&M. *Physical Review Special Topics—Physics Education Research, 8*(2), Article 020107. https://doi.org/10.1103/PhysRevSTPER.8.020107

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970–977. https://doi.org/10.1119/1.1374249

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom.

*Proceedings of the National Academy of Sciences, 116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Fischer, K., Sullivan, A. M., Cohen, A. P., King, R. W., Cockrill, B. A., & Besche, H. C. (2023). Using cognitive load theory to evaluate and improve preparatory materials and study time for the flipped classroom. *BMC Medical Education, 23*, Article 345. https://doi.org/10.1186/s12909-023-04325-x

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Furqon, M. (2024). Online flipped classroom in electricity and magnetism course supported Google Classroom and Telegram: A case study. *Jurnal Riset dan Inovasi Pembelajaran, 4*(1), 562–578. https://doi.org/10.51574/jrip.v4i1.1354

Gong, J., Cai, S., & Cheng, M. (2024). Exploring the effectiveness of flipped classroom on STEM student achievement: A meta-analysis. *Technology, Knowledge and Learning, 29*, 1129–1150. https://doi.org/10.1007/s10758-023-09700-7

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods, 11*(2), 181–217. https://doi.org/10.1002/jrsm.1378

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285–291. https://doi.org/10.3233/EFI-180221

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), Article ar5. https://doi.org/10.1187/cbe.14-08-0129

Kattel, B., Hutchcraft, W. E., & Norell, E. (2024). Evaluating the effectiveness of flipped learning in an upper-division undergraduate electrical engineering course. *Discover Education, 3*, Article 268. https://doi.org/10.1007/s44217-024-00378-2

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Låg, T., & Sæle, R. G. (2019). Does the flipped classroom improve student learning and satisfaction? A systematic review and meta-analysis. *AERA Open, 5*(3). https://doi.org/10.1177/2332858419870489

Li, R., Lund, A., & Nordsteien, A. (2023). The link between flipped and active learning: A scoping review. *Teaching in Higher Education, 28*, 1993–2027. https://doi.org/10.1080/13562517.2021.1943655

Lo, C. K., & Hew, K. F. (2019). The impact of flipped classrooms on student achievement in engineering education: A meta-analysis of 10 years of research. *Journal of Engineering Education, 108*(4), 523–546. https://doi.org/10.1002/jee.20293

Long, H. A., French, D. P., & Brooks, J. M. (2020). Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. *Research Methods in Medicine & Health Sciences, 1*(1), 31–42. https://doi.org/10.1177/2632084320947559

Maiti, M., & Priyaadharshini, M. (2026). Redefining learning strategies in SRL for student's achievements in flipped classrooms. *Frontiers in Computer Science, 8*, Article 1770049. https://doi.org/10.3389/fcomp.2026.1770049

Meda, T. (2025). Efektivitas Physics Interactive Multimedia pada materi listrik bolak-balik untuk meningkatkan literasi digital menggunakan flipped learning. *Ideguru: Jurnal Karya Ilmiah Guru, 10*(3), 2048–2053. https://doi.org/10.51169/ideguru.v10i3.2176

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pepper, R. E., Chasteen, S. V., Pollock, S. J., & Perkins, K. K. (2012). Observations on student difficulties with mathematics in upper-division electricity and magnetism. *Physical Review Special Topics—Physics Education Research, 8*(1), Article 010111. https://doi.org/10.1103/PhysRevSTPER.8.010111

Serrano, M.-A., Vidaurre, A., Meseguer-Dueñas, J. M., Tort-Ausina, I., Quiles, S., Sabater i Serra, R., García-Sanchez, T., Bernal-Pérez, S., Gámiz-González, M. A., Molina-Mateo, J., Gómez-Tejedor, J. A., & Riera, J. (2023). Active methods in electricity and magnetism courses: Influence of degree, academic level and gender on student performance. *Heliyon, 9*(10), Article e20490. https://doi.org/10.1016/j.heliyon.2023.e20490

Silverajah, V. S. G., Wong, S. L., Govindaraj, A., Khambari, M. N. M., Rahmat, R. W. B. O. K., & Deni, A. R. M. (2022). A systematic review of self-regulated learning in flipped classrooms: Key findings, measurement methods, and potential directions. *IEEE Access, 10*, 20270–20294. https://doi.org/10.1109/ACCESS.2022.3143857

Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review, 30*, Article 100314. https://doi.org/10.1016/j.edurev.2020.100314

Suárez, Á., Martí, A. C., Zuza, K., & Guisasola, J. (2023). Electromagnetic field presented in introductory physics textbooks and consequences for its teaching. *Physical Review Physics Education Research, 19*(2), Article 020113. https://doi.org/10.1103/PhysRevPhysEducRes.19.020113

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*, Article 45. https://doi.org/10.1186/1471-2288-8-45

Tunggyshbay, M., Balta, N., & Admiraal, W. (2023). Flipped classroom strategies and innovative teaching approaches in physics education: A systematic review. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(6), Article em2283. https://doi.org/10.29333/ejmste/13258

van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review, 28*, Article 100281. https://doi.org/10.1016/j.edurev.2019.05.003

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. https://www.hup.harvard.edu/books/9780674576292

Wang, C., Song, L., & Jiang, J. (2026). Enhancing physics education through flipped problem-based learning (PBL): Improving comprehension and problem-solving skills in undergraduate students. *Educational Technology Research and Development, 74*, 1801–1818. https://doi.org/10.1007/s11423-026-10617-y

Wilcox, B. R., Caballero, M. D., Rehn, D. A., & Pollock, S. J. (2013). Analytic framework for students' use of mathematics in upper-division physics. *Physical Review Special Topics—Physics Education Research, 9*(2), Article 020119. https://doi.org/10.1103/PhysRevSTPER.9.020119

Yang, J., & Valcke, M. (2026). Flipped classroom design as a driver of digital transformation and sustainable education in higher education: A systematic review of reviews. *Sustainability, 18*(7), Article 3582. https://doi.org/10.3390/su18073582

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zuza, K., De Cock, M., van Kampen, P., Kelly, T., & Guisasola, J. (2020). Guiding students towards an understanding of the electromotive force concept in electromagnetic phenomena through a teaching-learning sequence. *Physical Review Physics Education Research, 16*(2), Article 020110. https://doi.org/10.1103/PhysRevPhysEducRes.16.020110

Zuza, K., van Kampen, P., De Cock, M., Kelly, T., & Guisasola, J. (2018). Introductory university physics students' understanding of some key characteristics of classical theory of the electromagnetic field. *Physical Review Physics Education Research, 14*(2), Article 020117. https://doi.org/10.1103/PhysRevPhysEducRes.14.020117