From Rules to Transferable Skills

A Research Foundation for Principle Centered Gamified Learning in College

A white paper prepared for BeastCampus faculty, academic leadership, and institutional partners.

Executive Summary

This paper evaluates a proposition rather than assuming it: knowing a rule is one component of learning, and durable, transferable skill develops when students understand the underlying principle, practice applying it across varied contexts, receive useful feedback, and demonstrate independent performance. The research reviewed here supports each component of that chain, with varying strength. Cognitive psychology distinguishes knowing what (declarative knowledge), knowing how (procedural knowledge), and knowing when and why (conditional knowledge), and treats these as interacting parts of competence rather than separate curricula (McGill and Volet; de Jong and Ferguson-Hessler). In mathematics, conceptual and procedural knowledge develop bidirectionally and iteratively, so teaching principles alongside procedures is better supported than teaching either alone (Rittle-Johnson and Schneider; Rittle-Johnson et al., “Not a One-Way Street”). Retrieval practice, spaced and interleaved practice, and worked examples are among the best-supported methods for building durable skill (Roediger and Butler; Rohrer and Pashler; Sweller and Ayres), and feedback is powerful but highly variable—roughly a third of feedback interventions in the classic meta-analysis actually reduced performance (Kluger and DeNisi, “Effects of Feedback Interventions”; Wisniewski et al.). Transfer, especially to unfamiliar tasks, remains the hardest outcome to produce and must be measured directly rather than inferred from performance on supported exercises (Perkins and Salomon, “Transfer of Learning”; Engle). Error-centered instruction has a growing evidence base, including the counterintuitive finding that confidently held errors are often corrected more readily than low-confidence errors (Metcalfe; Butterfield and Metcalfe). Gamification shows small-to-medium average effects on engagement and learning in meta-analyses, but effects vary with duration and design and do not guarantee learning (Sailer and Homner; Kim and Castelli; Dichev and Dicheva). Corequisite developmental education has strong causal evidence at the college level for gateway-course completion, while longer-term findings vary by study (Ran and Lin; Logue et al.; Miller et al.). None of this constitutes evidence that BeastCampus itself works; the platform's specific combination of methods remains untested and should be evaluated through the pilot proposed in Section 5.

1. The Instructional Problem, Defined Carefully

BeastCampus is designed to move students from recalling or mechanically applying rules toward understanding principles, exercising judgment, and transferring skills to their own work. Framing that goal requires precise definitions, because the relevant terms are overlapping dimensions of instruction, not established, mutually exclusive research categories.

Rules-based instruction refers to teaching that centers explicit statements of what to do—"multiply by the reciprocal," "a comma cannot join two independent sentences." Skills-based instruction centers fluent execution of procedures through practice. Principle-centered instruction centers the reasons a rule or procedure works—the concept of division as multiplicative comparison, or of the clause as a unit of predication—so that students can judge when and how to apply what they know. These are not rival

pedagogies validated as packages by research; they are emphases that can coexist in a single lesson. The closest research analogues are knowledge-type frameworks that decompose competence into declarative, procedural, and conditional knowledge (McGill and Volet; de Jong and Ferguson-Hessler), and mathematics-education work showing that the conceptual/procedural distinction, while useful analytically, describes interacting knowledge types rather than competing instructional camps (Hiebert and Lefevre; Kieran; Österman and Bråting).

Two clarifications prevent common misreadings. First, "rules-based" here has nothing to do with the term's meaning in artificial intelligence, where rule-based systems denote production architectures that encode knowledge as condition-action pairs; that technical sense concerns machine knowledge representation, not classroom teaching (Nilsson). Second, the historical claim that education has primarily emphasized rules should be evaluated rather than assumed. The tension between procedural skill and conceptual understanding is a long-running debate in mathematics education, not a recent discovery, and reform movements emphasizing meaning have recurred for decades (Hiebert and Lefevre; Kieran). Educational history is better described as recurring swings of emphasis than as a simple progression from rules to skills.

Finally, principle-centered instruction must not be equated with minimally guided discovery learning. A substantial body of evidence indicates that novices learn more from strongly guided instruction—explicit explanations, worked examples, and scaffolded practice—than from unguided exploration, because working memory limits make search inefficient for learners who lack prior knowledge (Kirschner et al.; van Merriënboer and Sweller). A recent meta-analysis of the expertise reversal effect confirms the pattern: learners with lower prior knowledge benefit more from high-assistance instruction, and the advantage of guidance recedes as expertise grows (Tetzlaff et al.). Explicit instruction, memorization, and procedural practice are therefore valuable components of a principle-centered system, not its opposite (Kirschner et al.; Roediger and Butler). The amount of guidance must also be considered in relation to task complexity and learners' existing knowledge (Kalyuga and Singh); the defensible conclusion is that guidance should be strong early and faded as competence develops (Renkl and Atkinson; Tetzlaff et al.).

2. Research Synthesis

2.1 Knowing What, How, When, and Why

Research on knowledge types distinguishes declarative knowledge (facts, concepts, principles), procedural knowledge (how to perform operations), and conditional knowledge (when and why to apply them) (McGill and Volet; de Jong and Ferguson-Hessler). The distinction applies directly to grammar, writing, and mathematics: a student may know the rule for apostrophes (declarative), apply it on a worksheet (procedural), and still fail to notice the error in their own draft—the conditional gap (van Velzen; Sterner et al.). A study of public understanding of stock-flow problems found that people who possessed both the declarative and procedural components still failed tasks that did not signal which knowledge to use, isolating conditional knowledge as the missing piece (Sterner et al.). Work in intelligent tutoring systems similarly treats progression from declarative through procedural to conditional knowledge as the pathway toward independent performance (Abdelshiheed et al.). Evidence strength: strong for the knowledge-type framework as an analytic tool; limited as a prescription for how to teach each type, since the types develop together rather than sequentially (Rittle-Johnson and Schneider; Rittle-Johnson et al., "Not a One-Way Street").

2.2 Principles and Conceptual Understanding

The strongest finding in this literature is that conceptual and procedural knowledge are bidirectionally related: gains in one support gains in the other, and they develop iteratively rather than in a fixed sequence (Rittle-Johnson and Schneider; Rittle-Johnson et al., “Not a One-Way Street”; Rittle-Johnson et al., “Developing Conceptual Understanding”). Longitudinal work on decimal fractions showed that initial conceptual knowledge predicted procedural gains, that procedural gains predicted conceptual improvement, and that improved problem representation mediated the relationship (Rittle-Johnson et al., “Developing Conceptual Understanding”). This undermines both the claim that procedures must precede concepts and the claim that concepts must precede procedures; what matters is that instruction deliberately connects them (Rittle-Johnson et al., “Not a One-Way Street”; Byrnes and Wasik). Scholars of algebra warn that treating concepts and procedures as a dichotomy is itself damaging, because the procedural in algebra is saturated with conceptual content (Kieran), and frameworks of mathematical proficiency treat conceptual understanding and procedural fluency as intertwining strands among several that must develop together (Groves). Evidence strength: strong for bidirectional relations; mixed and limited on optimal instructional sequencing, which has rarely been compared directly (Rittle-Johnson et al., “Not a One-Way Street”).

2.3 Practice and Feedback: Distinguishing the Methods

These methods are not interchangeable, and their evidence bases differ.

- Retrieval practice. Actively recalling information produces better long-term retention and transfer than restudying, an effect that generalizes broadly across populations, materials, and settings (Roediger and Butler; Zepeda et al.). A systematic review of classroom experiments found most reported medium or large benefits (Agarwal et al.). However, retrieval's advantage over other effective activities (such as elaborative encoding) is small ($g = 0.14$) and depends heavily on whether corrective feedback is provided (Gonçalves et al.). Evidence strength: strong, with feedback as an important moderator.
- Distributed and interleaved practice. Learning is more durable when practice is spread over time, and interleaving different problem types forces students to choose a strategy rather than execute a given one (Rohrer and Pashler; Roediger and Pyc). In a classroom experiment with seventh graders, interleaved mathematics practice produced higher scores on immediate ($d = 0.42$) and delayed ($d = 0.79$) tests than blocked practice (Rohrer et al.), and undergraduates in physics showed large improvements on criterial tests while believing they had learned less from interleaving (Samani and Pan). A meta-analysis puts the overall interleaving effect at $g = 0.42$, smaller for mathematics tasks ($g = 0.34$) (Brunmair and Richter). Evidence strength: strong in principle; classroom implementation remains sensitive to task and design; Ruitenburg and colleagues’ conference presentation provides a supplementary classroom-focused review (Ruitenburg et al.).
- Worked examples and fading. For novices, studying fully worked solutions is often more effective and efficient than solving equivalent problems, because it reduces extraneous working-memory load (Sweller and Ayres; Sweller and Cooper). Design principles include multiple varied examples, self-explanation prompts, and gradual fading of worked steps as learners gain competence (Atkinson et al.; Renkl; Renkl and Atkinson). The worked-example advantage persists even against well-supported tutored problem solving (Schwonke et al.), and meta-analytic work identifies faded steps and self-explanation training as the strongest design features (Crissman). Notably,

simply telling learners explanations has minimal benefit; learners must process them, and instructional explanations are not superior to prompting self-explanation (Wittwer and Renkl). Evidence strength: strong for novices; the effect reverses for experienced learners (Tetzlaff et al.; Leppink et al.).

- Formative assessment and feedback. Feedback is among the most powerful influences on learning, but its effects are highly variable, and the type, timing, and focus of feedback determine whether it helps (Hattie and Timperley; Wisniewski et al.). A meta-analysis of 435 studies found a medium average effect (d = 0.48) with substantial heterogeneity (Wisniewski et al.), and the classic feedback meta-analysis found that over a third of feedback interventions reduced performance, particularly when feedback directs attention to the self rather than the task (Kluger and DeNisi, “Effects of Feedback Interventions”). A rigorous systematic review of feedback-only interventions in schools estimated a much smaller pooled effect (g = 0.17) with significant heterogeneity, cautioning against blanket claims (Newman et al.). Feedback also corrects metacognitive errors, improving retention of low-confidence correct responses (Butler et al., “Correcting a Metacognitive Error”), and feedback is integral to self-regulated learning, where students must interpret and act on it (Butler and Winne). Evidence strength: strong that feedback matters; strong that "more feedback" is not a prescription—design determines outcome.
- Revision. In writing instruction, feedback can contribute to improvement when students use it to revise their work. Reviews of instruction for adolescents support several approaches, including strategy instruction, goal setting, and feedback, while an additional meta-analysis addresses younger children rather than college students (Graham and Perin; Graham et al.; Koster et al.). These findings support purposeful revision opportunities, but they do not establish that any particular automated feedback system improves college writing. Evidence strength: moderate and population-dependent; effects depend on what students do with feedback.

2.4 Transfer: Why Isolated Success Is Not Enough

Students can perform well on an isolated exercise yet fail to apply the same skill in an essay or an unfamiliar problem because transfer is not automatic. Classic theory distinguishes "low-road" transfer, in which well-practiced routines trigger in similar contexts, from "high-road" transfer, which requires deliberate abstraction of a principle and mindful search for connections; conventional instruction often establishes the conditions for neither (Perkins and Salomon, “Transfer of Learning”; Perkins and Salomon, “Transfer and Teaching Thinking”). Contemporary researchers similarly define transfer as the productive, supported use of knowledge in new contexts and note that traditional measures—near-identical test items—systematically overstate it (De Corte; Engle). Salmerón’s literature review discusses instructional activities intended to support transfer, including attention to underlying principles and connections across contexts; their usefulness depends on how the activities are implemented (Salmerón). Self-explanation studies in mathematics show improved procedural transfer when knowledge is assessed immediately, with weaker evidence for delayed transfer (Rittle-Johnson, “Promoting Transfer”; Rittle-Johnson et al., “Promoting Self-Explanation”). Evidence strength: strong on the diagnosis (transfer is hard and often absent); mixed on remedies; far transfer remains genuinely difficult and should be measured directly with unfamiliar tasks (Perkins and Salomon, “Transfer of Learning”; Engle).

2.5 Misconceptions and Targeted Instruction

Identifying specific errors and targeting them with focused practice has a distinctive evidence base. Counterintuitively, errors committed with high confidence are more likely to be corrected by feedback than low-confidence errors—the hypercorrection effect—because surprising feedback captures attention (Metcalfe; Butterfield and Metcalfe). The effect appears in authentic college classrooms (Carpenter et al.), though high-confidence errors that are not successfully corrected tend to return after a delay, so retesting matters (Butler et al., "Hypercorrection Effect"). Refutation-based instruction that explicitly activates a misconception and corrects it produces hypercorrection in reading contexts (van Loon et al.), and prior knowledge predicts error correction at least as strongly as confidence itself (Sitzman et al.). In mathematics, incorporating incorrect examples and prompting students to compare multiple examples improves learning, and framing errors as a functional part of learning is recommended (Booth et al.; Rittle-Johnson, "Encouraging Students"). Evidence strength: moderate to strong in laboratory and some classroom settings; the mechanisms are well replicated, but long-term durability of corrections requires spaced retesting (Butler et al., "Hypercorrection Effect"; Metcalfe).

2.6 Gamification: Separating It from Game-Based Learning

Gamification—applying game-design elements such as points, badges, and challenges in non-game contexts—is distinct from game-based learning, in which a complete game is the instructional medium (Plass et al.; Zhang and Yu). Meta-analyses of gamification report small-to-medium effects: cognitive outcomes $g = 0.49$, motivational $g = 0.36$, and behavioral $g = 0.25$ in one rigorous synthesis (Sailer and Homner); $g = 0.464$ for learning outcomes in another (Huang et al.); and medium effects on affective ($g = 0.574$) and behavioral ($g = 0.740$) outcomes in a third (Ritzhaupt et al.). A separate meta-analysis reports positive effects of digital game-based learning in school settings (Barz et al.). These are different interventions and study pools, so their average effects should not be read as a direct head-to-head comparison (Zhang and Yu). Design features matter: game fiction and competition combined with collaboration improved behavioral outcomes in the Sailer and Homner meta-analysis (Sailer and Homner), and design-element comparisons also showed variation across studies (Huang et al.). The cautions are equally important. A critical review concluded there is insufficient evidence for long-term benefits and that practice has outpaced understanding of mechanisms (Dichev and Dicheva). A meta-analysis of behavioral change found larger effects in shorter interventions and weaker or negative estimates in some longer-duration categories (Kim and Castelli). These are comparisons across studies, not proof that a given intervention inevitably loses effectiveness over time. In computerized cognitive training, gamified tasks were more motivating ($g = 0.72$) but showed no effect on the trained domain itself, and game elements can distract (Vermeir et al.). A higher-education meta-analysis also reports positive average effects with substantial heterogeneity, underscoring variation across implementations (Gyedu et al.). Evidence strength: mixed—positive average short-term effects in several reviews on engagement and some learning outcomes; novelty and duration effects, reward dependence, and learner-preference differences remain live concerns (Kim and Castelli; Dichev and Dicheva; Vermeir et al.).

2.7 Metacognition and Independence

Self-explanation—generating one's own explanations of why and how—has a meta-analytic effect of $g = 0.55$ across instructional conditions (Bisra et al.). In mathematics specifically, prompted self-explanation yields small-to-moderate immediate gains in procedural and conceptual knowledge and transfer, but evidence for classroom effectiveness and delayed retention is much more limited (Rittle-Johnson et al., "Promoting Self-Explanation"; Rittle-Johnson, "Encouraging Students"). Self-explanation is not uniformly

beneficial: prompts focus attention and can crowd out other content, and explaining one's own incorrect method can be harmful, so prompt design matters (Rittle-Johnson and Loehr). Compared with spending the same time on additional practice, self-explanation's unique benefits are modest (McEldoon et al.). Classroom implementations are feasible: students who explained their steps in a cognitive tutor learned with greater understanding and better transfer (Aleven and Koedinger), and self-explanation training improved far transfer in secondary mathematics (Wong et al.). In college students learning statistics, prompted self-explanation improved initial learning but not retention a semester later (Ryan and Koppenhofer). Confidence judgments support monitoring: feedback corrects both memory errors and metacognitive errors, and metacognitive experiences such as surprise and overconfidence are tractable instructional variables (Butler et al., "Correcting a Metacognitive Error"; Efklides). Self-regulated learning interventions in higher-education e-learning environments show positive outcomes, and comprehensive interventions with school populations show durable achievement effects (Rowe; Cousins et al.). Evidence strength: moderate to strong for self-explanation and monitoring activities as components; limited for their durability without continued support.

2.8 College and Corequisite Relevance

Populations must be labeled honestly. Much of the foundational evidence above comes from children, adolescents, or laboratory studies with college samples (Chi et al., "Eliciting Self-Explanations"; Chi et al., "Self-Explanations"; Brunstein et al.), and applying it to college students is reasonable but indirect. Directly collegiate evidence exists for several claims: interleaved practice in undergraduate physics (Samani and Pan), worked-example advantages for low-prior-knowledge statistics students (Leppink et al.), hypercorrection in a college classroom (Carpenter et al.), self-explanation with college students (Ryan and Koppenhofer), peer feedback in STEM higher education ($d = 0.421$) (Pereira et al.), and gamification effects in higher education (Gyedu et al.; Barata et al.).

Corequisite developmental education includes college-specific experimental and quasi-experimental evidence. A statewide regression-discontinuity study in Tennessee found corequisite placement increased gateway math passage by 15 percentage points and gateway English by 13 points within one year, with no significant effects on persistence or degree completion (Ran and Lin). A randomized trial found that students assigned to college-level statistics with workshops passed at a rate 16 percentage points higher than students assigned to traditional remedial algebra and subsequently accumulated more credits (Logue et al.). A seven-year follow-up reports degree-completion and earnings outcomes (Douglas et al.). Separately, Smith's comparison of College Algebra groups found an association between corequisite participation and course success; it was not the randomized statistics trial (Smith). A randomized trial in Texas found corequisite English increased first college-level English completion by 24 percentage points within one year, with positive effects across models and for underrepresented groups (Miller et al.). A meta-analysis of regression-discontinuity studies found traditional developmental placement itself associated with negative effects on college-level passage, credits, and attainment (Valentine et al.). Heterogeneity matters: effects are stronger for higher-scoring remedial students in math, and some analyses find increased dropout among the lowest-scoring students (Ran and Lee); design elements such as shared instructors and support-course credits predict better outcomes (Ryu et al.); and the method of acceleration matters, with corequisite approaches tending to help most in math (Mokher et al.; Schudde and Keisler). Crucially for BeastCampus, what happens inside the support course is understudied: reviews of developmental mathematics pedagogy find most instructional studies methodologically weak and results inconclusive (Hodara). Evidence strength: strong for the corequisite structure as a gateway-

completion intervention; limited for the pedagogy of the support course itself—which is precisely where a platform like BeastCampus operates and where it must generate its own evidence.

3. From Rule to Principle: Two College Examples

Writing: comma splices and fragments. A rules-based presentation states that a comma alone conventionally does not join two independent clauses in formal academic prose. A principle-centered presentation explains the underlying unit—the clause as a complete predication—and how punctuation and conjunctions express relationships between complete clauses. A research-informed illustrative progression would be: study a worked (annotated) example of the principle applied (Atkinson et al.; Renkl); practice on short, focused challenges with immediate corrective feedback and retries (Roediger and Butler; Hattie and Timperley); self-explain why the error violates the principle (Bisra et al.; Rittle-Johnson et al., “Promoting Self-Explanation”); then detect and revise the error in a parallel example before scanning one's own draft—an independent-application task, because success on an isolated exercise does not by itself establish independent application (Sterner et al.; Perkins and Salomon, “Transfer of Learning”). This sequence is consistent with evidence that contextualized grammar instruction can improve writing in some secondary-school settings, with outcomes varying by learner (Jones et al.), that isolated grammar instruction historically showed null or negative effects on writing quality (Graham and Perin; Lancaster and Olinger), and that written corrective feedback can improve second-language writing accuracy (Brown et al.).

Mathematics: fraction division or signed-number operations. A rules-based presentation gives "invert and multiply." A principle-centered presentation treats division as multiplicative comparison and connects the procedure to why it works. An illustrative progression: study a process-oriented worked example that includes the "why" (van Gog et al.; Sweller and Ayres); complete faded examples that progressively remove steps (Renkl and Atkinson; Crissman); practice with interleaved problem types so the student must choose the operation, not merely execute it (Rohrer et al.; Samani and Pan); receive feedback that corrects both the error and the confidence judgment (Butler et al., “Correcting a Metacognitive Error”; Butterfield and Metcalfe); then solve an unfamiliar, context-embedded problem without cues—the transfer test that isolated practice does not guarantee (Perkins and Salomon, “Transfer of Learning”; Engle). Research suggests that conceptual and procedural knowledge can support one another; the effectiveness of this particular sequence remains to be tested (Rittle-Johnson and Schneider; Rittle-Johnson et al., “Not a One-Way Street”).

4. Mapping BeastCampus Design Intentions to Research Evidence

The following table maps each stated design intention to the most relevant research, with an evidence-strength judgment and key limitations. These are design intentions, not demonstrated outcomes of the platform.

Design intention	Relevant research	Evidence strength	Limitations
Understanding the principle behind a rule	Conceptual-procedural knowledge develop bidirectionally and iteratively (Rittle-Johnson and Schneider; Rittle-Johnson et al., “Not a One-Way Street”; Rittle-Johnson et al., “Developing Conceptual Understanding”)	Strong (relations); mixed (sequencing)	Optimal ordering of concept/procedure instruction rarely tested directly (Rittle-Johnson et al., “Not a One-Way Street”)
Recognizing when, why, and how a principle applies	Conditional knowledge as a distinct component often missing while declarative/procedural knowledge are present (McGill and Volet; Sterner et al.; Abdelshiheed et al.)	Indirect to moderate	Framework is analytic; fewer intervention studies target conditional knowledge directly (de Jong and Ferguson-Hessler; Sterner et al.)
Short, focused challenges	Retrieval practice improves retention and transfer broadly (Roediger and Butler; Zepeda et al.; Agarwal et al.)	Strong	The advantage over elaborative study activities depends on feedback and task design (Gonçalves et al.).
Explanatory feedback with retry opportunities	Feedback powerful but variable; over a third of interventions reduced performance; design determines effect (Hattie and Timperley; Kluger and DeNisi, “Effects of Feedback Interventions”; Wisniewski et al.)	Strong (with conditions)	Pooled effects in rigorous school studies are modest (g = 0.17); self-focused feedback can harm (Newman et al.; Kluger and DeNisi, “Effects of Feedback Interventions”)
Identifying weaknesses; targeted practice	Deliberate practice targets weaknesses with feedback and repetition (Ericsson; Moulaert et al.); formative assessment practice (Beesley)	Mixed	Formative-assessment evidence base has been criticized on methodological grounds (Dunn and Mulvenon); adaptive guidance must track expertise (Tetzlaff et al.)
Parallel example before one's own work	Worked examples for novices; fading to independent problem solving (Sweller and Ayres; Atkinson et al.; Renkl and Atkinson; Schwonke et al.)	Strong for novices	Reverses for experienced learners; explanations must be processed, not just provided (Tetzlaff et al.; Wittwer and Renkl)
Demonstrating independent application	Transfer requires deliberate abstraction and varied practice; must be measured on unfamiliar tasks (Perkins and Salomon, “Transfer of Learning”; Engle; Salmerón)	Mixed; hardest outcome	Far transfer frequently fails; supported performance does not establish independent mastery (Perkins and Salomon, “Transfer of Learning”; Bransford and Schwartz)
Named "beasts" for recurring errors	Hypercorrection of high-confidence errors; refutation and incorrect-example instruction (Metcalfe; van Loon et al.; Booth et al.)	Indirect for character naming	Error-learning studies do not establish that named characters improve memory or transfer. That design choice requires direct evaluation (Metcalfe; Booth et al.).
Game elements for participation and persistence	Gamification meta-analyses: small-to-medium cognitive/behavioral effects; design-dependent (Sailer and Homner; Huang et al.; Ritzhaupt et al.)	Mixed	Effects vary across duration categories and contexts; engagement does not establish learning. Cross-study differences do not demonstrate inevitable decay (Kim and Castelli; Dichev and Dicheva; Vermeir et al.).

5. A Proposed Pilot Evaluation

Because BeastCampus's effectiveness has not been formally evaluated, the pilot should test the platform as an instructional supplement, not merely measure satisfaction or time on task. The design should be a preregistered randomized or well-matched quasi-experimental study within corequisite and full-course sections, comparing BeastCampus-supported practice with the existing support structure, with randomization at the section level where feasible and baseline covariate adjustment regardless.

Outcomes should form a hierarchy that mirrors the platform's own logic: (a) proximal mastery—posttest on trained skills; (b) retention—unannounced delayed test at four to six weeks, since durability is where many interventions fail (Rittle-Johnson et al., "Promoting Self-Explanation"; Butler et al., "Correcting a Metacognitive Error"); (c) near transfer—novel problems of the same type requiring strategy selection, a demand addressed by interleaved practice (Rohrer et al.; Samani and Pan); (d) independent application—the decisive outcome: unaided detection and revision of errors in the student's own essay, and unaided solution of context-embedded mathematics problems, because performance on isolated exercises is insufficient evidence of conditional knowledge and transfer (Sterner et al.; Perkins and Salomon, "Transfer of Learning"); and (e) gateway course outcomes—pass rates in the paired college-level course, aligning with the outcome conventions of the corequisite literature (Ran and Lin; Miller et al.). Secondary measures should include calibration accuracy (confidence judgments against performance) (Butler et al., "Correcting a Metacognitive Error"; Efklides), persistence and dosage, and accessibility outcomes across student subgroups, given documented heterogeneity in who benefits from reforms (Ran and Lee). Satisfaction and engagement metrics should be reported as descriptive data, never as evidence of learning, since gamified practice can raise motivation without affecting the trained domain (Vermeir et al.).

6. Conclusion: What the Research Supports and What BeastCampus Must Still Test

The research supports the components of BeastCampus's approach with uneven but genuine strength: connecting principles to procedures (Rittle-Johnson and Schneider; Rittle-Johnson et al., "Not a One-Way Street"), retrieval-based and interleaved practice (Roediger and Butler; Rohrer et al.), worked examples with fading for novices (Sweller and Ayres; Renkl and Atkinson), carefully designed explanatory feedback (Hattie and Timperley; Wisniewski et al.), error- and misconception-centered instruction (Metcalfe; van Loon et al.), and metacognitive activities such as self-explanation and confidence judgment (Bisra et al.; Butler et al., "Correcting a Metacognitive Error"). It supports none of the stronger claims: that gamification guarantees improved learning (Dichev and Dicheva; Kim and Castelli), that engagement demonstrates mastery (Vermeir et al.), or that any platform's specific combination of methods works as assembled. The honest position for institutional partners is that BeastCampus is built from evidence-aligned ingredients whose combined effect—especially on delayed retention, transfer, and independent application in college writing and mathematics—remains an open empirical question. The pilot in Section 5 is the appropriate next step: measure learning, retention, transfer, and independent performance, preregister the analysis, report null results faithfully, and let the evidence, not the design intentions, determine the platform's claims.

Works Cited

Abdelshiheed, Mark, et al. “Bridging Declarative, Procedural, and Conditional Metacognitive Knowledge Gap Using Deep Reinforcement Learning.” arXiv, 23 Apr. 2023. Preprint. https://arxiv.org/abs/2304.11739.

Agarwal, Pooja K., et al. “Retrieval Practice Consistently Benefits Student Learning: A Systematic Review of Applied Research in Schools and Classrooms.” Educational Psychology Review, vol. 33, no. 4, 2021, pp. 1409–1453. https://doi.org/10.1007/s10648-021-09595-9.

Aleven, Vincent A.W.M.M., and Kenneth R. Koedinger. “An effective metacognitive strategy: learning by doing and explaining with a computer-based Cognitive Tutor.” Cognitive Science, vol. 26, no. 2, 2002, pp. 147-179, https://doi.org/10.1207/s15516709cog2602_1.

Atkinson, Robert K., et al. “Learning from Examples: Instructional Principles from the Worked Examples Research.” Review of Educational Research, vol. 70, no. 2, 2000, pp. 181-214, https://doi.org/10.3102/00346543070002181.

Barata, Gabriel, et al. “Improving participation and learning with gamification.” Proceedings of the First International Conference on Gameful Design, Research, and Applications, 2013, pp. 10-17, https://doi.org/10.1145/2583008.2583010.

Barz, Nathalie, et al. “The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis.” Review of Educational Research, vol. 94, no. 2, 2024, pp. 193-227, https://doi.org/10.3102/00346543231167795.

Beesley, Andrea D. Measuring Classroom Assessment with a Work Sample. American Educational Research Association annual meeting paper, 2009. ERIC, ED508465. https://files.eric.ed.gov/fulltext/ED508465.pdf.

Bisra, Kiran, et al. “Inducing Self-Explanation: a Meta-Analysis.” Educational Psychology Review, vol. 30, no. 3, 2018, pp. 703-725, https://doi.org/10.1007/s10648-018-9434-x.

Booth, Julie L., et al. “Simple Practice Doesn’t Always Make Perfect.” Policy Insights from the Behavioral and Brain Sciences, vol. 2, no. 1, 2015, pp. 24-32, https://doi.org/10.1177/2372732215601691.

Bransford, John D., and Daniel L. Schwartz. “Chapter 3: Rethinking Transfer: A Simple Proposal With Multiple Implications.” Review of Research in Education, vol. 24, no. 1, 1999, pp. 61-100, https://doi.org/10.3102/0091732x024001061.

Brown, Dan, et al. “Effectiveness of written corrective feedback in developing L2 accuracy: A Bayesian meta-analysis.” Language Teaching Research, vol. 30, no. 3, 2026, pp. 1357-1389, https://doi.org/10.1177/13621688221147374.

Brunmair, Matthias, and Tobias Richter. “Similarity matters: A meta-analysis of interleaved learning and its moderators.” Psychological Bulletin, vol. 145, no. 11, 2019, pp. 1029-1052, https://doi.org/10.1037/bul0000209.

Brunstein, Angela, et al. “Practice enables successful learning under minimal guidance.” Journal of Educational Psychology, vol. 101, no. 4, 2009, pp. 790-802, https://doi.org/10.1037/a0016656.

Butler, Andrew C., et al. “Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses.” Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 34, no. 4, 2008, pp. 918-928, https://doi.org/10.1037/0278-7393.34.4.918.

Butler, Andrew C., et al. “The hypercorrection effect persists over a week, but high-confidence errors return.” Psychonomic Bulletin & Review, vol. 18, no. 6, 2011, pp. 1238-1244, https://doi.org/10.3758/s13423-011-0173-y.

Butler, Deborah L., and Philip H. Winne. “Feedback and Self-Regulated Learning: A Theoretical Synthesis.” Review of Educational Research, vol. 65, no. 3, 1995, pp. 245-281, https://doi.org/10.3102/00346543065003245.

Butterfield, Brady, and Janet Metcalfe. “The correction of errors committed with high confidence.” Metacognition and Learning, vol. 1, no. 1, 2006, pp. 69-84, https://doi.org/10.1007/s11409-006-6894-z.

Byrnes, James P., and Barbara A. Wasik. “Role of conceptual knowledge in mathematical procedural learning.” Developmental Psychology, vol. 27, no. 5, 1991, pp. 777-786, https://doi.org/10.1037/0012-1649.27.5.777.

Carpenter, Shana K., et al. “Hypercorrection of high-confidence errors in the classroom.” Memory, vol. 26, no. 10, 2018, pp. 1379-1384, https://doi.org/10.1080/09658211.2018.1477164.

Chi, Michelene T.H., et al. “Eliciting Self-Explanations Improves Understanding.” Cognitive Science, vol. 18, no. 3, 1994, pp. 439-477, https://doi.org/10.1207/s15516709cog1803_3.

Chi, Michelene T.H., et al. “Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems.” Cognitive Science, vol. 13, no. 2, 1989, pp. 145-182, https://doi.org/10.1207/s15516709cog1302_1.

Cousins, Erin, et al. “Exploring Long-Term Impacts of Self-Regulated Learning Interventions in K-12 Contexts: A Systematic Review.” Current Issues in Education, vol. 23, no. 1, 2022, https://doi.org/10.14507/cie.vol23iss1.2013.

Crissman, Jennifer K. The Design and Utilization of Effective Worked Examples: A Meta-Analysis. 2006. University of Nebraska–Lincoln, doctoral dissertation. https://digitalcommons.unl.edu/dissertations/AAI3208114/.

De Corte, Erik. “Transfer as the Productive Use of Acquired Knowledge, Skills, and Motivations.” Current Directions in Psychological Science, vol. 12, no. 4, 2003, pp. 142-146, https://doi.org/10.1111/1467-8721.01250.

de Jong, Ton, and Monica G.M. Ferguson-Hessler. “Types and qualities of knowledge.” Educational Psychologist, vol. 31, no. 2, 1996, pp. 105-113, https://doi.org/10.1207/s15326985ep3102_2.

Dichev, Christo, and Darina Dicheva. “Gamifying education: what is known, what is believed and what remains uncertain: a critical review.” International Journal of Educational Technology in Higher Education, vol. 14, no. 1, 2017, https://doi.org/10.1186/s41239-017-0042-5.

Douglas, Daniel, et al. “The Long-Term Impacts of Corequisite Mathematics Remediation with Statistics: Degree Completion and Wage Outcomes.” Educational Researcher, vol. 52, no. 1, 2023, pp. 7-15, https://doi.org/10.3102/0013189X221138848.

Dunn, Karee E., and Sean W. Mulvenon. “A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education.” Practical Assessment, Research & Evaluation, vol. 14, article 7, 2009. https://openpublishing.library.umass.edu/pare/article/1520/galley/1471/download/.

Efklides, Anastasia. “Commentary: How readily can findings from basic cognitive psychology research be applied in the classroom?.” Learning and Instruction, vol. 22, no. 4, 2012, pp. 290-295, https://doi.org/10.1016/j.learninstruc.2012.01.001.

Engle, Randi A.. “The Resurgence of Research Into Transfer: An Introduction to the Final Articles of the Transfer Strand.” Journal of the Learning Sciences, vol. 21, no. 3, 2012, pp. 347-352, https://doi.org/10.1080/10508406.2012.707994.

Ericsson, K. Anders. “Deliberate Practice and Acquisition of Expert Performance: A General Overview.” Academic Emergency Medicine, vol. 15, no. 11, 2008, pp. 988-994, https://doi.org/10.1111/j.1553-2712.2008.00227.x.

Gonçalves, Ariel de Oliveira, et al. “Retrieval Practice Versus Elaborative Encoding: A Systematic and Meta-analytic Review.” Educational Psychology Review, vol. 37, no. 4, 2025, https://doi.org/10.1007/s10648-025-10076-6.

Graham, Steve, and Dolores Perin. “A meta-analysis of writing instruction for adolescent students.” Journal of Educational Psychology, vol. 99, no. 3, 2007, pp. 445-476, https://doi.org/10.1037/0022-0663.99.3.445.

Graham, Steve, et al. “Effective writing instruction for students in grades 6 to 12: a best evidence meta-analysis.” Reading and Writing, vol. 38, no. 4, 2025, pp. 1-46, https://doi.org/10.1007/s11145-024-10539-2.

Groves, Susie. “Developing Mathematical Proficiency.” Journal of Science and Mathematics Education in Southeast Asia, vol. 35, no. 2, 2012, pp. 119–145. https://recsam.edu.my/sub_JSMESEA/images/journals/YEAR2012/dec2012vol-2/119-145.pdf.

Gyedu, Francis Obeng, et al. “Gamification Effect: Enhancing Student Learning Outcomes in Higher Education: A Meta-Analysis and Policy Implications.” Sage Open, vol. 16, no. 1, 2026, https://doi.org/10.1177/21582440261421375.

Hattie, John, and Helen Timperley. “The Power of Feedback.” Review of Educational Research, vol. 77, no. 1, 2007, pp. 81-112, https://doi.org/10.3102/003465430298487.

Hiebert, James, and Patricia Lefevre. “Conceptual and Procedural Knowledge in Mathematics: An Introductory Analysis.” Conceptual and Procedural Knowledge: The Case of Mathematics, edited by James Hiebert, Lawrence Erlbaum Associates, 1986, pp. 1–27. https://doi.org/10.4324/9780203063538.

Hodara, Michelle. Reforming Mathematics Classroom Pedagogy: Evidence-Based Findings and Recommendations for the Developmental Math Classroom. Community College Research Center, Teachers College, Columbia University, 2011. Working Paper no. 27, Assessment of Evidence Series. https://academiccommons.columbia.edu/doi/10.7916/D8W66V0S.

Huang, Rui, et al. “The impact of gamification in educational settings on student learning outcomes: a meta-analysis.” Educational Technology Research and Development, vol. 68, no. 4, 2020, pp. 1875-1901, https://doi.org/10.1007/s11423-020-09807-z.

Jones, Susan, et al. “Grammar for writing? An investigation of the effects of contextualised grammar teaching on students’ writing.” Reading and Writing, vol. 26, no. 8, 2013, pp. 1241-1263, https://doi.org/10.1007/s11145-012-9416-1.

Kalyuga, Slava, and Anne-Marie Singh. “Rethinking the Boundaries of Cognitive Load Theory in Complex Learning.” Educational Psychology Review, vol. 28, no. 4, 2016, pp. 831-852, https://doi.org/10.1007/s10648-015-9352-0.

Kieran, Carolyn. “The False Dichotomy in Mathematics Education Between Conceptual Understanding and Procedural Skills: An Example from Algebra.” Vital Directions for Mathematics Education Research, 2013, pp. 153-171, https://doi.org/10.1007/978-1-4614-6977-3_7.

Kim, Jihoon, and Darla M. Castelli. “Effects of Gamification on Behavioral Change in Education: A Meta-Analysis.” International Journal of Environmental Research and Public Health, vol. 18, no. 7, 2021, article 3550, https://doi.org/10.3390/ijerph18073550.

Kirschner, Paul A., et al. “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, vol. 41, no. 2, 2006, pp. 75-86, https://doi.org/10.1207/s15326985ep4102_1.

Kluger, Avraham N., and Angelo DeNisi. “Feedback Interventions.” Current Directions in Psychological Science, vol. 7, no. 3, 1998, pp. 67-72, https://doi.org/10.1111/1467-8721.ep10772989.

Kluger, Avraham N., and Angelo DeNisi. “The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory.” Psychological Bulletin, vol. 119, no. 2, 1996, pp. 254–284. https://doi.org/10.1037/0033-2909.119.2.254.

Koster, M., et al. “Teaching children to write: A meta-analysis of writing intervention research.” Journal of Writing Research, vol. 7, no. 2, 2015, pp. 249-274, https://doi.org/10.17239/jowr-2015.07.02.2.

Lancaster, Zak, and Andrea R. Olinger. Teaching Grammar-in-Context in College Writing Instruction: An Update on the Research Literature. WPA-CompPile Research Bibliographies, no. 24, Apr. 2014. https://ir.library.louisville.edu/faculty/374/.

Leppink, Jimmie, et al. “Self-explanation in the domain of statistics: an expertise reversal effect.” Higher Education, vol. 63, no. 6, 2012, pp. 771-785, https://doi.org/10.1007/s10734-011-9476-1.

Logue, A. W., et al. “Should Students Assessed as Needing Remedial Mathematics Take College-Level Quantitative Courses Instead? A Randomized Controlled Trial.” Educational Evaluation and Policy Analysis, vol. 38, no. 3, 2016, pp. 578–598. https://doi.org/10.3102/0162373716649056.

McEldoon, Katherine L., et al. “Is self-explanation worth the time? A comparison to additional practice.” British Journal of Educational Psychology, vol. 83, no. 4, 2013, pp. 615-632, https://doi.org/10.1111/j.2044-8279.2012.02083.x.

McGill, Tanya J., and Simone E. Volet. “A Conceptual Framework for Analyzing Students’ Knowledge of Programming.” Journal of Research on Computing in Education, vol. 29, no. 3, 1997, pp. 276-297, https://doi.org/10.1080/08886504.1997.10782199.

Metcalfe, Janet. “Learning from Errors.” Annual Review of Psychology, vol. 68, no. 1, 2017, pp. 465-489, https://doi.org/10.1146/annurev-psych-010416-044022.

Miller, Trey, et al. “Assessing the Effect of Corequisite English Instruction Using a Randomized Controlled Trial.” Journal of Research on Educational Effectiveness, vol. 15, no. 1, 2022, pp. 78-102, https://doi.org/10.1080/19345747.2021.1932000.

Mokher, Christine G., et al. “Does the Method of Acceleration Matter? Exploring the Likelihood of College Coursetaking Success Across Four Developmental Education Instructional Strategies.” Teachers College Record: The Voice of Scholarship in Education, vol. 123, no. 9, 2021, pp. 3-27, https://doi.org/10.1177/01614681211052000.

Moulaert, Véronique, et al. “The effects of deliberate practice in undergraduate medical education.” Medical Education, vol. 38, no. 10, 2004, pp. 1044-1052, https://doi.org/10.1111/j.1365-2929.2004.01954.x.

Newman, Mark, et al. The Impact of Feedback on Student Attainment: A Systematic Review. Education Endowment Foundation, 2021. https://discovery.ucl.ac.uk/id/eprint/10138571/.

Nilsson, Nils J. “Knowledge Representation and Reasoning.” The Quest for Artificial Intelligence, Cambridge University Press, 2009, pp. 149–161. https://www.cambridge.org/core/books/abs/quest-for-artificial-intelligence/knowledge-representation-and-reasoning/F3AC72728DD18F040F4CA297B172ABE5.

Österman, Tove, and Kajsa Bråting. “Dewey and mathematical practice: revisiting the distinction between procedural and conceptual knowledge.” Journal of Curriculum Studies, vol. 51, no. 4, 2019, pp. 457-470, https://doi.org/10.1080/00220272.2019.1594388.

Pereira, Priyanka, et al. “A meta-analysis of the effect of peer feedback on academic achievement in the STEM fields of higher education.” Towards a new future in engineering education, new scenarios that european alliances of tech universities open up, 2022, pp. 2138-2145, https://doi.org/10.5821/conference-9788412322262.1287.

Perkins, D. N., and Gavriel Salomon. “Transfer and Teaching Thinking.” Thinking: The Second International Conference, 2018, pp. 285-303, https://doi.org/10.4324/9781315802015-22.

Perkins, David N., and Gavriel Salomon. “Transfer of Learning.” Manuscript prepared for the International Encyclopedia of Education, 2nd ed., 2 Sept. 1992. https://openlab.citytech.cuny.edu/fywpd/files/2020/01/transferoflearning.pdf.

Plass, Jan L., et al. “Foundations of Game-Based Learning.” Educational Psychologist, vol. 50, no. 4, 2015, pp. 258-283, https://doi.org/10.1080/00461520.2015.1122533.

Ran, Florence Xiaotao, and Hojung Lee. “Does Corequisite Remediation Work for Everyone? An Exploration of Heterogeneous Effects and Mechanisms.” Education Finance and Policy, vol. 20, no. 4, 2025, pp. 534-561, https://doi.org/10.1162/edfp_a_00451.

Ran, Florence Xiaotao, and Yuxin Lin. “The Effects of Corequisite Remediation: Evidence From a Statewide Reform in Tennessee.” Educational Evaluation and Policy Analysis, vol. 44, no. 3, 2022, pp. 458-484, https://doi.org/10.3102/01623737211070836.

Renkl, Alexander, and Robert K. Atkinson. “Structuring the Transition From Example Study to Problem Solving in Cognitive Skill Acquisition: A Cognitive Load Perspective.” Educational Psychologist, vol. 38, no. 1, 2003, pp. 15-22, https://doi.org/10.1207/S15326985EP3801_3.

Renkl, Alexander. “The Worked-Out Examples Principle in Multimedia Learning.” The Cambridge Handbook of Multimedia Learning, edited by Richard E. Mayer, Cambridge University Press, 2005, pp. 229–246. https://doi.org/10.1017/CBO9780511816819.016.

Rittle-Johnson, Bethany, and Abbey M. Loehr. “Eliciting explanations: Constraints on when self-explanation aids learning.” Psychonomic Bulletin & Review, vol. 24, no. 5, 2017, pp. 1501-1510, https://doi.org/10.3758/s13423-016-1079-5.

Rittle-Johnson, Bethany, and Michael Schneider. “Developing Conceptual and Procedural Knowledge of Mathematics.” The Oxford Handbook of Numerical Cognition, Oxford University Press, 2015, pp. 1118–1134. https://academic.oup.com/edited-volume/34494/chapter/292696975.

Rittle-Johnson, Bethany, et al. “Developing conceptual understanding and procedural skill in mathematics: An iterative process.” Journal of Educational Psychology, vol. 93, no. 2, 2001, pp. 346-362, https://doi.org/10.1037/0022-0663.93.2.346.

Rittle-Johnson, Bethany, et al. “Not a One-Way Street: Bidirectional Relations Between Procedural and Conceptual Knowledge of Mathematics.” Educational Psychology Review, vol. 27, no. 4, 2015, pp. 587-597, https://doi.org/10.1007/s10648-015-9302-x.

Rittle-Johnson, Bethany, et al. “Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles.” ZDM, vol. 49, no. 4, 2017, pp. 599-611, https://doi.org/10.1007/s11858-017-0834-z.

Rittle-Johnson, Bethany. “Encouraging students to explain their ideas when learning mathematics: A psychological perspective.” The Journal of Mathematical Behavior, vol. 76, 2024, article 101192, https://doi.org/10.1016/j.jmathb.2024.101192.

Rittle-Johnson, Bethany. “Promoting Transfer: Effects of Self-Explanation and Direct Instruction.” Child Development, vol. 77, no. 1, 2006, pp. 1-15, https://doi.org/10.1111/j.1467-8624.2006.00852.x.

Ritzhaupt, Albert D., et al. “A meta-analysis on the influence of gamification in formal educational settings on affective and behavioral outcomes.” Educational Technology Research and Development, vol. 69, no. 5, 2021, pp. 2493-2522, https://doi.org/10.1007/s11423-021-10036-1.

Roediger, Henry L., and Andrew C. Butler. “The critical role of retrieval practice in long-term retention.” Trends in Cognitive Sciences, vol. 15, no. 1, 2011, pp. 20-27, https://doi.org/10.1016/j.tics.2010.09.003.

Roediger, Henry L., and Mary A. Pyc. “Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice.” Journal of Applied Research in Memory and Cognition, vol. 1, no. 4, 2012, pp. 242-248, https://doi.org/10.1016/j.jarmac.2012.09.002.

Rohrer, Doug, and Harold Pashler. “Recent Research on Human Learning Challenges Conventional Instructional Strategies.” Educational Researcher, vol. 39, no. 5, 2010, pp. 406-412, https://doi.org/10.3102/0013189x10374770.

Rohrer, Doug, et al. “Interleaved practice improves mathematics learning.” Journal of Educational Psychology, vol. 107, no. 3, 2015, pp. 900-908, https://doi.org/10.1037/edu0000001.

Rowe, Frances Anne. A Meta-Analysis of Self-Regulated Learning Interventions and Learning Outcomes in Higher Education E-Learning Environments. 2023. Durham University, doctoral thesis. https://etheses.durham.ac.uk/id/eprint/15132/.

Ruitenburg, Sterre K., et al. “Retrieval-, Distributed-, and Interleaved Practice in the Classroom: A Systematic Review.” EARLI 2021, 23–27 Aug. 2021, Gothenburg, Sweden. Conference presentation. https://research.ou.nl/en/publications/retrieval-distributed-and-interleaved-practice-in-the-classroom-a/.

Ryan, Robert S., and James A. Koppenhofer. “Prompted Self-Explanations Improve Learning in Statistics but Not Retention.” Teaching of Psychology, vol. 51, no. 4, 2024, pp. 402-413, https://doi.org/10.1177/00986283221114196.

Ryu, Wonsun, et al. “Constructing Corequisites: How Community Colleges Structure Corequisite Math Coursework and the Implications for Student Success.” AERA Open, vol. 8, 2022, https://doi.org/10.1177/23328584221086664.

Sailer, Michael, and Lisa Homner. “The Gamification of Learning: a Meta-analysis.” Educational Psychology Review, vol. 32, no. 1, 2020, pp. 77-112, https://doi.org/10.1007/s10648-019-09498-w.

Salmerón, Ladislao. “Actividades que promueven la transferencia de los aprendizajes: Una revisión de la literatura” [Activities That Promote Transfer of Learning: A Review of the Literature]. Revista de Educación, special issue, 2013, pp. 34–53. https://www.uv.es/lasalgon/papers/2013_salmeron_transferencia.pdf.

Samani, Joshua, and Steven C. Pan. “Interleaved practice enhances memory and problem-solving ability in undergraduate physics.” npj Science of Learning, vol. 6, no. 1, 2021, https://doi.org/10.1038/s41539-021-00110-x.

Schudde, Lauren, and Katherine Keisler. “The Relationship Between Accelerated Dev-Ed Coursework and Early College Milestones: Examining College Momentum in a Reformed Mathematics Pathway.” AERA Open, vol. 5, no. 1, 2019, https://doi.org/10.1177/2332858419829435.

Schwonke, Rolf, et al. “The worked-example effect: Not an artefact of lousy control conditions.” Computers in Human Behavior, vol. 25, no. 2, 2009, pp. 258-266, https://doi.org/10.1016/j.chb.2008.12.011.

Sitzman, Danielle M., et al. “Prior knowledge is more predictive of error correction than subjective confidence.” Memory & Cognition, vol. 42, no. 1, 2014, pp. 84-96, https://doi.org/10.3758/s13421-013-0344-3.

Smith, Amy D. “Relationship between Required Corequisite Learning and Success in College Algebra.” Georgia Journal of College Student Affairs, vol. 35, no. 1, 2019, https://doi.org/10.20429/gcpa.2019.350103.

Sterner, Erik O., et al. “Knowing how and knowing when: unpacking public understanding of atmospheric CO2 accumulation.” Climatic Change, vol. 154, no. 1-2, 2019, pp. 49-67, https://doi.org/10.1007/s10584-019-02423-8.

Sweller, John, and Graham A. Cooper. “The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra.” Cognition and Instruction, vol. 2, no. 1, 1985, pp. 59-89, https://doi.org/10.1207/s1532690xci0201_3.

Sweller, John, and Paul Ayres. “Worked Examples.” The Oxford Handbook of Educational Psychology, Oxford University Press, 2022, pp. 331–356. https://academic.oup.com/edited-volume/35466/chapter/371820471.

Tetzlaff, Leonard, et al. “A cornerstone of adaptivity – A meta-analysis of the expertise reversal effect.” Learning and Instruction, vol. 98, 2025, article 102142, https://doi.org/10.1016/j.learninstruc.2025.102142.

Valentine, Jeffrey C., et al. “What Happens to Students Placed Into Developmental Education? A Meta-Analysis of Regression Discontinuity Studies.” Review of Educational Research, vol. 87, no. 4, 2017, pp. 806-833, https://doi.org/10.3102/0034654317709237.

van Gog, Tamara, et al. “Process-Oriented Worked Examples: Improving Transfer Performance Through Enhanced Understanding.” Instructional Science, vol. 32, no. 1-2, 2004, pp. 83-98, https://doi.org/10.1023/B:TRUC.0000021810.70784.b0.

van Loon, Mariëtte H., et al. “Refutations in science texts lead to hypercorrection of misconceptions held with high confidence.” Contemporary Educational Psychology, vol. 42, 2015, pp. 39-48, https://doi.org/10.1016/j.cedpsych.2015.04.003.

Van Merriënboer, Jeroen J G, and John Sweller. “Cognitive load theory in health professional education: design principles and strategies.” Medical Education, vol. 44, no. 1, 2010, pp. 85-93, https://doi.org/10.1111/j.1365-2923.2009.03498.x.

van Velzen, Joke. “Metacognitive Knowledge in Theory.” Metacognitive Learning, 2016, pp. 13-25, https://doi.org/10.1007/978-3-319-24433-4_2.

Vermeir, Julie F., et al. “The Effects of Gamification on Computerized Cognitive Training: Systematic Review and Meta-Analysis.” JMIR Serious Games, vol. 8, no. 3, 2020, article e18644. https://doi.org/10.2196/18644.

Wisniewski, Benedikt, et al. “The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research.” Frontiers in Psychology, vol. 10, 2020, https://doi.org/10.3389/fpsyg.2019.03087.

Wittwer, Jörg, and Alexander Renkl. “How Effective are Instructional Explanations in Example-Based Learning? A Meta-Analytic Review.” Educational Psychology Review, vol. 22, no. 4, 2010, pp. 393-409, https://doi.org/10.1007/s10648-010-9136-5.

Wong, Regina M.F, et al. “The effects of self-explanation training on students' problem solving in high-school mathematics.” Learning and Instruction, vol. 12, no. 2, 2002, pp. 233-262, https://doi.org/10.1016/s0959-4752(01)00027-5.

Zepeda, Cristina D, et al. “The Mnemonic Effects of Retrieval Practice.” Oxford Research Encyclopedia of Education, 2024, https://doi.org/10.1093/acrefore/9780190264093.013.858.

Zhang, Qi, and Zhonggen Yu. “Meta-Analysis on Investigating and Comparing the Effects on Learning Achievement and Motivation for Gamification and Game-Based Learning.” Education Research International, vol. 2022, 2022, pp. 1-19, https://doi.org/10.1155/2022/1519880.