

Article

Transforming Learning through Productive Failure: A Multi-Dimensional Framework for Fairness in Education

Jing Liu¹, Junjie Yan¹, Xiaonan Luo¹, Wanni Chen² and Xueliang Li^{3,*}

¹ Department of Educational Technology, College of Education, Wenzhou University, Wenzhou 325035, China

² Wenzhou University Affiliated School, Wenzhou 325035, China

³ Aksu Vocational and Technical College, Wensu Town, Wensu County, Aksu 843100, China

* Correspondence: 1808610551@qq.com; Tel.: +86-15258789292

How To Cite: Liu, J., Yan, J., Luo, X., Chen, W., & Li, X. (2026). Transforming Learning through Productive Failure: A Multi-Dimensional Framework for Fairness in Education. *Journal of Educational Technology and Innovation*, 8(3), 35–51. <https://doi.org/10.61414/bsq6e774>

Received: 8 April 2026

Revised: 12 July 2026

Accepted: 20 August 2026

Published: 30 September 2026

Abstract: High-stakes testing and standardized instruction have long dominated education environments. They often place students with learning difficulties on learning tracks with low-cognitive tasks and low classroom participation. This trend further aggravates educational inequality. Research on fairness in education has gradually shifted its focus from resource allocation to the quality of learning processes. However, most current classroom practices still adopt an “error-prevention instruction” approach oriented toward reducing mistakes and maintaining accuracy rates. This approach limits learners’ deep cognitive development. Against this background, this study takes the theory of productive-failure instruction (Productive Failure, PF) as its core. It re-examines the transition of students with learning difficulties from the perspective of fairness in education. It combines literature analysis, conceptual analysis, and theoretical deduction to systematically explore the internal mechanisms through which productive-failure instruction promotes the cognitive development of students with learning difficulties. This study holds that students with learning difficulties are not a group with fixed ability deficits. They are a group of “developmental differences” that long stay in environments with low-quality cognitive engagement. Compared with traditional error-prevention instruction, productive-failure instruction works through three mechanisms: cognitive conflict, delayed instruction, and knowledge reconstruction. It promotes learners’ deep conceptual understanding and transfer of learning. On this basis, this study further proposes a four-dimensional classroom transformation model of productive failure for fairness in education. The model reshapes the classroom learning ecology. It provides new theoretical explanations and practical pathways for the transition of students with learning difficulties and the practice of fairness in education.

Keywords: productive-failure instruction; transition of students with learning difficulties; fairness in education; classroom reform

1. Introduction

In the process of education modernization, narrowing learning outcome gaps while expanding educational opportunities has become a key topic in research on fairness in education. Education access continues to improve, but ability gaps between learners have not narrowed accordingly. The gaps even widen in education contexts with high-stakes testing systems (Strello et al., 2021) and obvious socioeconomic stratification (Sandsør et al., 2023). Against this background, policy adjustments on education quality and fairness across countries essentially reflect the structural limitations of standardized instruction systems in addressing student diversity.



Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher’s Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

In specific teaching practices, this problem manifests prominently as a “low-cognitive-load orientation” in classroom logic. Teachers oversimplify knowledge content to prevent students from making mistakes. This approach formally ensures the completion of learning tasks, but keeps students with weak foundations stuck in low-level repetitive training for a long time. They can hardly achieve substantive development of their cognitive structure (Li et al., 2021). This phenomenon is not limited to a single country. It widely exists in education systems that emphasize unified progress and standardized evaluation. A fairness in education report by De Bortoli et al. (2023) shows that the coefficient of academic variation in China’s basic education stage is still higher than the average of Nordic countries. This difference not only reflects uneven regional development but also reveals the general dilemma of unified teaching paradigms in addressing cognitive differences.

As an important theoretical breakthrough in cognitive science, productive-failure instruction follows a three-stage mechanism: “cognitive conflict–independent inquiry–collaborative argumentation” (Kapur, 2015). It provides a new path to break through this dilemma. The theory emphasizes that introducing complex problems at the early stage of learning and allowing learners to experience failure can activate higher-level cognitive processing and thus promote knowledge reconstruction and transfer. However, when this student-centered, trial-and-error-based learning approach is introduced into real classrooms, it will generate a series of structural conflicts with existing teaching logic. On the one hand, traditional instruction treats knowledge as a stable, transferable object. Productive-failure instruction understands knowledge as a dynamic cognitive construction process generated in problem situations (Loibl et al., 2017). This difference changes the definition of “knowledge sources” in the classroom. On the other hand, teacher-centered instruction relies on clear guidance and standard answers to ensure the certainty of learning outcomes. Failure-based learning emphasizes exploration processes and multiple solution paths. Its learning value often emerges gradually in delayed feedback (Schwartz et al., 2011). This requirement clearly mismatches the demand for immediate measurability in high-stakes testing systems. In this context, failure can be both an opportunity for cognitive development and a perceived learning risk. For this reason, people systematically avoid failure in teaching practice. Existing studies have revealed the constraints on the transformation of productive failure theory from the perspectives of emotional culture (Chen et al., 2024), classroom power structure (Lee & Kim, 2019), and examination system constraints (Xu & Lu, 2022). But on the whole, there is still a lack of an analytical framework that integrates the cognitive mechanisms of students with learning difficulties and the goal of fairness in education.

The rapid development of artificial intelligence technology has further changed the ways of knowledge acquisition and learning support structures in classrooms. On one hand, artificial intelligence can provide personalized feedback and dynamic scaffolding based on learners’ behavioral data. It offers more precise cognitive support for students with learning difficulties (Zawacki-Richter et al., 2019). On the other hand, if people only use artificial intelligence for quick error correction and answer generation, it may reduce the space for learners to experience cognitive conflict and generate diverse solutions (Khalil & Er, 2023). This will weaken the deep cognitive processing process on which productive failure depends. Therefore, against the background of deep integration of artificial intelligence into classrooms, balancing the relationship between “support and exploration” and between “efficiency and cognitive generation” has become an important premise for re-examining productive-failure instruction. It is necessary to elevate “productive failure” from a single teaching strategy to a mechanism of fairness in education for investigation.

This study aims to reveal the internal logic of the transition of students with learning difficulties driven by productive-failure instruction, and provide a theoretical basis and a practical framework for classroom reform. Its significance is reflected in three aspects:

- (1) Theoretical level: It breaks through the limitation of traditional research on students with learning difficulties that focuses on compensatory support. It constructs a three-in-one theoretical model of “cognitive reconstruction–social negotiation–fair development”, and expands the practical connotation of fairness in education.
- (2) Strategic level: Based on productive failure theory, formative assessment, and research on artificial intelligence-supported learning, it constructs a “four-dimensional classroom transformation model” covering space, time, technology, and evaluation. It addresses the long-term problem of low cognitive challenge and low classroom participation for students with learning difficulties, and provides operable classroom implementation paths for teachers to carry out deep learning-oriented instruction.
- (3) Practical level: Focusing on teacher education, school evaluation, and classroom implementation mechanisms, it proposes practical pathways such as cultivating implementation capacity for fairness in education, building a value-added fairness assessment system, and optimizing productive-failure instruction. It promotes productive-failure instruction to shift from a partial classroom strategy to a systematic education reform practice.

2. Literature Review

2.1. Causes of Learning Difficulties among Students

The issue of students with learning difficulties is a long-standing topic in education. Its causes show multi-dimensional interactive characteristics. Most existing studies agree that learning difficulties result from the combined effect of individual internal factors and external environmental conditions (Yan et al., 2025). With the continuous penetration of artificial intelligence technology into education, this mechanism is being further reshaped.

From the perspective of internal factors, the lack of learning goals, insufficient self-regulation ability, and weak cognitive structure form the core mechanism of learning difficulties. Vague learning goals weaken the directionality and sustainability of learning behaviors. Goal setting and reflection can significantly improve academic performance (Morisano et al., 2010). Insufficient self-regulation ability manifests as defects in planning, monitoring, and metacognitive regulation. It has a direct impact on the use of learning strategies and learning effectiveness (Panadero et al., 2017). Weak cognitive structure is reflected in insufficient prior knowledge and solidified misconceptions. It limits the integration and transfer of new knowledge (Simonsmeier et al., 2022). In digital learning environments, these problems are often further amplified. For example, online learning requires learners to have stronger self-regulation ability. Otherwise, they easily fall into inefficient participation under information overload and fragmented learning (Broadbent & Poon, 2015). Meanwhile, although artificial intelligence-driven personalized recommendations can make up for learners' knowledge gaps, their adaptation mechanism based on existing performance may solidify low-level learning paths and limit the improvement of students' cognitive level (Pelánek et al., 2024).

From the perspective of external factors, insufficient family support, a single school teaching mode, and uneven distribution of social resources constitute important situational conditions for learning difficulties. Lack of family supervision affects students' learning engagement by reducing learning support and educational expectations (Mao et al., 2020). The teaching mode in schools, characterized by unified progress and mechanical training, keeps some students on low-gain learning tracks for a long time by limiting cognitive challenges and differentiated support (Furtak et al., 2012). Differences in socioeconomic status and educational resources affect learning opportunities at a more macro level and may lead to the continuous widening of academic gaps (Nieuwenhuis & Hooimeijer, 2016). In addition, against the background of technology intervention, the digital divide has gradually shifted from “access differences” to “use quality differences”. Disadvantaged student groups still remain in an unfavorable position in obtaining and effectively using high-quality learning resources (Hadjar & Naumann, 2025).

2.2. Theoretical Progress of Productive-Failure Instruction

Manu Kapur proposed the productive failure (PF) theory in 2006. Its core assumption is that “experiencing failure in problem solving without scaffolding support can promote deep transfer of subsequent conceptual learning”. The development of this theory can be divided into three stages: Mechanism exploration stage (2006–2012): Through mathematics problem-solving experiments, researchers confirmed that the PF group performed significantly better than the direct instruction group in conceptual understanding and transfer ability. Key design principles include “delayed instruction”, “diverse solution generation,” and “cognitive conflict stimulation” (Kapur, 2014). Interdisciplinary verification stage (2013–2020): Research expanded to STEM education, language learning, and maker education (Nachtigall et al., 2020). Refined design stage (2021–present): Focusing on the boundary conditions of PF, scholars put forward the concept of “micro-failure” and a task difficulty calibration model (Sinha & Kapur, 2021).

Although PF shows the potential to promote deep learning in many studies, its applicability still faces important controversies. Cognitive load theory points out that without necessary scaffolding, complex problems may exceed beginners' cognitive capacity and reduce learning efficiency (Kirschner et al., 2006). For example, in learning environments with minimal guidance, beginners perform significantly worse than learners receiving direct instruction in complex problem-solving tasks. The reason is that cognitive resources are overconsumed in ineffective search processes (Ashman et al., 2020). The constructivist perspective argues that failure experience is an important condition for knowledge reconstruction. Moderate cognitive conflict helps activate deep processing (Kapur, 2008). A study by Loibl et al. (2020) shows that moderate cognitive conflict prompts learners to re-examine their original cognitive framework. By identifying structural differences and conceptual gaps between solutions, it enhances learners' understanding of deep features of knowledge and transfer ability. This disagreement essentially reflects the balance between “intensity of guidance” and “space for exploration”.

Further studies show that the effect of PF highly depends on learner characteristics and task design conditions (Nachtigall et al., 2020; Sinha & Kapur, 2021). In recent years, research has gradually shifted from “whether it works” to “under what conditions it works”. For example, learners at different age stages have different perceptions

and coping styles toward failure. Younger students are more likely to attribute failure to insufficient ability and develop avoidance behaviors. Older students are more likely to regard failure as an opportunity for strategy adjustment (Metcalf, 2017). In addition, Loibl et al. (2017) found through experimental research that subsequent instruction can effectively promote conceptual understanding only when learners generate diverse solutions during the failure stage. Otherwise, failure may only bring frustration rather than learning benefits. Today, artificial intelligence technology can provide result-oriented prompts or even directly generate solutions based on learners' immediate performance. Although this can improve task completion efficiency to a certain extent, it may reduce learners' independent exploration time during the failure stage. It weakens their comparison of different solution paths and deep cognitive processing. Conversely, when artificial intelligence is used to record different problem-solving paths, delay key feedback, and support peer comparison of solutions, it helps preserve cognitive conflict in the failure process and thus promotes subsequent conceptual construction (Ouyang & Zhang, 2024).

2.3. Realizing Productive-Failure Classrooms for Fairness in Education

Realizing productive failure in classroom situations depends not only on the teaching strategy itself, but also on the redistribution of learning opportunities and cognitive engagement oriented by fairness in education. Fairness in education is not an external condition for PF. It is the internal mechanism for its realization. Only by expanding learning opportunities at the starting point, improving cognitive engagement in the process, and ensuring development space in evaluation can classrooms truly provide sustainable conditions for “productive failure” to occur.

At the learning starting point level, a fairness orientation requires lowering the threshold for learners to enter complex tasks through the redistribution of resources and support (Tan, 2024). For example, China's “three types of classrooms” (delivery classrooms, famous teacher classrooms, and famous school online classrooms) expand opportunities for students in disadvantaged areas to access high-quality problem situations and diverse solutions through online sharing of high-quality resources (Sun et al., 2024). This provides basic conditions for the “multiple solution generation” on which PF depends. The key to this mechanism does not lie in resource transmission itself. It lies in the opportunity space it provides for students to access complex problems and try multi-path solutions.

At the learning process level, the effective implementation of PF depends on the reconstruction of the learning participation structure. Studies show that with online collaboration platforms and learning management systems, teachers can organize inter-school collaborative learning, share generative tasks, and support multi-solution comparison. This helps enhance the depth of students' cognitive engagement during the failure stage (Meyer et al., 2023). Meanwhile, technology-supported formative assessment can provide immediate feedback for students. It helps them turn failure experience into reflective learning resources, so as to prevent failure from remaining as “ineffective frustration”.

At the evaluation and outcome level, fairness in education requires weakening the constraints of a single summative assessment on learning paths and providing institutional space for exploratory learning (Tai et al., 2018). Relevant studies show that excessive reliance on standardized testing and stratified teaching tends to limit students' opportunities to participate in high-challenge tasks and strengthen the differentiation of learning paths (Terrin & Triventi, 2023). In contrast, an evaluation system combining process evaluation and dynamic feedback is more conducive to supporting students' continuous development in the cycle of attempt–failure–revision.

3. Research Methodology

This study belongs to conceptual/theoretical research. It aims to re-examine the internal logic of the transition of students with learning difficulties driven by productive-failure instruction from the perspective of fairness in education, and construct a practical framework for classroom reform on this basis. Compared with empirical research centered on variable testing, conceptual research emphasizes forming explanatory mechanisms and practical models for complex educational phenomena through systematic literature integration, theoretical deduction, and framework construction (Jaakkola, 2020).

On the basis of literature analysis, the theoretical analysis part mainly adopts a combination of conceptual analysis and theoretical deduction. Relevant studies show that theoretical deduction can help educational research shift from fragmented phenomenon description to mechanistic explanation, thus enhancing the theoretical explanatory power and conceptual innovation of research (Allen-Platt et al., 2021). Therefore, this study does not regard “students with learning difficulties” as a group with fixed ability deficits. It redefines the concept of “students with learning difficulties” by deducing the paradigm of fairness in education, and puts forward the explanatory framework of “developmental differences”. Meanwhile, based on productive failure theory, cognitive conflict theory, formative assessment theory, and fairness in education theory, this study conducts theoretical deduction on the internal logic among “cognitive conflict–knowledge reconstruction–developmental fairness”.

The focus of this part is not to verify variable relationships. It explains why traditional “error-prevention fairness” leads to cognitive impoverishment of students with learning difficulties, and why productive-failure instruction can realize procedural fairness at the classroom level through cognitive activation, delayed instruction, and social negotiation mechanisms.

In the strategic analysis part, this paper mainly adopts the inductive approach. Inductive framework construction in educational research helps transform scattered teaching strategies into systematic, practical models, thus improving the classroom explanatory power and application value of theory (Mu et al., 2022). Therefore, on the basis of the aforementioned theoretical analysis, combined with domestic and foreign research results on classroom cognitive activation, spatial organization, artificial intelligence-supported learning, and formative assessment, this study summarizes and integrates the key implementation elements in productive-failure classrooms. It further proposes the “four-dimensional classroom transformation model”, which systematically reconstructs productive-failure classrooms from four dimensions: space, time, technology, and evaluation. Among them, the spatial dimension emphasizes the reconstruction of classroom power relations. The temporal dimension emphasizes stepped failure teaching design. The technical dimension emphasizes precise support for cognitive exploration by generative artificial intelligence. The evaluation dimension emphasizes the establishment of a value-added fairness assessment system. This model not only responds to the long-standing problem in PF research of “effective in theory but difficult to implement in classrooms”, but also expands the practical path of fairness in education from resource compensation to cognitive empowerment.

In the practical analysis part, this paper adopts a combination of “theoretical deduction–practical derivation” to systematically analyze the implementation path of productive-failure instruction in educational practice. Some studies point out that the important value of educational theory research lies not only in explaining educational phenomena, but also in providing operable institutional paths and practical solutions for education reform (Hanushek, 2003). Therefore, based on the aforementioned theoretical framework and classroom transformation model, this study further carries out practical derivation from the aspects of teacher education curriculum reform, school quality evaluation mechanism reconstruction, and education policy support. It elevates productive-failure instruction from a partial classroom strategy to a practical path of fairness in education with institutional value.

In summary, through the research path combining literature analysis, theoretical deduction, and induction, this study gradually forms a three-in-one research framework of “theoretical mechanism–classroom model–practical path”. It is expected to provide explanatory and practical theoretical support for research on the transition of students with learning difficulties in the context of fairness in education.

4. Theoretical Analysis of the Transition of Students with Learning Difficulties Driven by Productive-Failure Instruction

4.1. Strategic Guidance: The Policy Coordination of the Transition of Students with Learning Difficulties in the Context of Building a Leading Country in Education

4.1.1. Decoding the View of Fairness in Education

The contemporary framework of fairness in education takes “education for all” as its core orientation. People no longer narrowly equate fairness with the allocation of enrollment opportunities and educational resources. Instead, fairness focuses simultaneously on the process quality and outcome quality of teaching. For students with learning difficulties, the view of fairness in education should evolve from coverage-oriented opportunity fairness to “quality justice”. This quality justice centers on the three-in-one process quality of classroom interaction, cognitive activation, and formative assessment. Evidence-based designs such as productive failure serve as actionable approaches to translate fairness into effective learning gains (Sinha & Kapur, 2021).

First, high-quality classroom interaction (including teacher-student emotional support, expectation communication, and constructive interaction) is the basic channel to promote the participation of disadvantaged students, maintain learning motivation, and ensure their access to teaching support. A large body of meta-analytic evidence shows that emotional support and positive teacher-student relationships significantly improve student engagement and academic achievement. The impact is particularly important for low-achieving students (Roorda et al., 2011).

Second, cognitive activation refers to deliberately stimulating students’ cognitive processing through task design (for example, triggering conceptual conflicts through complex problems, promoting generative thinking, and multi-solution exploration). It is the key mechanism to realize deep learning and knowledge transfer (Lazonder & Harmsen, 2016). Productive failure is built on the cognitive activation logic of “first letting students experience cognitive conflict in problem solving in generative tasks, then giving integrated guidance”. Multiple studies show that it has a robust positive effect on students’ conceptual mastery and transfer (Kapur, 2008; Sinha & Kapur, 2021).

Finally, formative assessment and high-quality feedback serve as a mechanistic tool that turns classroom interaction and cognitive activation into sustained learning gains. A large number of reviews and meta-analyses show that explanatory, actionable feedback and continuous formative assessment can significantly improve learning outcomes, and help narrow gaps between students (Black & Wiliam, 1998; Hattie & Timperley, 2007).

4.1.2. Semantic Reconstruction of the Connotation of Students with Learning Difficulties

The terminological shift from “underachievers” to “developmental differences” reflects three paradigm shifts in the concept of fairness in education. The “compensatory fairness” in the 1980s emphasized resource inclination (Borman et al., 2001). Practices in this stage focused on “who receives extra resources” rather than differentiated support in classroom processes. The “procedural fairness” in the early 2000s focused on teaching intervention. It emphasized the importance of classroom interaction quality, teacher-student relationships, formative assessment, and teaching appropriateness in promoting the learning of disadvantaged students (Florian & Black-Hawkins, 2011). This paradigm shifted “compensation” from transfusion-style aid at the institutional level to scaffolded teaching and dynamic monitoring at the classroom level. It significantly reduced the stigmatization effect of “moralized interpretation” on students with learning difficulties. The “developmental fairness” in the new era focuses on activating individual potential. It emphasizes activating the learning potential of every student, rather than only focusing on the equal distribution of resources or processes (Gracia-Calandín & Tamarit-López, 2021). This semantic reconstruction breaks the traditional prejudice that regards academic difficulties as moral defects. It reinterprets the formation mechanism of individual differences among students through the theory of multiple intelligences.

In the process of education modernization, one of the root causes for the systematic neglect of students with learning difficulties is the quantitative and utilitarian tilt of the education evaluation system. It squeezes the essence of education by “measurable indicators” (Polesel et al., 2014). When standardized assessment and test-based accountability systems become the main driving force of teaching, teachers and schools tend to prioritize time and resources into assessed subjects and quantifiable indicators. As a result, learning activities involving practical innovation, art, or cross-domain skills are marginalized. Students with such abilities are judged as “substandard” or “problem students” because they fall outside the evaluation dimensions. This structural dilemma requires education policies to build a differentiated value recognition mechanism, rather than staying at the mechanical allocation of material resources (Sirin, 2005).

4.2. Paradigm Breakthrough: Theoretical Innovation of Productive-Failure Instruction

4.2.1. Localized Integration of Western “Productive Failure” Theory

Kapur (2015) proposed the productive failure theory. This theory shares deep cultural connections with traditional Chinese education. The Book of Learning puts forward a classic teaching principle: guide students instead of pulling them forward, encourage them instead of holding them back. This principle aligns with the “disequilibrium and reconstruction” mechanism in cognitive conflict theory, even though they come from different times and cultural backgrounds. This cultural connection shows itself in three aspects. First, the Confucian idea of “gaining wisdom through hardships and acting with diligence” lays a cultural foundation for the value of frustration education. It gives positive meaning to failure experiences. Second, the dialectical thinking of Taoism, “reversal is the movement of the Dao,” matches the cognitive value-added logic behind failure-focused instructional design. Third, the Chan Buddhism concept of “sudden enlightenment” emphasizes individual differences in the threshold of cognitive breakthroughs. It provides traditional wisdom support for differentiated failure-based teaching design.

To realize the localized transformation of productive-failure instruction theory, educators must address three cultural challenges. The first is how to turn the sense of shame around failure in collectivist traditions into learning motivation (Mesquita & Karasawa, 2004). The second is how to reshape teacher-student interactions in generative learning within the framework of teacher authority (Tweed & Lehman, 2002). The third is how to reconstruct procedural evaluation mechanisms in an exam-oriented culture (Dee & Jacob, 2011).

First, the socialization of failure shame in collectivist cultures makes students avoid risks in public trial-and-error situations. They become less willing to participate or show their mistakes. This weakens the cognitive exploration motivation required for generative tasks (Fang et al., 2023). For productive failure learning to take root effectively in Chinese classrooms, education authorities need to reduce students’ sense of failure shame through emotionally and culturally sensitive interventions.

Second, teacher authority and hierarchical teacher–student relations, rooted in Confucian traditions, create cultural and practical barriers to students’ generative exploration (Sum & Kwon, 2020). In teacher-centered classrooms, teachers find it hard to shift classroom control to student-led exploration spaces. Teachers also hold

professional expectations of maintaining classroom order and avoiding “loss of face”. These expectations limit their willingness to adopt teaching strategies that allow public failure. Therefore, the localization of productive failure is not just about introducing new teaching tasks. It also requires the reshaping of teacher roles and the advancement of teacher professional development.

Finally, the widespread high-stakes testing culture fixes teaching goals and resource allocation on measurable indicators. It squeezes the space for procedural learning and formative assessment. Schools and teachers thus prefer to “teach testable question types” instead of supporting fault-tolerant cognitive exploration (Dee & Jacob, 2011). Research by Polesel et al. (2014) shows that high-pressure testing leads to curriculum narrowing. It tilts teaching toward testable content and changes the priorities of school resources and teaching practices. This creates institutional barriers to fairness in education and teaching innovation. Therefore, we need to reform evaluation and examination mechanisms. This will leave institutional space for formative assessment and bearable failure.

4.2.2. Value Shift of Productive-Failure Instruction

Traditional “error-prevention fairness” provides superficial equality by reducing cognitive difficulty. But it easily leads to “cognitive impoverishment” of students with learning difficulties. Productive-failure instruction can promote three shifts in fairness in education: The focus of teaching shifts from remediating learning outcomes to optimizing cognitive processes (Sinha & Kapur, 2021). Support methods shift from external resource supply to stimulating students’ self-regulation and intrinsic motivation (Zimmerman, 2002). Training objectives shift from guaranteeing the right to acquire knowledge to building transferable learning abilities (Lazonder & Harmsen, 2016). The theoretical basis for this shift lies in the following: moderate frustration experience and delayed instruction can activate students’ reflective learning, promote their knowledge reorganization and transfer, and thus turn “allowing failure” into long-term learning gains (Hattie & Timperley, 2007).

The fairness value of productive-failure instruction design is reflected in its dynamic compensation mechanism: First, by creating cognitive conflict within the zone of proximal development, productive-failure instruction helps students with learning difficulties expose their cognitive misunderstandings in problem solving (Nachtigall et al., 2020). Second, by building a dialogue mechanism of a learning community, productive-failure instruction breaks the discourse hegemony of top students and reconstructs the power structure of knowledge production in the classroom (Mitra & Serriere, 2012). Third, productive-failure instruction guides individuals to achieve cognitive self-regulation through tools such as reflection journals and realizes substantive fairness in differentiated development (Zhang et al., 2022). Productive-failure instruction elevates fairness in education from resource redistribution to the dimension of “thinking rights”. It provides an operable mechanism for realizing procedural justice in classrooms.

5. Strategic Analysis of the Transition of Students with Learning Difficulties Driven by Productive-Failure Instruction

Based on the previous literature review and theoretical analysis, we can find that the formation of learning difficulties is not caused by a single ability defect. It results from the interaction of multiple factors such as cognitive structure, self-regulation, classroom participation, resource acquisition, and evaluation mechanism. Meanwhile, research on productive failure also shows that its learning gains do not come from “failure” itself. They depend on whether learners can complete knowledge reconstruction in moderate cognitive conflict, continuous exploration, and subsequent integrated support (Kapur, 2014). However, the implementation of PF in real classrooms is still restricted by traditional teaching structures: Fixed classroom space easily solidifies students’ participation identities (Vaughn, 2021). Unified teaching rhythm compresses exploration time (Kozanitis & Nenciovici, 2023). Standardized evaluation inhibits trial-and-error behavior (Simon et al., 2024). Improper intervention of technical tools may prematurely terminate the cognitive exploration process (Bastani et al., 2024).

Therefore, regarding PF only as a teaching method can hardly truly respond to the deep-seated needs of the transition of students with learning difficulties and fairness in education. Against this background, based on the aforementioned research on fairness in education, cognitive activation, formative assessment, and artificial intelligence-supported learning, this study further constructs the “four-dimensional classroom transformation model”. It systematically reconstructs productive-failure classrooms from four dimensions: space, time, technology, and evaluation, as shown in Figure 1. The four dimensions are not isolated from each other. They together form the classroom implementation framework for the transition of students with learning difficulties driven by productive-failure instruction. They promote fairness in education to further shift from resource compensation to cognitive empowerment and procedural justice.

5.1. Spatial Dimension: Reconstructing Classroom Space to Eliminate the “Back Row Phenomenon”

Classroom space is a visual projection of the educational power structure. Seating arrangement, sight distribution, and teachers’ spatial attention jointly determine which students can gain the right to speak and cognitive support (Smidekova et al., 2020). A large number of reviews and empirical studies show that teachers tend to place students with distracted attention or low behavioral compliance in the back row of the classroom. This causes this area to gradually evolve into a marginal zone of teaching participation (Dong et al., 2021). This spatial differentiation not only intensifies the identity solidification effect of students but also forms structural obstacles to cognitive development by limiting interaction opportunities (Wannarka & Ruhl, 2008). Building a dynamic learning community requires breaking through traditional spatial barriers and realizing a deep reconstruction of teaching relations. Specific measures can be implemented from the following three aspects:

Spatial deconstruction and reorganization: Adopt U-shaped, round-table, or project-island layouts to eliminate the hierarchical difference between front and back rows. For example, the inter-school learning community practice of Pudong Institute of Educational Development shows that through random grouping and role rotation, students transform from “passive recipients” to “knowledge co-constructors” (Nehyba et al., 2023).

Balanced allocation of cognitive resources: Build virtual learning spaces with platforms such as Moodle, and turn marginal students in physical space into “opinion contribution nodes” in online discussions (Ho et al., 2023). This forms a virtual-real integrated cognitive network. Students can increase their speaking opportunities and peer feedback without being restricted by seating positions.

Reconstruction of power relations: Teachers should transform from “space dominators” into collaborative facilitators. They can break the link between seating position and discourse power by patrolling the classroom and using differentiated questioning strategies. For instance, they can prioritize students in the back row to share their problem-solving ideas.

5.2. Temporal Dimension: Application of the Stepwise Productive-Failure Instruction Design Model

Based on the productive failure theory of Kapur (2008), building a stepped teaching design model requires precise control of the timing of cognitive conflict delivery and support intensity. Teaching design can be carried out in three stages:

First, in the pre-failure stage, teachers should create cognitive conflict situations and deliver cognitive conflicts precisely. By deliberately creating cognitive conflicts and complex problems within learners’ zone of proximal development, productive-failure instruction helps students with learning difficulties expose and become aware of their cognitive misunderstandings. This provides diagnostic information for subsequent guiding instruction and promotes deep knowledge reorganization (Henningsen & Stein, 1997). For example, in mathematics concept teaching, teachers can create students’ cognitive tension through preconception exposure strategies. In the teaching of “fraction size comparison”, teachers can first present an intuitive diagram of “ $1/3 > 1/4$ ”, and guide students to debate the contradiction between empirical cognition and mathematical principles of “the larger the denominator, the smaller the fraction when numerators are the same”. This design forces students with learning difficulties to face knowledge gaps directly and lays the foundation for subsequent conceptual reconstruction.

Second, in the controlled failure stage, teachers should provide externalized thinking and progressive scaffolding while maintaining the complexity of the problem itself (van de Pol et al., 2010). For example, set up stepped task groups in mathematics problem solving. The first-level task retains the complexity of the original problem, but provides 3D modeling tools to assist spatial imagination. The second-level task requires students to analyze wrong solutions against standard answers through data comparison scaffolding.

Finally, in the knowledge reconstruction stage, social negotiation and formative feedback should be completed to turn individual failure into shared cognitive resources (Hattie & Timperley, 2007). For example, hold discussion meetings on failed physics experiment schemes. Students can reconstruct their own cognitive structure by learning from others’ experiences. This process turns individual failure experience into group cognitive resources, and forms a virtuous cycle of “failure–attribution–iteration”.

5.3. Technical Dimension: GAI Empowering Precise Support for Productive Failure

The development of generative artificial intelligence (GAI) has gradually turned productive-failure instruction from a partial teaching strategy that relies on teacher experience into an intelligent learning support system with dynamic adjustment and continuous tracking functions. Traditional digital tools focus on “improving accuracy”. PF-oriented GAI focuses more on preserving learners’ exploration process during the failure stage. It creates personalized failure situations, provides cognitive scaffolding and adaptive assessment, and turns failure into a real resource for cognitive reconstruction.

For the personalized creation of failure situations, GAI can identify learners' preconception misunderstandings based on their historical homework, error records, and classroom behavior data. It then dynamically generates cognitive conflict tasks. For example, in physics learning, the AI system can generate differentiated experimental situations for students who confuse sliding friction and rolling friction (Kieser et al., 2024). In mathematics learning, it can automatically generate counterexample tasks for cognitive biases in fraction size rules (Ross & Andreas, 2024). This dynamic task generation based on learning weaknesses allows students to experience "bearable failure" in problem situations highly related to their own cognitive structure. It increases the possibility of subsequent knowledge reconstruction.

For the cognitive scaffolding link in the failure stage, GAI can maintain students' exploration depth through Socratic questioning, instead of providing answers directly. When students get stuck in complex problems, the AI system does not give standard solutions immediately. It guides them to re-examine their existing ideas through continuous enlightening questions. Lee et al. (2024) point out that a generative AI tutor based on Socratic questioning can guide learners to identify conceptual conflicts and revise reasoning paths independently through continuous reflective questions. This maintains cognitive engagement in complex problem solving. Compared with direct error correction systems, AI support based on heuristic questioning better promotes the development of learning transfer and complex problem-solving abilities (Shabrina et al., 2024; Zhang et al., 2025). This interaction mode of "delayed answers–strengthened reflection" helps preserve the cognitive conflict process on which PF depends. It also helps learners develop metacognitive monitoring ability.

For the adaptive assessment of failure experience, GAI can continuously track the learning process through multimodal learning analytics. It breaks the limitation of traditional evaluation that only focuses on right or wrong answers. For example, in language classrooms, the system can comprehensively analyze students' text expression, voice discussions, and revision trajectories. It identifies weak links in their argumentation structure and language organization (Schneider, 2024). In programming classrooms, the system can judge students' problem-solving strategies and persistence by recording code modification processes, debugging frequencies, and error correction paths (Xu et al., 2023). Some studies have combined facial expression recognition, eye movement trajectories, and behavior log data to conduct a comprehensive analysis of learners' emotional fluctuations, cognitive load, and learning engagement. This helps more accurately distinguish two different learning states: "high engagement but temporary failure" and "low participation and low understanding" (Zawacki-Richter et al., 2019). This feedback mechanism shifts from "error identification" to "growth support". It turns failure experience into a real cognitive resource for the development of learning ability. On this basis, the AI system can generate development-oriented feedback reports instead of simple scores. For example, in writing instruction, when the system finds a lack of logical connection between students' arguments, it not only reminds them of "chaotic structure". It further generates a visual framework of "claim–evidence–connection", and recommends corresponding revision paths and sample texts (Ranalli et al., 2017).

5.4. Evaluation Dimension: Constructing a Value-Added Fairness Assessment System

Value-added assessment takes students' original learning foundation as a reference. It focuses on measuring their growth rate in a specific learning cycle, rather than making horizontal comparisons only based on absolute scores under unified standards. For example, one student improves from 60 to 75 points, and another improves from 90 to 95 points. In traditional evaluation, the latter may still be regarded as "better". But value-added assessment pays more attention to the significant improvement of the former in learning ability and cognitive structure (Morris et al., 2021). For productive-failure instruction, this evaluation method can more truly reflect students' development process in the "attempt–failure–revision" cycle. It prevents students with learning difficulties from being fixed in evaluation due to their long-term low starting point. Therefore, building a "three-in-one" value-added fairness assessment system within the framework of productive-failure instruction requires a fundamental breakthrough from the traditional result-oriented evaluation model. It should turn into an evaluation index system that covers three dimensions of opportunity, process, and outcome, and follows a value-added orientation.

Designing opportunity fairness indicators: Traditional practices often only focus on the average distribution of physical resources, such as equipment, textbooks, and teacher allocation. Value-added assessment breaks through this narrow focus on static resource allocation. It turns to evaluate the actual ability of students with learning difficulties to obtain and use "cognitive opportunities" (OECD, 2012). For example, traditional evaluation may only record whether students participate in classroom discussions. The value-added opportunity fairness indicator evaluates the frequency and quality of students with learning difficulties participating in high-order thinking activities through multi-dimensional dynamic monitoring. Such activities include putting forward hypotheses, participating in reasoning, and conducting self-explanation.

Designing process fairness indicators: Value-added assessment breaks through the traditional limitation of only focusing on unified teaching processes. It turns to focus on the real-time cognitive and emotional state changes of students with learning difficulties during the learning process (Roorda et al., 2011). Previous evaluations tended to focus on “whether formative assessment is carried out”. Now we should consider “to what extent formative assessment promotes the improvement of metacognition and strategy use of students with learning difficulties”. That is, we turn process quality into value-added indicators of students’ learning ability.

Designing outcome fairness indicators: Value-added assessment completely abandons the traditional practice of only taking absolute scores and unified standards as the measurement basis. It adopts the value-added model and focuses on the progress range and growth trajectory of students with learning difficulties relative to their own starting point. Traditional evaluation usually compares students’ achievements through standardized test scores. The value-added outcome fairness indicator tracks the growth of students with learning difficulties in multiple dimensions, such as conceptual understanding, strategy application, and academic resilience. For example, the research of Pounder (2018) emphasizes that outcome evaluation of productive-failure instruction should calculate the ability gain value of students with learning difficulties after experiencing a failure learning cycle. This can objectively reflect their real progress. This evaluation not only reduces prejudice based on horizontal comparison, but also identifies the subtle progress of students with learning difficulties that is not easy to detect. It provides fairer developmental feedback.



Figure 1. Four-dimensional Classroom Transformation Model.

6. Practical Analysis of the Transition of Students with Learning Difficulties Driven by Productive-Failure Instruction

6.1. Value Transmission from the Micro to Macro Level

At the micro level, meaningful failure situations can reconstruct the cognitive schemas of students with learning difficulties and cultivate their resilience and self-efficacy. Social cognitive theory points out that individuals form beliefs about their own abilities through mastery experience and reflective interpretation of failure. They then show higher persistence and coping strategies when facing challenges (Bandura, 1977). Evidence from developmental psychology further shows that resilience is not an innate talent. People can shape it through systematic learning experiences and supportive environments (Masten, 2001). In educational practice, classroom design based on productive failure deliberately exposes students to controllable cognitive conflicts and then provides integrated guidance. Multiple studies have confirmed that this design can improve conceptual understanding and transfer ability. It also enhances learners’ ability to cope with complex tasks (Kapur, 2008; Sinha & Kapur, 2021). Therefore, the core of the transformation path from “frustration tolerance” to “innovative

breakthrough” lies in designing failure as a learnable and reflective experience. This can cultivate cognitive resilience for unknown problems.

At the macro level, international assessments and policy analyses have begun to include “academic resilience/proportion of resilient students” into the indicator system for measuring fairness in education and system compensation effectiveness (Jin et al., 2022). PISA-related studies take the proportion of students who “achieve high achievements despite disadvantaged socioeconomic backgrounds” as an operational indicator of academic resilience. They believe this ratio directly reflects the institutional capacity of a country or region in reshaping the development trajectory of disadvantaged students (Agasisti et al., 2018). Accordingly, China can include classroom-level transition rate of students with learning difficulties, classroom participation balance, and higher-order thinking development index into a micro-meso-macro coupled monitoring framework. It can realize a policy closed loop through the three-level linkage of “school effectiveness–regional balance–national strategy”. Elevating individual gains from classroom interventions (such as PF) to systematic evaluation through value-added and long-term tracking indicators can not only confirm the immediate effect of classroom reform, but also provide operable fairness indicators and evaluation paths for realizing fairness in education (McCaffrey et al., 2004).

6.2. Adding an “Implementation Capacity for Fairness in Education” Module to Teacher Education Curricula

The current teacher training system has an imbalance in ability structure. It overemphasizes teaching skill training, but ignores the practical wisdom of fairness in education. We suggest adding a compulsory module of “implementation capacity for fairness in education” in pre-service teacher training and in-service teacher training. The module focuses on three core abilities: First, differential diagnosis ability. We train teachers to accurately identify the cognitive styles and learning obstacles of students with learning difficulties (van de Pol et al., 2010). Second, culturally responsive ability. We guide teachers to turn the life experiences of students with learning difficulties into teaching resources. For example, they can reconstruct mathematical problem situations through folk culture cases (Villegas & Lucas, 2002). Third, failure of design ability. We train teachers to scientifically plan the threshold of cognitive conflict and design moderate challenges on the premise of protecting the learning motivation of students with learning difficulties (Desimone, 2009). This kind of ability training needs to break through the traditional lecture mode. It adopts a practical path of “clinical diagnosis + situation simulation”. For example, we can develop a case library of urban-rural differences and require teachers to design layered intervention schemes for different types of learning difficulties. This can reshape teachers’ action power for fairness in education.

6.3. Establishing a Linkage Mechanism Between Transition Effectiveness of Students with Learning Difficulties and School Quality Evaluation

The invisible exclusion of students with learning difficulties in the current education evaluation system makes schools more inclined to “cultivate top students” rather than “improve bottom students”. We need to reconstruct school quality evaluation standards and build a “three-dimensional dynamic evaluation model.” The basic dimension examines the academic qualification rate of students with learning difficulties. It focuses on guaranteeing the bottom-line fairness of education (Au, 2007). The development dimension tracks the progress of students with learning difficulties. It uses value-added evaluation to calculate their relative progress value (Koedel & Rockoff, 2015). The innovation dimension introduces resilience assessment. It observes the persistence and strategic flexibility of students with learning difficulties in coping with challenges through situational tasks. Evaluation results should be deeply linked to the allocation of educational resources. For example, schools with a continuously improved transition rate of students with learning difficulties can get targeted support in teacher staffing and teaching research funds. Meanwhile, “pseudo high-quality schools” that simply eliminate students with learning difficulties will receive yellow card warnings (Dee & Jacob, 2011). At the same time, we should also pay attention to the negative effects of this linkage. For example, schools may over-concentrate teaching resources on examination subjects. This will lead to curriculum narrowing and the tendency of exam-oriented teaching (Dee et al., 2013). Some schools may also maintain performance indicators by reducing the participation of low-achieving students or implicit streaming. This will further aggravate educational inequality (Jennings & Lauen, 2016). In addition, continuous performance pressure may also lead schools to form a short-term performance orientation. It will reduce teachers’ teaching autonomy and space for educational innovation (Klein, 2020). Therefore, although the combination of evaluation results and resource inclination can strengthen schools’ motivation for improvement to a certain extent, it may also induce educational alienation and new fairness risks without reasonable supervision and multi-dimensional evaluation mechanisms.

6.4. Optimizing the Design of Productive-Failure Instruction

The promotion of productive-failure instruction needs to solve the double dilemma of “daring not to use failure” and “abusing failure” in practice. We need to optimize the existing standardized teaching design. The optimization can be carried out in three directions: First, in terms of task design, we should follow the layered design logic of “high challenge–accessible”. That is, we maintain the cognitive complexity of tasks and reduce the entry threshold for students with learning difficulties through progressive scaffolding. The research of Roll et al. (2011) points out that progressive scaffolding and layered support for learning tasks can effectively relieve cognitive load in complex tasks. It allows students with learning difficulties to maintain continuous participation in high-challenge tasks. Second, in terms of intervention tools, we provide a standard paradigm of cognitive scaffolding, such as the “four-step guidance method” (exposing misconceptions–comparing conflicts–reconstructing connections–transfer verification). This paradigm helps teachers grasp the intervention timing more accurately. It avoids the interruption of students’ exploration process caused by premature error correction, and improves the quality of cognitive construction in complex problem learning (van de Pol et al., 2019). Third, in terms of evaluation feedback, we should emphasize the continuous recording of “process evidence chain”, rather than only focusing on the final result (Hattie & Timperley, 2007). Teachers can track students’ cognitive changes in the “attempt–failure–revision” process through learning logs, mind maps, scheme iteration records, or classroom dialogue texts. For example, Novak and Treagust (2022) find in science inquiry classrooms that students’ continuous revision of the causes of experimental phenomena in multiple rounds of explanation iteration significantly promotes their scientific reasoning ability and knowledge transfer level. The evaluation goal in PF classrooms is not to reduce errors. It is to identify how students use errors to promote knowledge reconstruction and strategy optimization.

Author Contributions

J.L.: Writing—original draft, Methodology, Investigation, Funding acquisition, Methodology, Formal analysis, Conceptualization; J.Y. and X.L. (Xiaonan Luo): Conceptualization, Writing—original draft, Writing—review & editing; X.L. (Xueliang Li): Conceptualization, Methodology, Investigation, Writing—original draft, Writing—review & editing; W.C.: Methodology, Writing—original draft. All authors have read and agreed to the published version of the manuscript.

Funding

2024 Provincial Teaching Reform Program for Graduate Students in the Second Batch of the 14th Five-Year Plan of Zhejiang Provincial Office of Education: Innovation and Practice of “Six Synergistic” Graduate Teaching Guided by Educator’s Spirit (Project No:JGCG2024406); 2025 Wenzhou University Graduate Student Innovation Fund Project: Research on Improving Junior High School Students’ Practical Knowledge Through Human-Machine Collaborative Learning Decision (3162025003019); 2025 Wenzhou University Graduate Student Innovation Fund Project: Research on Enhancing Junior High School Students’ Deep Knowledge Learning Effect via Data-Driven Learning Decision (3162025004025); 2026 Zhejiang Provincial College Students’ Science and Technology Innovation Activity Plan (Xinmiao Talent Program) College Students’ Scientific and Technological Achievements Promotion Project: Research on Improving Students’ Deep Knowledge Learning Effect Through Data-Driven Learning Decision-Making (Project No.: 2026R444B038).

Institutional Review Board Statement

Ethical approval for this study was granted by the Wenzhou University (WZU-2025-068). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The research has gained approval from institutional ethics committee.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The data and materials are available upon request to the corresponding author.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilized for this paper.

References

- Agasisti, T., Avvisati, F., Borgonovi, F., & Longobardi, S. (2018). *Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA* (OECD Education Working Papers, No. 167). OECD Publishing. <https://doi.org/10.1787/e22490ac-en>.
- Allen-Platt, C., Gerstner, C.-C., Boruch, R., & Ruby, A. (2021). Toward a science of failure analysis: A narrative review. *Review of Research in Education*, 45(1), 223–252. <https://doi.org/10.3102/0091732x20985074>.
- Ashman, G., Kalyuga, S., & Sweller, J. (2020). Problem-solving or explicit instruction: Which should go first when element interactivity is high? *Educational Psychology Review*, 32(1), 229–247. <https://doi.org/10.1007/s10648-019-09500-5>.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267. <https://doi.org/10.3102/0013189x07306523>.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295x.84.2.191>.
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2024). *Generative AI can harm learning* (SSRN Scholarly Paper No. 4895486). SSRN. <https://doi.org/10.2139/ssrn.4895486>.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>.
- Borman, G. D., Stringfield, S. C., & Slavin, R. E. (2001). *Title I: Compensatory education at the crossroads*. Routledge.
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>.
- Chen, Y., Chui, H., Huang, Y., & King, R. B. (2024). Adaptation and validation of the Chinese version of fear of failure in learning scale. *Journal of Psychoeducational Assessment*, 42(7), 866–881. <https://doi.org/10.1177/07342829241257867>.
- De Bortoli, L., Underwood, C., & Thomson, S. (2023). *PISA 2022. Reporting Australia's results. Volume I: Student performance and equity in education*. Australian Council for Educational Research. <https://doi.org/10.37517/978-1-74286-725-0>.
- Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418–446. <https://doi.org/10.1002/pam.20586>.
- Dee, T. S., Jacob, B., & Schwartz, N. L. (2013). The effects of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis*, 35(2), 252–279. <https://doi.org/10.3102/0162373712467080>.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189x08331140>.
- Dong, Z., Liu, H., & Zheng, X. (2021). The influence of teacher-student proximity, teacher feedback, and near-seated peer groups on classroom engagement: An agent-based modeling approach. *PLoS ONE*, 16(1), e0244935. <https://doi.org/10.1371/journal.pone.0244935>.
- Fang, J., Brown, G. T. L., & Hamilton, R. (2023). Changes in Chinese students' academic emotions after examinations: Pride in success, shame in failure, and self-loathing in comparison. *British Journal of Educational Psychology*, 93(1), 245–261. <https://doi.org/10.1111/bjep.12552>.
- Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal*, 37(5), 813–828. <https://doi.org/10.1080/01411926.2010.501096>.
- Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. *Review of Educational Research*, 82(3), 300–329. <https://doi.org/10.3102/0034654312457206>.
- Gracia-Calandín, J., & Tamarit-López, I. (2021). Education as a common good from the capability approach. *Journal of Philosophy of Education*, 55(4–5), 817–828. <https://doi.org/10.1111/1467-9752.12575>.
- Hadjar, A., & Naumann, I. K. (2025). The digital divide. *sozialpolitik.ch*, 2025(2). <https://doi.org/10.18753/2297-8224-8576>.
- Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal*, 113(485), F64–F98. <https://doi.org/10.1111/1468-0297.00099>.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.

- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524–549. <https://doi.org/10.2307/749690>.
- Ho, D. G., Sa'adi, M., He, D., & Hoon, C. Y. (2023). Silence over the wire: Student verbal participation and the virtual classroom in the digital era. *Asia Pacific Education Review*, 24(4), 599–615. <https://doi.org/10.1007/s12564-023-09834-4>.
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>.
- Jennings, J. L., & Lauen, D. L. (2016). Accountability, inequality, and achievement: The effects of the No Child Left Behind Act on multiple measures of student learning. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2(5), 220–241. <https://doi.org/10.7758/rsf.2016.2.5.11>.
- Jin, S., Fang, G., Cheung, K. C., & Sit, P. S. (2022). Factors associated with academic resilience in disadvantaged students: An analysis based on the PISA 2015 B-S-J-G (China) sample. *Frontiers in Psychology*, 13, 846466. <https://doi.org/10.3389/fpsyg.2022.846466>.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. <https://doi.org/10.1080/07370000802212669>.
- Kapur, M. (2014). Productive failure in learning math. *Cognitive Science*, 38(5), 1008–1022. <https://doi.org/10.1111/cogs.12107>.
- Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, 1(1), 51–65. <https://doi.org/10.1080/23735082.2015.1002195>.
- Khalil, M., & Er, E. (2023). *Will ChatGPT get you caught? Rethinking of plagiarism detection*. Center for Open Science. <https://doi.org/10.35542/osf.io/fnh48>.
- Kieser, F., Tschisgale, P., Rauh, S., Bai, X., Maus, H., Petersen, S., Stede, M., Neumann, K., & Wulff, P. (2024). David vs. Goliath: Comparing conventional machine learning and a large language model for assessing students' concept use in a physics problem. *Frontiers in Artificial Intelligence*, 7, 1408817. <https://doi.org/10.3389/frai.2024.1408817>.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1.
- Klein, E. D. (2020). School accountability. In *Oxford research encyclopedia of education*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190264093.013.677>.
- Koedel, C., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47, 180–195. <https://doi.org/10.1016/j.econedurev.2015.01.006>.
- Kozanitis, A., & Nenciovici, L. (2023). Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: A meta-analysis. *Higher Education*, 86(6), 1377–1394. <https://doi.org/10.1007/s10734-022-00977-8>.
- Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning. *Review of Educational Research*, 86(3), 681–718. <https://doi.org/10.3102/0034654315627366>.
- Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21(1), 27. <https://doi.org/10.1186/s41239-024-00447-4>.
- Lee, J.-A., & Kim, C.-J. (2019). Teaching and learning science in authoritative classrooms: Teachers' power and students' approval in Korean elementary classrooms. *Research in Science Education*, 49(5), 1367–1393. <https://doi.org/10.1007/s11165-017-9659-6>.
- Li, H., Liu, J., Zhang, D., & Liu, H. (2021). Examining the relationships between cognitive activation, self-efficacy, socioeconomic status, and achievement in mathematics: A multi-level analysis. *British Journal of Educational Psychology*, 91(1), 101–126. <https://doi.org/10.1111/bjep.12351>.
- Loibl, K., Roll, I., & Rummel, N. (2017). Towards a theory of when and how problem solving followed by instruction supports learning. *Educational Psychology Review*, 29(4), 693–715. <https://doi.org/10.1007/s10648-016-9379-x>.
- Loibl, K., Tillema, M., Rummel, N., & van Gog, T. (2020). The effect of contrasting cases during problem solving prior to and after instruction. *Instructional Science*, 48(2), 115–136. <https://doi.org/10.1007/s11251-020-09504-7>.
- Mao, M., Zang, L., & Zhang, H. (2020). The effects of parental absence on children development: Evidence from left-behind children in China. *International Journal of Environmental Research and Public Health*, 17(18), 6770. <https://doi.org/10.3390/ijerph17186770>.
- Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, 56(3), 227–238. <https://doi.org/10.1037/0003-066x.56.3.227>.

- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29(1), 67–101. <https://doi.org/10.3102/10769986029001067>.
- Mesquita, B., & Karasawa, M. (2004). Self-conscious emotions as dynamic cultural processes. *Psychological Inquiry*, 15(2), 161–166. <https://www.jstor.org/stable/20447222>.
- Metcalf, J. (2017). Learning from errors. *Annual Review of Psychology*, 68(1), 465–489. <https://doi.org/10.1146/annurev-psych-010416-044022>.
- Meyer, A., Kleinknecht, M., & Richter, D. (2023). What makes online professional development effective? The effect of quality characteristics on teachers' satisfaction and changes in their professional practices. *Computers & Education*, 200, 104805. <https://doi.org/10.1016/j.compedu.2023.104805>.
- Mitra, D. L., & Serriere, S. C. (2012). Student voice in elementary school reform. *American Educational Research Journal*, 49(4), 743–774. <https://doi.org/10.3102/0002831212443079>.
- Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of Applied Psychology*, 95(2), 255–264. <https://doi.org/10.1037/a0018478>.
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), e3292. <https://doi.org/10.1002/rev3.3292>.
- Mu, J., Bayrak, A., & Ufer, S. (2022). Conceptualizing and measuring instructional quality in mathematics education: A systematic literature review. *Frontiers in Education*, 7, 994739. <https://doi.org/10.3389/feduc.2022.994739>.
- Nachtigall, V., Serova, K., & Rummel, N. (2020). When failure fails to be productive: Probing the effectiveness of productive failure for learning beyond STEM domains. *Instructional Science*, 48(6), 651–697. <https://doi.org/10.1007/s11251-020-09525-2>.
- Nehyba, J., Juhaňák, L., & Cigán, J. (2023). Effects of seating arrangement on students' interaction in group reflective practice. *The Journal of Experimental Education*, 91(2), 249–277. <https://doi.org/10.1080/00220973.2021.1954865>.
- Nieuwenhuis, J., & Hooimeijer, P. (2016). The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *Journal of Housing and the Built Environment*, 31(2), 321–347. <https://doi.org/10.1007/s10901-015-9460-7>.
- Novak, A. M., & Treagust, D. F. (2022). Supporting the development of scientific understanding when constructing an evolving explanation. *Disciplinary and Interdisciplinary Science Education Research*, 4(1), 2. <https://doi.org/10.1186/s43031-021-00043-w>.
- Organisation for Economic Co-operation and Development. (2012). *Equity and quality in education: Supporting disadvantaged students and schools*. OECD Publishing. <https://doi.org/10.1787/9789264130852-en>.
- Ouyang, F., & Zhang, L. (2024). AI-driven learning analytics applications and tools in computer-supported collaborative learning: A systematic review. *Educational Research Review*, 44, 100616. <https://doi.org/10.1016/j.edurev.2024.100616>.
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74–98. <https://doi.org/10.1016/j.edurev.2017.08.004>.
- Pelánek, R., Effenberger, T., & Jarušek, P. (2024). Personalized recommendations for learning activities in online environments: A modular rule-based approach. *User Modeling and User-Adapted Interaction*, 34(4), 1399–1430. <https://doi.org/10.1007/s11257-024-09396-z>.
- Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of Education Policy*, 29(5), 640–657. <https://doi.org/10.1080/02680939.2013.865082>.
- Pounder, J. S. (2018). Measuring student success: A value-added approach. In *The future of higher education in the Middle East and Africa* (pp. 205–210). Springer. https://doi.org/10.1007/978-3-319-64656-5_17.
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1), 8–25. <https://doi.org/10.1080/01443410.2015.1136407>.
- Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, 21(2), 267–280. <https://doi.org/10.1016/j.learninstruc.2010.07.004>.
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement. *Review of Educational Research*, 81(4), 493–529. <https://doi.org/10.3102/0034654311421793>.

- Ross, A., & Andreas, J. (2024). Toward in-context teaching: Adapting examples to students' misconceptions. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 13283–13310). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.acl-long.718>.
- Sandsør, A. M. J., Zachrisson, H. D., Karoly, L. A., & Dearing, E. (2023). The widening achievement gap between rich and poor in a Nordic country. *Educational Researcher*, 52(4), 195–205. <https://doi.org/10.3102/0013189x221142596>.
- Schneider, B. (2024). Three challenges in implementing multimodal learning analytics in real-world learning environments. *Learning: Research and Practice*, 10(1), 103–112. <https://doi.org/10.1080/23735082.2023.2270611>.
- Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, 103(4), 759–775. <https://doi.org/10.1037/a0025140>.
- Shabrina, P., Mostafavi, B., Abdelshiheed, M., Chi, M., & Barnes, T. (2024). Investigating the impact of backward strategy learning in a logic tutor: Aiding subgoal learning towards improved problem solving. *International Journal of Artificial Intelligence in Education*, 34(3), 825–861. <https://doi.org/10.1007/s40593-023-00338-1>.
- Simon, P. D., Fryer, L. K., & Nakao, K. (2024). Assessing class participation in physical and virtual spaces: Current approaches and issues. *Frontiers in Education*, 8, 1306568. <https://doi.org/10.3389/educ.2023.1306568>.
- Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist*, 57(1), 31–54. <https://doi.org/10.1080/00461520.2021.1939700>.
- Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, 91(5), 761–798. <https://doi.org/10.3102/00346543211019105>.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>.
- Smidekova, Z., Janik, M., Minarikova, E., & Holmqvist, K. (2020). Teachers' gaze over space and time in a real-world classroom. *Journal of Eye Movement Research*, 13(4), 1. <https://doi.org/10.16910/jemr.13.4.1>.
- Strello, A., Strietholt, R., Steinmann, I., & Siepmann, C. (2021). Early tracking and different types of inequalities in achievement: Difference-in-differences evidence from 20 years of large-scale assessments. *Educational Assessment, Evaluation and Accountability*, 33(1), 139–167. <https://doi.org/10.1007/s11092-020-09346-4>.
- Sum, E. S. W., & Kwon, O. N. (2020). Classroom talk and the legacy of Confucian culture in mathematics classroom. *Teaching and Teacher Education*, 88, 102964. <https://doi.org/10.1016/j.tate.2019.102964>.
- Sun, Y., Zhu, X., Zhou, X., Yao, B., Zhang, K., Wang, D., Chen, J., & He, L. (2024). *Who changed the destiny of rural students, and how? Unpacking ICT-mediated remote education in rural China*. arXiv. <https://arxiv.org/abs/2401.13799>.
- Tai, J., Ajajawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. <https://doi.org/10.1007/s10734-017-0220-3>.
- Tan, C. Y. (2024). Socioeconomic status and student learning: Insights from an umbrella review. *Educational Psychology Review*, 36(4), 99. <https://doi.org/10.1007/s10648-024-09929-3>.
- Terrin, É., & Triventi, M. (2023). The effect of school tracking on student achievement and inequality: A meta-analysis. *Review of Educational Research*, 93(2), 236–274. <https://doi.org/10.3102/00346543221100850>.
- Tweed, R. G., & Lehman, D. R. (2002). Learning considered within a cultural context: Confucian and Socratic approaches. *American Psychologist*, 57(2), 89–99. <https://doi.org/10.1037/0003-066x.57.2.89>.
- van de Pol, J., Mercer, N., & Volman, M. (2019). Scaffolding student understanding in small-group work: Students' uptake of teacher support in subsequent small-group interaction. *Journal of the Learning Sciences*, 28(2), 206–239. <https://doi.org/10.1080/1058406.2018.1522258>.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296. <https://doi.org/10.1007/s10648-010-9127-6>.
- Vaughn, M. (2021). *Student agency in the classroom: Honoring student voice in the curriculum*. Teachers College Press.
- Villegas, A. M., & Lucas, T. (2002). Preparing culturally responsive teachers. *Journal of Teacher Education*, 53(1), 20–32. <https://doi.org/10.1177/0022487102053001003>.
- Wannarka, R., & Ruhl, K. (2008). Seating arrangements that promote positive academic and behavioural outcomes: A review of empirical research. *Support for Learning*, 23(2), 89–93. <https://doi.org/10.1111/j.1467-9604.2008.00375.x>.
- Xu, S., & Lu, C. (2022). Hybridization of curriculum integration and high-stakes testing: Chinese teachers' experiences and struggles in handling competing institutional logics. *Teaching and Teacher Education*, 119, 103845. <https://doi.org/10.1016/j.tate.2022.103845>.
- Xu, W., Wu, Y., & Ouyang, F. (2023). Multimodal learning analytics of collaborative patterns during pair programming in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 3. <https://doi.org/10.1186/s41239-022-00377-z>.

- Yan, J. J., Luo, X. N., Zeng, Q. F., Zhang, H. L., & Wu, J. N. (2025). Investigation and strategy research on the causes of middle school students' learning difficulties in the context of the leading country in education. *Journal of Educational Technology and Innovation*, 7(2), 49. <https://doi.org/10.61414/0shg4n49>.
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>.
- Zhang, H. L., Yan, J. J., Liu, J., Zou, Z. C., Zeng, Q. F., Jin, Y. L., & Wu, J. N. (2025). An empirical study on the reflective feedback model based on pedagogical agents to improve students' problem-solving ability. *Journal of Educational Technology and Innovation*, 7(1), 66–75. <https://doi.org/10.61414/tykx7317>.
- Zhang, Y., Li, M., Chieh, Y., & Han, S. (2022). On a path to becoming more self-regulated: Reflective journals' impact on Chinese English as a foreign language students' self-regulated writing strategy use. *Frontiers in Psychology*, 13, 1042031. <https://doi.org/10.3389/fpsyg.2022.1042031>.
- 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.