EDUCATIONAL PSYCHOLOGIST
https://doi.org/10.1080/00461520.2022.2149525

An integrated model of learning from errors

Qian Zhang and Logan Fiorella

Department of Educational Psychology, University of Georgia

ABSTRACT
Errors are inevitable in most learning contexts, but under the right conditions, they can be beneficial for learning. Prior research indicates that generating and learning from errors can promote retention of knowledge, higher-level learning, and self-regulation. The present review proposes an integrated theoretical model to explain two major phases of learning from self-generated errors: the Generating Errors (GE) phase, which contributes to learning via semantically related prior knowledge activation, and the Detecting and Correcting Errors (DCE) phase, which contributes to learning via self-explanation when processing and comparing one's responses with provided reference information to promote high-quality internal feedback. Our model identifies general design principles that support each phase based on prior empirical research. We conclude by identifying research gaps and future directions regarding specific design features of the GE and DCE phases and the role of students' emotion, motivation, and individual differences in learning from errors.

Errors are behaviors or responses to learning tasks that are different from learners' subjective expectations or a provided, objective reference criterion (Simpson et al., 2020). Conventional instruction typically aims to avoid or minimize student errors. The rationale for this approach is that if the goal is to achieve errorless performance, students should learn in an errorless environment where only correct behaviors are encouraged, only correct information is presented, and errors are ignored (Ausubel, 1968; Skinner, 1958; Terrace, 1963). Although an error-prevention approach might work well in some learning contexts, it has two major limitations. First, it is *impractical* because errors are inevitable in most learning contexts: They are natural by-products of attempting challenging learning tasks. Second, it is potentially *suboptimal* because errors can provide valuable opportunities for learning, in line with constructivist or generative views of learning. For example, in Piaget's (1952) theory of cognitive development, learning is triggered by a mental state of disequilibrium, in which children get stuck or experience contradictions (e.g., errors) in solving problems as a result of insufficient or inaccurate knowledge. Children resolve this disequilibrium by assimilating new knowledge into their existing schema or accommodating conflicting information. Similarly, VanLehn's (1988) impasse-driven learning theory proposed that when students get stuck in problem-solving, they are more aware of their deficient knowledge and motivated to resolve the impasse. Instruction designed to help students overcome the impasse can update students' knowledge and improve learning. In short, students are likely to make errors, and with proper support, errors can catalyze new learning.

In this review, we explore existing theoretical and empirical research concerning what and how students learn from errors. Our goal is not to compare the effects of errorless learning versus errorful learning. Rather, given the inevitability of errors, we focus on how student errors can be leveraged to maximize learning, whether errors are merely permitted or actively promoted (Wong & Lim, 2019). Although errors have the potential to initiate learning, many students struggle to learn well from errors. Moreover, the empirical research on errors is often disconnected and not clearly situated within a common theoretical model. Therefore, the first goal of this review is to synthesize representative research related to learning from errors in cognitive psychology and educational psychology to understand what and how students learn from errors. Specifically, we focus on how making and correcting errors affect (a) retention of knowledge, (b) higher-level learning, and (c) self-regulatory knowledge and skills. The second goal is to propose a theoretical model based on our synthesis of past theories and empirical research to explain the process of learning from errors. Then we use this model to derive two principles on how to support students to learn from errors by interpreting prior findings of instructional support for learning from errors. Finally, we discuss knowns and unknowns around the model to inform recommendations for practice and future research.

CONTACT Qian Zhang qz89230@uga.edu; Logan Fiorella lfiorella@uga.edu Department of Educational Psychology, University of Georgia, Athens, GA 30602-0002, USA.

What and how do students learn from errors?

Learning from errors to improve retention of knowledge

Retention is the ability to reproduce information presented during instruction at a later time (Anderson & Bloom, 2001). Retention alone is often considered lower-level rote learning, yet it serves as an important precondition for meaningful learning, which involves higher-order thinking such as knowledge comprehension, application, and transfer. Evidence that learning from errors improves retention of knowledge mainly comes from memory research in cognitive psychology, involving studies where students generate errors before or after initial learning of the target knowledge (i.e., in a pretest or a post-test). These studies mostly use cued recall or cued recognition tasks to elicit students' memory errors and measure learning. For instance, when using simple word pairs as learning materials (e.g., pond-frog), a cued recall task asks students to generate the target word to a given cue word (e.g., pond – ___), and learning is measured through the same cued recall task in a final test (e.g., Kornell et al., 2009, Exp 3–6). Alternatively, an error-eliciting task may use questions as cues for retrieving factual knowledge, such as trivia questions (e.g., "What is the only word the raven says in Edgar Allen Poe's poem 'The Raven'?"), and learning is measured through performance on the same questions in a final test (e.g., Kornell et al., 2009, Exp 1–2).

It is well-established that *successful* retrieval of information strengthens one's memory and slows forgetting of the retrieved information (i.e., the testing effect or retrieval-based learning; Roediger & Karpicke, 2006). Surprisingly, a growing body of memory research has shown that *unsuccessful* retrieval of information (i.e., making memory errors) followed by the target information as feedback also improves memory of the target information (Metcalfe, 2017). For example, research on the pretesting effect found that asking students to generate answers to questions *before* knowing anything about the questions (e.g., a pretest of the meaning of Euskara nouns) significantly improved their memory for the tested information, even though students' answers in the pretest were almost always wrong (Potts & Shanks, 2014). Similarly, research on retrieval failure *after* initial learning of target information showed that students retained more information from the initial learning phase when they took tests after learning and made memory errors followed by correct answers as feedback, compared to taking tests and then only restudying the information (e.g., Arnold & McDermott, 2013a; Kornell & Vaughn, 2016). Moreover, after receiving feedback, students are more likely to correct errors in questions they rated with high confidence compared to low-confidence errors (i.e., the hypercorrection effect, Butterfield & Metcalfe, 2006).

How does learning from errors facilitate retention of knowledge?

Theories of retrieval-based learning can be useful to explain the mechanisms of learning from errors because memory errors are synonymous with *unsuccessful* retrieval attempts. Kornell et al. (2015) and Kornell and Vaughn (2016) proposed a two-stage theoretical model to describe the process of learning from retrieval. In stage 1, people engage in retrieval attempts to search for the information they need to complete a learning task. Once people identify and retrieve the information as a response to the task, they enter stage 2 where people process the retrieved information (i.e., their responses) and/or externally provided feedback. The post-retrieval processing in stage 2 strengthens or weakens associations among cues in the task, retrieved responses, and information in external feedback, depending on whether the responses are consistent with the feedback. In this sense, learning from errors is one special case of learning from retrieval, in which one's retrieved responses are inconsistent with feedback. Prior research explaining how making errors facilitates memory generally focuses on one of the two stages of Kornell et al.'s model: What happens during retrieval (Stage 1) or what happens during post-retrieval processing (Stage 2) (Zawadzka & Hanczakowski, 2019).

First, attempting to retrieve information from memory (i.e., Stage 1) promotes active search or passive spreading activation of semantic networks, thereby activating a set of concepts associated with cues in a task. Even though one's final response is incorrect (i.e., people generate a memory error), the underlying activation processes during retrieval attempts create a fertile ground to facilitate the encoding of correct answers in the subsequent feedback or learning materials (e.g., the search set theory, Grimaldi & Karpicke, 2012; reconsolidation theory, Metcalfe & Xu, 2018; the elaborative retrieval theory, Carpenter, 2009; Richland et al., 2009). Accordingly, the retrieved erroneous information and associations might be weakened while the correct information and associations are strengthened. There is also evidence that retrieval attempts can improve the structural organization of previously learned knowledge, which is another way to facilitate the encoding of information in subsequent learning materials (Arnold & McDermott, 2013a). Moreover, when taking a future test, one's episodic memory of generating and correcting errors can serve as additional cues to improve retention performance (i.e., recursive reminding theory, Metcalfe & Huelser, 2020).

Next, because retrieval is more effortful than passive studying, retrieval attempts will increase one's attention to the information provided after retrieval, such as feedback or other learning materials (i.e., Stage 2; Potts & Shanks, 2014; Seabrooke, Mitchell, et al., 2019). This effect on attention is especially salient in the pretesting effect using complex, educational learning materials: Students retain pretested information better than untested information, suggesting that pretests direct students' attention to tested information in the learning materials (Toftness et al., 2018). In addition, when people realize their retrieved responses are wrong, the awareness of their knowledge deficiency can trigger emotions such as surprise and curiosity, which further boost attention and motivation to learn (Butterfield & Metcalfe, 2006; Carpenter & Toftness, 2017; Potts et al., 2019).

Additionally, recent research examined different measures of memory and suggested more nuanced mechanisms involved in the two stages, depending on how people's responses associate with cues in learning tasks as well as targets in provided feedback. For instance, measures of memory can be broken down into memory for targets only (in a cued recognition task) and memory for cue-to-target associations (in a cued recall task). A series of studies by Seabrooke, Mitchell, et al. (2019), Seabrooke, Hollins, et al. (2019), Seabrooke et al. (2021) and Zawadzka and Hanczakowski (2019) showed that generating errors helped strengthen memory for cue-to-target associations when people's responses associated with both cues and targets, consistent with findings in past research that generating errors had benefits for related word pairs but not for unrelated word pairs in *cued recall* tests (e.g., Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012). In contrast, when people's responses were associated with cues but not targets or when people's responses were associated with neither cues nor targets, generating errors had an impact on memory for targets only, consistent with findings in past research that generating errors had benefits for novel vocabulary learning in *recognition* tests (e.g., Potts & Shanks, 2014; Potts et al., 2019). Furthermore, Clark et al. (2021) showed that error generation benefited memory when people's responses were *semantically* associated with cues and/or targets but not when their responses were *phonologically* associated with cues and/or targets.

Taken together, these findings suggest that the semantic relationship between errors and targets determines the extent to which generating errors benefits memory. It seems that both activation of semantic networks and increased attention through motivation and/or emotion contribute to memory for cue-to-target associations when people generate errors *semantically* associated with targets. When people's errors are not semantically associated with targets, increased attention through motivation and/or emotion can still contribute to recall for targets only, but that attention is not beneficial for building cue-to-target associations.

Learning from errors to improve higher-level learning

Learning from errors can also facilitate higher-level learning such as conceptual understanding, knowledge application, and transfer, which are often the primary goals of authentic academic tasks (Anderson & Bloom, 2001). These learning outcomes reflect the ability to apply one's knowledge flexibly, which is especially challenging in STEM education (Clement, 2000; Fiorella & Mayer, 2016; Rittle-Johnson et al., 2001). Evidence of learning from errors facilitating higher-level learning mostly comes from research in educational psychology (e.g., Loibl et al., 2017). Researchers typically use complex problem-solving tasks to elicit students' errors and measure learning through tests of comprehension, knowledge application, conceptual understanding, and knowledge transfer.

As in memory research, students can improve higher-level learning by generating errors before or after initial learning of to-be-learned information. Research on problem-solving *before* direct instruction (PS-I, Loibl et al., 2017) suggests permitting or promoting errors during initial problem-solving can be superior for conceptual learning to the traditional tell-and-practice (instruction-first, or I-PS) method (e.g., Kapur, 2010, 2012, 2014; Kapur & Bielaczyc, 2012; Schwartz et al., 2011; for a review see Loibl et al., 2017). Generally, PS-I research includes studies of inventing to prepare for future learning (Schwartz & Martin, 2004) and productive failure (Kapur, 2016). Students may generate errors in the inventing or problem-solving phase (i.e., the PS phase) before formal instruction or in a practice session of knowledge application during the formal instruction phase (i.e., the I phase). However, PS-I researchers typically focus on the benefits of students' failure and errors in the PS phase. Specifically, researchers of inventing to prepare for future learning typically ask students to invent a method, index, or formula and examine it with provided data *before* instructions on relevant concepts and expert solutions (Schwartz & Martin, 2004).[1] Similarly, productive failure researchers provide students with challenging and rich problems that enable students to generate diverse representations and solutions *before* instructions of canonical solutions and relevant concepts (Kapur & Bielaczyc, 2012).

Another line of research focuses more on students' errors *after* initial learning in formal instructions. For example, students can benefit from self-assessment/diagnosis activities where they learn from errors by grading and reflecting on their learning products, such as assignments, exams, or projects after class. Researchers generally define self-assessment as one form of formative assessment to improve learning through feedback generated from comparing one's competency, learning process, or learning product to internal/external standards (Andrade, 2018; Panadero et al., 2019). Two recent reviews of self-assessment research showed that self-assessment activities can improve students' *subsequent* academic performance including conceptual understanding, knowledge application, and knowledge transfer (Andrade, 2019; Sanchez et al., 2017). Self-assessment researchers in non-STEM education mainly use report writing, essay writing, and oral presentations to elicit students' errors (e.g., Andrade et al., 2008), whereas self-assessment researchers in STEM education mainly use problem-solving tasks (e.g., Safadi & Saadi, 2021). Problem-solving tasks typically require students to apply what they know to solve familiar or unfamiliar problems in multiple-choice questions, word problems, computational problems, and/or open-ended conceptual questions. A few studies also included drawing tasks to elicit students' errors, although tasks using visual representations are relatively rare in research on learning from errors (e.g., Jax et al., 2019; Loibl & Leuders, 2018, 2019; Zamora et al., 2018a). Basic activities in self-assessment include simply asking students to self-evaluate, grade, or

[1]Researchers of inventing to prepare for future learning put less emphasis on students' failure in inventing activities. In many studies that show advantages of inventing before instruction, students performed quite well in the inventing activities (e.g., Chin et al., 2016; Lamnina & Chase, 2019; Loehr et al., 2014; Schalk et al., 2018).

reflect on their works or performance, but students often self-assess poorly and do not learn from errors (Andrade, 2019). Researchers agree that instructional support is necessary for self-assessment to be accurate and effective to improve learning (Andrade & Valtcheva, 2009; Brown & Harris, 2013; Nicol & Macfarlane-Dick, 2006; Panadero et al., 2017; Sanchez et al., 2017). Past research has found effective support such as self-assessment training through modeling (Kostons et al., 2012; Raaijmakers et al., 2018), providing students with detailed/step-by-step reference criteria (Panadero et al., 2013; Safadi & Saadi, 2021), and providing teachers' feedback on students' self-assessment (Andrade et al., 2010).

How does learning from errors facilitate higher-level learning?

A review of PS-I studies by Loibl et al. (2017) found three major cognitive mechanisms of PS-I, which inform how generating errors *before* formal instruction may facilitate higher-level learning. First, exploratory activities before formal instruction help students activate and differentiate their prior knowledge to support knowledge integration during subsequent instruction, even though students' limited knowledge will lead them to errors in exploratory activities. For example, productive failure studies have shown that students generated diverse representations and solutions during exploratory problem-solving before instruction. Although students mostly failed to generate canonical solutions, there was a significant positive correlation between the number of solutions students generated and their conceptual understanding and transfer performance after instructions (DeCaro & Rittle-Johnson, 2012; Kapur, 2012; Kapur & Bielaczyc, 2012; Kapur, 2010).

Second, PS-I helps students become more aware of their knowledge gaps. Loibl and Rummel (2014a) found that students who solved problems before instruction perceived more knowledge gaps than students who received instruction first. Such awareness and identification of knowledge gaps is important because students need to realize flaws in their knowledge and skills before they can repair them (Chi, 2000). Moreover, some PS-I design features can help students identify knowledge gaps more accurately and effectively, such as by using contrasting cases in the problem-solving phase or building knowledge on students' common errors in the instruction phase (Loibl & Rummel, 2014b; Roll et al., 2012).

Finally, carefully designed activities in PS-I direct students' attention to critical features of concepts, which enable students to perform well in tests of conceptual understanding and knowledge transfer. For instance, Schwartz et al. (2004, 2011) found students who invented formulas using contrasting cases recalled the deep structure of learning materials better than students taught by the conventional tell-and-practice method. Additionally, PS-I may benefit learning through affective and motivational factors. Students reported higher curiosity and tended to adopt mastery-oriented learning goals in PS-I, which are important predictors of academic performance (Belenky & Nokes-Malach, 2012; Elliot, 1999; Glogger-Frey et al., 2015; Lamnina & Chase, 2019; Loibl & Rummel, 2014b).

One can find similar explanations of learning from errors in self-assessment research, in which students learn from errors *after* instruction. Self-assessment may trigger students' awareness of knowledge gaps, increase attention to the knowledge presented in assessment criteria, and develop positive motivational beliefs such as a growth mindset and intrinsic motivation (Nicol & Macfarlane-Dick, 2006; Safadi, 2018; Sanchez et al., 2017). Yan and Brown (2017) proposed a model of students' actions in self-assessment based on interviews with undergraduate students. The model suggests that self-assessment/evaluation encourages students to generate reflective self-feedback or *internal feedback* on their understanding of content knowledge and self-regulatory knowledge (Butler & Winne, 1995; Nicol, 2020). Specifically, internal feedback consists of self-generated information about one's feelings, current knowledge, learning process, and learning products based on how one's performance compares to some reference criteria, such as personal goals or rubrics (Butler & Winne, 1995; Nicol, 2020). Students can generate internal feedback on their understanding of content knowledge as well as self-regulatory knowledge and skills via a process of *self-explaining*, in which they fill knowledge gaps and/or repair flaws in their mental models (Butler & Winne, 1995; Chi, 2000; Nicol, 2020). Moreover, internal feedback on self-regulatory knowledge and skills (e.g., regulation of motivation, emotion, and learning strategies) initiates a subsequent cycle of *self-regulation* to improve learning by closing gaps between students' current performance and learning goals (Andrade, 2019; Andrade & Valtcheva, 2009; Kostons et al., 2012; Panadero et al., 2019; Panadero & Romero, 2014). Thus, internal feedback generated from self-assessment activities has the potential to improve learning through a direct process of self-explaining and/or an indirect process of self-regulation following self-explaining.

External and internal feedback play distinct but complementary roles in learning from errors. On one hand, external feedback is necessary to initiate learning from errors: Only when students are aware of errors after receiving external feedback can they learn from errors. On the other hand, internal feedback is the proximate cause of learning from errors, in line with the constructive view of feedback (Carless, 2019; Mory, 2004; Nicol, 2020; Van der Kleij et al., 2019). Importantly, regardless of the presence or absence of external feedback, students continually generate internal feedback to update their content knowledge and/or self-regulatory knowledge and skills to improve future performance (Andrade, 2009; Brown & Harris, 2014; Butler & Winne, 1995; Mory, 2004; Nicol, 2020).

Learning from errors to develop self-regulated learning knowledge and skills

Beyond the learning of content knowledge, research suggests that self-assessing and diagnosing one's assignments, exams, or projects based on provided reference information can improve students' self-regulated learning (SRL) knowledge and skills (Butler & Winne, 1995). Although not equal to self-assessment, learning from errors is an inherent self-assessment/diagnosis context when students make errors in their initial performance. Therefore, evidence from self-assessment research on SRL naturally applies to the specific context of learning from errors.

SRL is a cyclical process, in which students set goals and make plans (i.e., forethought phase), metacognitively monitor the progress and regulate their cognition, motivation, and emotion (i.e., performance phase) and adjust knowledge or behaviors based on self-evaluation (i.e., self-reflection phase; Zimmerman, 2000). SRL knowledge and skills are not only closely related to academic achievement in school but also prepare students to become independent and lifelong learners outside of school (Jansen et al., 2019; Wang & Sperling, 2020). Panadero et al. (2017) reported in a meta-analysis that self-assessment interventions with appropriate instructional support had a positive effect on students' SRL knowledge and skills, as measured through questionnaires and think-aloud protocols. For instance, Panadero et al. (2012) counted the number of self-regulation propositions in secondary school students' think-aloud protocols when students self-assessed their landscapes analysis report in a Geography class. Self-regulation propositions included negative emotional self-regulations, planning, help-seeking, and questions for clarification. They found that students who self-assessed while using provided scripts and rubrics expressed significantly more self-regulation propositions compared with students who self-assessed without scripts.

Self-assessment can also improve more specific components of SRL knowledge and skills, such as self-efficacy motivational beliefs and self-monitoring metacognitive skills. For example, undergraduate students who self-assessed their learning by comparing project outcomes with provided criteria and rating the amount of progress they have made developing their computer skills reported significantly higher self-efficacy after class than students who did not self-assess their progress (Schunk & Ertmer, 1999). Furthermore, self-monitoring involves retrospectively assessing how well one performed on a task one just completed (i.e., self-assessment) or predicting how well one will perform in similar tasks in the future (i.e., judgments of learning). Researchers calculate *monitoring accuracy* typically by comparing students' self-assessment and/or judgments of learning with their actual task performance. Research has shown that self-assessment interventions that provided students with objective reference criteria (e.g., normative solutions, rubrics, or scripts) or self-assessment training can improve students' *subsequent* monitoring accuracy on new tasks (Baars et al., 2014; Kostons et al., 2012; Sadler & Good, 2006; Zamora et al., 2018b).[2] Research on monitoring accuracy assumes that accurate self-monitoring is necessary for students to effectively regulate subsequent learning, such as by selecting which learning materials to restudy. However, there is little systematic research to directly support this assumption (Andrade, 2019; Raković et al., 2022).

How does learning from errors facilitate the development of SRL knowledge and skills?

There is no existing theory to explain how learning from errors might facilitate the development of SRL knowledge and skills. However, the link between self-assessment and SRL via *internal feedback* likely plays an important role (Andrade, 2019; Yan & Brown, 2017). For example, when students self-assess their performance against provided criteria and detect performance gaps or errors, they generate internal feedback to clarify learning goals, adjust their goals as necessary, and process/internalize reference criteria, which can facilitate the development of learning goals in similar tasks and goal-setting skills. The comparison between students' performance and reference criteria also enables students to generate internal feedback on their judgments of learning or ability, which can improve their self-monitoring accuracy through calibration. When students reflect on their learning process that leads to errors during self-assessment, they generate internal feedback on the effectiveness of their learning strategies and change learning strategies to improve performance based on self-analysis of errors, which can enrich their knowledge about learning strategies. Therefore, self-assessment contributes to SRL knowledge and skills throughout the three phases of SRL suggested by most models: the preparatory phase, the performance phase, and the appraisal phase (Panadero, 2017; Panadero et al., 2017).

There are other theoretical explanations of how learning from errors through self-assessment might affect students' SRL knowledge and skills. For instance, based on the self-determination theory and mindset theory of motivation, Sanchez et al. (2017) proposed that self-assessment affords a sense of autonomy and emphasizes revision and progress so that students might develop higher intrinsic motivation and a growth mindset for learning. Moreover, self-assessment interventions that include instructional guidance and training can improve students' self-efficacy by increasing their confidence in performing self-assessment as well as the tasks being assessed (Panadero et al., 2017). Finally, self-assessment makes the comparison process explicit and mindful, which is critical for students to effectively learn from comparison and internal feedback (Gentner et al., 2003; Nicol, 2020).

Synthesis of prior research

Taken together, students can learn from self-generated errors before or after initial learning of target knowledge.[3]

[2]Some studies in memory research found no benefits of learning from errors on judgment accuracy (e.g., Huelser & Metcalfe, 2012; Potts & Shanks, 2014) probably because self-assessment studies that reported judgment accuracy benefits provided students with instructional support whereas memory studies did not.

[3]The current state of the evidence suggests the mechanisms proposed and examined by past research may explain the process of learning from errors in both contexts, which is the focus of our model. That said, it is still possible

Table 1. What and how students learn from errors in prior research.

Learning goals	Learning outcome measures	Research topics	Common generating errors tasks	Mechanisms
Retention of knowledge	• Cued recall test (memory for cue-to-target associations) • (Cued) Recognition test (memory for targets only)	• Pretesting effect • Post-test errors followed by correct answers • Hypercorrection effect	• Before or after initial learning • Cued recall task e.g., cued recall word pairs, trivia questions, etc.	• Error generation activates and reorganizes prior knowledge to prepare for subsequent information encoding. • Awareness of errors/incompetency increases/directs attention increases motivation and triggers emotions such as surprise or curiosity. • When errors semantically associate with targets, generating errors benefit memory for cue-to-target associations; when errors are not semantically associated with targets, generating errors benefit memory for targets only. • Episodic memory of generating and correcting errors serves as additional memory cues.
Higher-level learning	• Comprehension test • Conceptual understanding test • Knowledge application test • Knowledge transfer test	• Problem-solving before direct instruction (PS-I): - Inventing to prepare for learning - Productive failure • Self-assessment	• Before or after formal instruction • Problem-solving task • Inventing task	• Error generation activates and differentiates prior knowledge to prepare for subsequent information integration. • Awareness of errors/incompetency increases/directs attention increases motivation and triggers emotions such as surprise or curiosity. • Self-assessment directly affects learning through internal feedback on content knowledge and indirectly affects learning through internal feedback on self-regulated learning knowledge and skills.
SRL knowledge and skills	• Questionnaire of self-regulated learning knowledge and skills • Think-aloud protocol	• Self-assessment	• After formal instruction • Report writing, essay writing, and oral presentations • Problem-solving task	• Awareness of errors/explicit and mindful comparison against reference information encourages internal feedback generation that updates SRL knowledge and skills. • Self-assessment affords a sense of autonomy to increase intrinsic motivation; self-assessment focuses on learning progress to nurture a growth mindset; self-assessment support increases students' confidence to improve self-efficacy.

Depending on learning goals, learning from errors can facilitate retention of knowledge, higher-level learning, and the development of SRL knowledge and skills, as summarized in Table 1. Our synthesis of prior research identifies common and unique mechanisms of learning from errors across separate lines of research that echo and complement one another.

Two common mechanisms surfaced from all lines of research. First, error generation prepares students for subsequent information encoding and integration by activating, reorganizing, and differentiating students' prior knowledge (e.g., Arnold & McDermott, 2013b; Grimaldi & Karpicke, 2012; Loibl et al., 2017). Students must retrieve information from prior knowledge to respond to a learning task, which not only makes prior knowledge accessible but may also create more organized knowledge structures. Thus, students are more likely to better encode and integrate incoming information with activated and organized prior knowledge. Second, error detection encourages students to engage in subsequent learning more actively by triggering students' metacognitive awareness of knowledge deficiencies or incompetence to increase motivation, direct attention, and provoke positive emotions (e.g., Butterfield & Metcalfe, 2006; Loibl & Rummel, 2014b; Sanchez et al., 2017; Seabrooke, Hollins, et al., 2019). Identifying knowledge gaps after making errors may increase students' motivation to learn and provoke emotions such as surprise and curiosity, resulting in higher subsequent engagement with the provided external feedback and the learning materials.

Furthermore, unique mechanisms identified in one area of research may apply to other areas to provide complementary explanations. For example, memory research suggests that when errors *semantically* associate with targets, generating errors benefits memory for cue-to-target associations through both activation of prior knowledge and increased attention, motivation, and/or positive emotions (e.g., Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012). When errors are *not semantically* associated with targets, generating errors benefits memory for targets only through increased attention, motivation, and/or positive emotions (e.g., Clark et al., 2021; Potts et al., 2019). Similar to memory research, PS-I and self-assessment research on higher-level learning has equivalent components as "cue" and "target": Generating errors tasks in PS-I and self-assessment research can be seen as groups of different cues; instructions in PS-I research and provided reference criteria in

that the mechanisms operate uniquely when errors occur before vs. after initial learning.

self-assessment research can be seen as groups of to-be-learned targets. However, different from memory research, measures of higher-level learning inherently require memory for "cue-to-target associations" due to the complexity of higher-level learning tasks. Therefore, the nuanced mechanisms proposed in memory research can provide explanations for finding no benefits of generating errors in higher-level learning when students have too little prior knowledge to generate meaningful responses semantically associated with target knowledge (Loibl & Leuders, 2019).

Also, the central role of internal feedback emphasized by self-assessment research can extend explanations of how learning from errors affects retention of content knowledge and higher-level learning. For instance, both memory research and PS-I research explain learning from errors by focusing on the benefits of generating errors, such as activating prior knowledge and triggering awareness of incompetence to increase motivation or direct attention. However, the two lines of research do not provide more detailed explanations about what happens after generating errors. According to self-assessment research, when comparing erroneous responses with reference criteria, students generate *internal feedback* to update activated prior knowledge and regulate motivation, emotion, and attention to facilitate learning from errors (e.g., Andrade, 2019; Yan & Brown, 2017). Thus, the concept of internal feedback provides more complete and detailed explanations about students' learning process after generating errors.

A theoretical model of learning from self-generated errors

From our synthesis of prior research, we developed a theoretical model of how students learn from self-generated errors to improve content knowledge as well as develop SRL knowledge and skills. Our model not only covers common mechanisms identified across separate lines of research but also links and integrates unique mechanisms from different research areas to provide a coherent and comprehensive picture of learning from errors.

In academic settings, students' behaviors or responses are partly guided by their underlying knowledge and understanding of situations in the moment. We define students' observable, erroneous behaviors or responses as the *surface* error and their underlying knowledge that leads to their erroneous behaviors or responses as the knowledge-based *deep* error, similar to Reason's (1990) taxonomy of errors. For example, when a student fails to solve a physics problem, she may first notice that her responses (results and/or solving processes) were different from the normative answers provided in the textbook (i.e., surface error). If she further reflected on her surface error, she may notice missing pieces and flaws in her knowledge or understanding of some physics concepts applied in the problem (i.e., deep error) that led to her erroneous responses. Each type of error can happen naturally due to students' deficient knowledge, or students can deliberately generate errors when they know the correct answers (e.g., Wong & Lim, 2022). We only consider natural errors resulting from students' deficient knowledge or skills in our model.

As shown in Figure 1, there are two major phases in students' learning from self-generated errors: the Generating Errors (GE) phase and the Detecting & Correcting Errors (DCE) phase. In the GE phase, students activate their prior knowledge when they generate responses to a learning task. Prior knowledge refers to any knowledge in long-term memory students bring with them to the learning task including content knowledge as well as SRL knowledge. However, students may or may not have prior content knowledge of materials in GE tasks in some situations. Specifically, if the GE phase happens *before* initial learning or formal instructions of target content knowledge, students may have little (e.g., Loibl et al., 2017) or no prior content knowledge of materials in the GE tasks (e.g., Potts & Shanks, 2014). Nevertheless, students must activate other prior knowledge to generate responses. Depending on how much/little prior content knowledge students have, they may generate responses that are or are not semantically associated with target content knowledge. In research on higher-level learning, even when students have little prior content knowledge, they may still be able to generate responses semantically associated with target content knowledge due to rich cues and the complexity of knowledge in higher-level learning tasks.

As mentioned above, prior knowledge includes not only content knowledge, but also SRL knowledge such as goals,

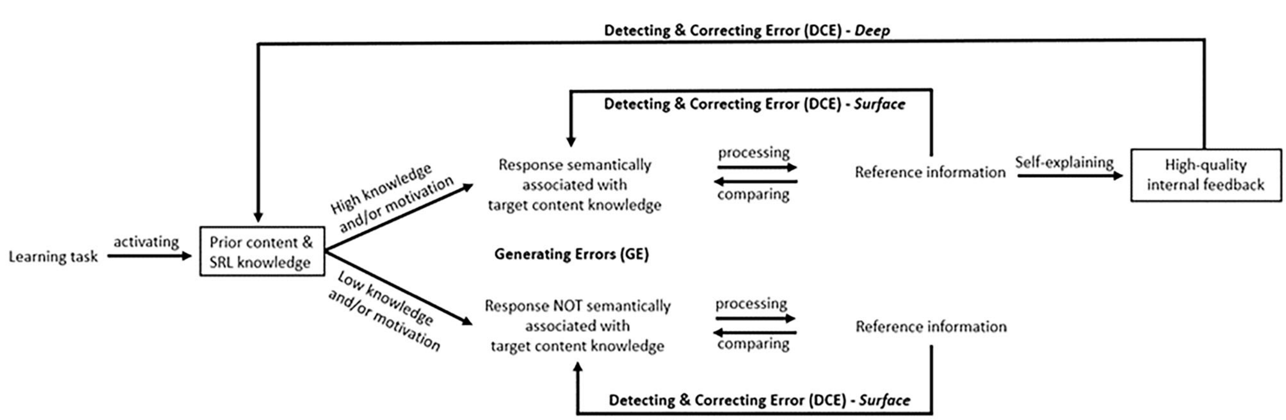

Figure 1. A theoretical model of learning from self-generated errors.

motivational beliefs, emotions, learning strategies, and metacognitive knowledge and skills, all of which will influence how students generate errors in the GE phase and learn from errors in the DCE phase (Tulis et al., 2016). For example, students' motivation may have an impact on whether their responses semantically associate with target content knowledge: When students have difficulty generating responses, lower-motivated students may not invest much mental effort and thus, generate random guesses that are not semantically associated with target content knowledge. In contrast, higher-motivated students may try their best to utilize relevant prior content knowledge and thus, generate meaningful responses semantically associated with target content knowledge. However, empirical research is needed to directly test this hypothesis. Moreover, individual differences in motivation and beliefs about errors can lead to different self-regulation strategies for dealing with errors (e.g., Reindl et al., 2020; Tulis et al., 2018). For instance, students who have a mastery goal orientation and positive beliefs that errors are potential learning opportunities will effectively regulate their negative emotions following errors through mastery self-talk, focusing on mastery goals, and directing their cognitive and metacognitive capabilities toward effective strategies to learn from errors. In contrast, students who have a performance goal orientation and negative beliefs about errors will ruminate about failure and worry about the negative consequences of making errors instead of directing their attention to learning from errors. Finally, the effectiveness of students learning from errors also depends on whether students have sufficient knowledge of effective learning strategies and adequate skills of self-monitoring and self-evaluation. In sum, students' activated SRL knowledge can influence their learning from errors.

In the DCE phase, students have access to subjective or objective reference information. In this model, we only focus on objective reference information provided by an external agent (e.g., teacher, peer, book, computer, etc.). Thus, reference information refers to any externally provided information following students' responses (e.g., external feedback in memory research, explicit instruction in PS-I research, and reference criteria in self-assessment research). It should be noted that external feedback is only one form of reference information.

As shown in Figure 1, there are different learning paths in the DCE phase. When students' responses from the GE phase are semantically associated with target content knowledge, comparing reference information to one's responses should at least trigger students' awareness of surface errors, enabling students to update initial responses by detecting and correcting surface errors. Detecting and correcting deep errors, however, depends on the quality of internal feedback students generate when they make sense of reference information *and* when they compare their responses with reference information. We propose that generating *high-quality* internal feedback to reconstruct prior knowledge requires students to self-explain when processing and comparing with reference information (e.g., Loibl & Leuders, 2018, 2019). This conjecture aligns with theoretical and empirical evidence from cognitive research on comparison and learning from performance errors. One key finding from research on analogical comparison is that although making comparisons is fundamental to human thought, the comparison process must be deliberate and effortful for it to serve as a learning tool (Alfieri et al., 2013; Gentner et al., 2003; Hoyos & Gentner, 2017). Moreover, computer simulation research suggests that people avoid making the same errors (i.e., errors are corrected) only when people specify faulty knowledge structures by explaining why their response is incorrect (Ohlsson, 1996).

When students' responses are not semantically associated with target content knowledge, comparing reference information to one's own responses will trigger students' awareness of *surface* errors to help them detect and correct *surface* errors. However, self-explaining during such comparisons is unlikely because no semantic relationships exist between students' responses and target content knowledge, and students generate self-explanations based on what they already know (Chi, 2000). This lack of self-explanation in processing and comparing with reference information makes it difficult for students to update prior content knowledge and SRL knowledge because no relevant prior knowledge in long-term memory is activated in the GE phase to facilitate knowledge integration. This updating process, if it happens at all, requires students to encode novel knowledge from scratch and has only limited benefits on recognition memory for target knowledge (e.g., Clark et al., 2021; Potts et al., 2019).

The present model lays out the process of how students can potentially learn from errors, but many students need support during this process to learn well from errors. The model implies that one can support students' learning from errors by designing effective learning tasks and activities in the GE phase and/or DCE phase. For example, prior knowledge is always activated to some extent in the GE phase. However, the design of GE tasks can affect what and how students' prior knowledge is activated and thus, whether students' responses semantically associate with target content knowledge, which influences subsequent learning in the DCE phase. Generating high-quality internal feedback is the key component of the DCE phase. Learning materials and activities in this phase can influence the quality of students' internal feedback by affecting students' processing of reference information as well as how they compare erroneous responses and reference information. In the following section, we present two fundamental principles for helping students learn from errors based on our model. We illustrate and validate the two principles through empirical evidence from prior research, as summarized in Table 2.

Two principles to support students' learning from errors

Principle 1: The GE phase should align with learning goals and support students to generate responses semantically associated with target content knowledge

The purpose of the GE phase is to activate students' deficient prior knowledge when students generate erroneous

Table 2. Two principles to support students' learning from errors.

Principles	Recommended design in past research
Principle 1: The GE phase should align with learning goals and support students to generate responses semantically associated with target content knowledge.	• When the GE phase happens *before* initial learning of target content knowledge: • - Increase the quantity and complexity of cues in memory questions in a pretest GE task (e.g., multiple-choice questions with alternatives mentioned in the learning texts, integrative factual questions combining facts across learning texts) - Design challenging yet feasible high-level learning GE tasks (e.g., tasks using abstract contrasting cases for invention) - Provide guidance and practice opportunities for challenging GE tasks (e.g., guided inventing activities, practice inventing activities in multiple sessions) - Group students in GE tasks based on their abilities and knowledge to encourage effective collaboration • When the GE phase happens *after* initial learning of target content knowledge: - Encourage students to invest more retrieval effort (e.g., using short answer questions, embedding questions in authentic contexts, repeated post-tests before studying feedback, delayed post-tests)
Principle 2: The DCE phase should help students self-explain to generate high-quality internal feedback when processing and comparing with reference information.	• The design of reference information: - Content: Provide informative or elaborative external feedback for post-test; Choose or combine different forms of reference information as appropriate (e.g., rubrics, scripts, normative/expert solutions, worked examples) - Presentation: Present only parts of correct answers for the pretest; Provide immediate feedback • The design of learning activities: - Teachers and students interpret or co-generate rubrics together before self-assessment - Students observe teachers modeling the reflective comparison process - Teachers lead discussions and/or design specific instructions on the reflective comparison process - Teachers provide feedback on students' self-assessment - Teachers design self-regulation activities following self-assessment (e.g., select more problems to solve, create action plans, rewrite the essay) - Support students to apply effective learning strategies (e.g., self-testing after reading reference information, self-explaining during reflective comparison

Note: GE: generating errors; DCE: detecting & correcting error.

responses to GE tasks, which lays the foundation for the subsequent DCE phase. An important question is how to activate students' prior knowledge to facilitate learning from errors. First, the design of GE tasks should align with specific learning goals. For example, to improve retention of knowledge, GE tasks can simply involve questions and/or activities that trigger the cognitive process of retrieval. To improve higher-level learning of knowledge application and transfer, GE tasks should require activating students' conceptual understanding and higher-order thinking through problem-solving, inventing, report writing, etc. Beyond learning goals, GE tasks should encourage and support students to generate responses semantically associated with target content knowledge. This is important because our model suggests that when students' responses are not semantically associated with target content knowledge, it is unlikely for students to successfully update prior knowledge in the DCE phase and thus, learn from errors.

When the GE phase happens *before* students' initial learning of target content knowledge, increasing the quantity and complexity of cues in memory questions may help broaden the benefits of pretesting to untested information (Carpenter et al., 2018). For example, Little and Bjork (2016) found that a pretest using multiple-choice questions with incorrect alternatives mentioned in a subsequent text facilitated students' memory for both pretested and non-pretested information in the text. In contrast, a pretest asking the same question in a short answer format had limited memory benefits for pretested information only. A more recent study by Hilaire et al. (2019) found that integrative factual questions that required combining facts located across a text improved memory for both pretested and non-pretested information. However, when a pretest used factual questions asking for an isolated fact stated in the text, there was a limited beneficial effect of pretesting on tested information.

Similarly, in higher-level learning tasks, students benefit from challenging yet feasible tasks *before* any formal instruction of target content knowledge. These tasks encourage or guide students to generate multiple solutions and representations by activating rich and relevant prior knowledge (Kapur & Bielaczyc, 2012). For example, using contrasting cases can help activate relevant prior knowledge by highlighting similarities and/or differences across cases and thus, direct students' attention to important features of problems or concepts (Schwartz et al., 2011; Schwartz & Martin, 2004). Beneficial contrasting cases can differ in surface features but share the same deep features of a complex concept (e.g., Jacobson et al., 2020) or they can share the same cover story or problem context but vary in deep features of a concept (e.g., Holmes et al., 2014; Schwartz et al., 2011).

What students do with GE tasks also impacts learning from errors. For instance, when using contrasting cases, encouraging students to invent a single method or representation may improve learning better than asking students to simply compare similarities and differences across cases, or to self-explain contrasting cases without any invention (e.g., Chin et al., 2016; Schalk et al., 2018). Moreover, invention activities work best with abstract contrasting cases that strip away as many contextual details as possible (e.g., Schalk et al., 2018). Finally, students may need guidance and practice with GE tasks. For instance, Holmes et al. (2014) guided students by explicitly prompting them to compare pairs of contrasting cases, explain their thoughts, and evaluate their invention. Guided students were more likely to explore different solutions during the invention and sustained long-term conceptual understanding than unguided students. Also, when students had only one inventing phase, they might not benefit from inventing activities because they were unfamiliar with the activities, which imposed a high cognitive load. However, after students practiced inventing activities in multiple inventing phases, they learned better compared with studying worked examples (Glogger-Frey et al., 2015).

Kapur and Bielaczyc (2012) also proposed other critical features for designing students' generation activities such as grouping students based on their abilities and knowledge to encourage effective collaboration (e.g., Westermann & Rummel, 2012) as well as providing students with motivational and emotional support. All the designs discussed above can encourage and support students to generate responses semantically associated with target content knowledge. This is especially important when students have limited prior content knowledge *before* initial learning of target content knowledge, in which students are more likely to generate random guesses or irrelevant errors that are unlikely to benefit learning.

When the GE phase happens *after* initial learning of target content knowledge, students' responses should be semantically associated with target content knowledge, yet the more retrieval effort students invest in generating responses, the better students learn from errors. For example, a post-test with short answer questions and external feedback appears to boost students' retention of knowledge more than multiple-choice questions (e.g., Kang et al., 2007). It can also be helpful to embed post-test questions in an authentic context (e.g., Larsen et al., 2013). Other memory research has explored the number and timing of post-tests. Arnold and Mcdermott (2013b) showed that, after initial learning of Russian-English word pairs, students who were tested five times before restudying learned faster and more effectively than students who were tested only once before restudying, despite higher error rates in both conditions. Similarly, inserting post-test questions with correct answer feedback at different locations of a textbook chapter resulted in better long-term retention of tested information than putting all the same questions and correct answers at the end of the chapter (Uner & Roediger, 2018). There is also evidence that a longer delay between initial learning and the post-test supported better long-term memory of tested information than a shorter delay (Carpenter et al., 2009). All the GE tasks in these studies encouraged students to invest retrieval effort to activate rich and related prior knowledge as much as possible.

Principle 2: The DCE phase should help students self-explain to generate high-quality internal feedback when processing and comparing with reference information

According to our model, high-quality internal feedback during the DCE phase determines what and how much students learn from errors. The quality of internal feedback depends on the design of reference information and students' engagement with reference information, such as how students make sense of reference information and how they compare their responses with reference information. Regarding the design of reference information, one should consider two dimensions of design features: the content and the presentation of reference information.

The content of reference information should at least include the correct/expert solutions or standards for students to compare with their erroneous responses. Providing explanations of correct answers or standards may yield additional benefits in some learning contexts. For example, much memory research has found external feedback that included correct answers enhanced learning from errors better than no feedback or corrective feedback that only conveyed whether a response was correct (e.g., Fazio et al., 2010; Marsh & Eliseev, 2019; Metcalfe & Kornell, 2007; Metcalfe, 2017; Vojdanoska et al., 2010; for an exception, see Anderson & McDaniel, 2021). However, providing correct answer feedback to a *pretest* can potentially hinder learning because students may not pay attention to the subsequent learning materials after knowing the correct answers (Sana et al., 2021). Moreover, for external feedback on *post-tests*, there is evidence that, compared to corrective feedback, providing elaborative feedback, also called explanatory feedback, of explaining correct or incorrect answers may further boost students' memory for tested knowledge (Enders et al., 2021).

Self-assessment research also found benefits of various forms of reference information, such as rubrics, scripts, normative/expert solutions, and worked examples. Rubrics usually consist of a list of dimensions/goals of performance, a scale (i.e., qualitative or quantitative) for grading different levels of achievement in each dimension/goal, and a description for each achievement level. Scripts are metacognitive prompts in the form of structured checklists, steps, or questions based on models of expert performance. Normative/expert solutions are the final products that meet standards, whereas worked examples not only show the final products but also demonstrate the step-by-step process of performing a learning task. On one hand, different forms of reference information make them appropriate as self-assessment tools for different learning tasks. Rubrics and scripts are more appropriate and often used to help students self-assess open-ended tasks such as essay writing (Andrade & Boulay, 2003), oral presentations (Hafner & Hafner, 2003), or experiment reports (Memis & Seven, 2015), whereas normative/expert solutions and worked examples are more appropriate and often used to help self-assess close-ended tasks such as problem-solving with a definite answer (Safadi, 2017a). On the other hand, different forms of reference information can be combined to complement each other. For instance, model essays (i.e., normative/expert solutions) may serve as additional examples to support students' understanding of provided rubrics (e.g., Andrade et al., 2008, 2010). Rubrics can also be integrated with worked examples to encourage students to actively process worked examples by self-grading their levels of proficiency on each step, which is more beneficial for conceptual understanding than using worked examples alone (Safadi & Saadi, 2021).

The presentation of reference information also influences learning from errors. For example, students may benefit from external feedback presenting only parts of correct answers. Finn and Metcalfe (2010) used scaffolded feedback for a *pretest* that provided students with one letter of the correct answer at a time until students could generate the rest of the answer. They found that students in the scaffolded feedback condition retained more correct answers than students who were provided with standard feedback that showed full answers. However, other studies suggest that when students generate errors in a *post-test*, scaffolded feedback and standard feedback have similar effects on long-term retention (Fiechter & Benjamin, 2019; Leggett et al., 2019). Moreover, there might be an advantage of immediate feedback over delayed feedback in short-term learning from errors. For example, memory research on the learning of word pairs showed that providing feedback immediately after students' errors improved retention more than delayed feedback (Grimaldi & Karpicke, 2012; Hays et al., 2013). Yet, the advantage of immediate over delayed feedback to correct errors may diminish in a delayed test of long-term retention (Butler et al., 2007; Smith & Kimball, 2010).

Even if reference information is well-designed in content and presentation, generating high-quality internal feedback via self-explaining requires a deep understanding of reference information as well as reflective comparisons between erroneous responses and reference information. Instructors can play an important role to support these processes. For example, teachers and students interpreting or co-generating rubrics together before self-assessment can help students better process reference information to appreciate and understand standards for learning tasks, and thus, improve the learning of content knowledge (e.g., Andrade et al., 2008; Sadler & Good, 2006). Students can also learn how to effectively self-assess in the DCE phase by observing a model performing reflective comparison while thinking aloud (e.g., Kostons et al., 2012; Raaijmakers et al., 2018).

Moreover, teachers can lead class discussions and design specific instructions on reflective comparison with reference information in the DCE phase (e.g., Loibl & Rummel, 2014a; Ross et al., 2002; Sadler & Good, 2006). For instance, in addition to providing rubrics, Andrade et al. (2008, 2010) guided students to underline key phrases in the rubrics and mark sentences in their draft essays as evidence of meeting standards in the underlined rubrics. If students failed to find evidence in their draft essays that matched the rubrics, they were further instructed to write a reminder to improve on the final essay. Finally, teachers can provide feedback on students' self-assessment (e.g., Andrade et al., 2010) and design follow-up self-regulation activities such as asking students to select more problems to work on (Zamora et al., 2018a), creating action plans for future learning (Ross & Starling, 2008), or improving the first draft essay by writing the essay again (Andrade et al., 2010). All these teachers' supports encourage students to generate high-quality internal feedback and provide subsequent opportunities for students to act on their internal feedback.

Instead of relying on teachers, students can help themselves generate high-quality internal feedback by applying effective learning strategies in the DCE phase. For example, self-testing after reading reference information, even without any feedback, can be a promising learning strategy that protects against the return of errors by deeply processing external feedback (e.g., Metcalfe & Miele, 2014). Another way to support students to apply effective learning strategies is to use self-explanation prompts. For instance, in a series of studies by Safadi (2017a, 2017b, 2018), students compared their problem-solving processes with worked examples. Meanwhile, students were prompted to identify and circle mistakes, correct mistakes by writing down the correct physics principles or laws, explain why their erroneous responses are incorrect, and give advice to a friend to avoid the same mistakes in the future. Results showed that students who were prompted to self-explain performed significantly better than unprompted students in conceptual understanding.

Future research directions

Designing the GE phase to activate and externalize related prior knowledge

Prior studies have provided abundant evidence to show the positive effects of activating rich and related prior knowledge in the GE phase on learning from errors. However, past research mainly focused on examining GE tasks to facilitate learning of *content knowledge* such as retention of knowledge and higher-level learning. Although self-assessment research has shown learning from errors can facilitate the development of SRL knowledge and skills, no research has investigated the effects of designing different GE tasks to support this particular learning goal. This is probably because SRL knowledge and skills are inherently embedded in all learning tasks and self-assessment is itself an SRL skill (Andrade, 2019; Butler & Winne, 1995; Panadero et al., 2017). Nevertheless, future research on the design of GE tasks specifically for activating SRL knowledge and skills may be promising for developing training programs that help students become independent learners. For instance, students are notorious for using ineffective cues or strategies to make judgments of learning (Bjork et al., 2013). One can design GE tasks to activate students' prior knowledge of adopting ineffective cues or strategies to monitor learning. Such knowledge activation might provide greater opportunities for students to learn from the following reference information that presents SRL knowledge of effective cues or strategies for accurate judgments of learning.

Also, SRL is a complex process involving different phases, subphases, and numerous constructs. Learning from errors may benefit different SRL constructs to a different extent. More research is needed to clarify which specific SRL constructs benefit more or less from self-generated errors and corresponding intricate mechanisms. In this future research, it is important to examine different SRL constructs under the same designs or implementations of experiments to avoid confounding explanations. Additionally, the relationship between learning content knowledge and improving SRL knowledge and skills in the context of learning from errors is unknown. Effective learning from errors in content knowledge might be necessary for effective learning from errors to improve SRL knowledge and skills. For example, a student who does not know what is wrong with their learning of a physics concept might not be able to detect what is wrong with their previous learning strategies to study the concept.

Moreover, when students have insufficient prior knowledge to understand the GE tasks or generate meaningful responses semantically related to the target content knowledge, they may not learn from errors effectively (e.g., Loibl & Leuders, 2019; Seabrooke et al., 2021). Future research should not only consider designing effective GE tasks to activate rich and relevant prior knowledge but also the moderating effect of different levels of prior knowledge on learning from errors. A related research gap is the potential moderating effect of prior knowledge on the quantity and quality of errors. Although past research found a positive association between the number of students' erroneous solutions and learning (e.g., Kapur, 2012; Kapur & Bielaczyc, 2012), a measure of the quality of students' errors might provide insights into how different levels of prior knowledge influence learning from errors. For instance, the semantic relationship between response and target content knowledge in our model can provide a starting point to measure the quality of errors.

Finally, researchers should systematically investigate specific design features of GE tasks, such as the type or format of questions in GE tasks and the timing or presentation of GE tasks. Particularly, there is a lack of research on GE tasks that elicit responses in multiple representations. Most GE tasks in past research require responses in verbal representations. However, generating visual representations might have unique benefits in activating and externalizing prior knowledge to facilitate learning from errors. For example, visual representations such as drawings, concept/knowledge maps, and graphic organizers represent information in a nonlinear, spatial way, making the relationships among concepts perceptually salient and transparent (Larkin & Simon, 1987). Accordingly, research suggests learner-generated visual representations can better reveal gaps and misconceptions in students' conceptual knowledge, facilitate comparisons between correct and incorrect representations, and direct students' attention toward key features of concepts (Becker et al., 2016; Bobek & Tversky, 2016; Horiguchi et al., 2014; Loibl & Leuders, 2018; Ryan & Stieff, 2019; Valanides et al., 2013; Van Meter & Riley, 1999). Therefore, future research should test the effects of GE tasks that elicit responses in multiple representations on students' learning from errors, especially for higher-level learning in STEM domains involving complex spatial relationships or interrelated ideas.

Designing the DCE phase to encourage deep processing of and reflective comparison to reference information

First, there is limited research on how to present reference information. To facilitate retention of knowledge through learning from errors, more evidence is needed to confirm the benefits of immediate feedback to post-tests after initial learning of target knowledge. However, there have been inconsistent definitions of 'immediate' or 'delayed' in research on the timing of external feedback. In some studies, the timing of feedback was relative to completing the entire learning task, such as a whole test (e.g., Smith & Kimball, 2010), whereas in other studies the timing of feedback was relative to each item of a test (e.g., Grimaldi & Karpicke, 2012). Thus, immediate feedback to a whole test in one study could be delayed feedback to each test item in another study. Such inconsistent use of terminology might be partly responsible for mixed empirical findings regarding the optimal timing of external feedback (Mason & Bruning, 2001; Mory, 2004; Shute, 2008). Therefore, in future studies, researchers should clarify and explicitly define the timing of feedback under investigation. Beyond retention of knowledge, higher-level learning and SRL researchers should also examine the optimal timing of reference information for learning from errors in GE tasks before and after initial learning of target knowledge.

In addition, more systematic research is needed on other content- or presentation-related design features of reference information. For example, multimedia and multiple representations are widely used in academic learning materials. Surprisingly, no research on students' learning from errors has systematically tested the effects of adding multiple representations in reference information and applying multimedia learning principles to present reference information that contains visualizations or multiple representations, although the multimedia learning effect and multimedia learning principles have been investigated and well supported in various learning contexts (Mayer, 2014; Mayer & Fiorella, 2022). For instance, there is initial evidence that, after generating drawings, students learned better by comparing their inaccurate drawings with reference information in a visual format than in a verbal format (Zhang & Fiorella, 2021).

Finally, in past research, teachers took the dominant roles in instructional support in the DCE phase. There is a lack of research on how to help students apply effective learning strategies themselves in the DCE phase. Fiorella and Mayer (2015, 2016) identified eight strategies intended to promote generative learning by encouraging students to actively make sense of to-be-learned information: summarizing, mapping, drawing, imagining, self-testing, self-explaining, teaching, and enacting. Few studies in memory research involve helping students use generative strategies to actively process external feedback or make reflective comparisons (e.g., Metcalfe & Miele, 2014), likely because the learning materials are relatively straightforward to process or compare (e.g., word pairs, trivia questions, etc.). Yet the way students process external feedback and compare their answers to external feedback, even in simple materials, might influence their retention of knowledge. Moreover, no research on higher-level learning or SRL has examined the effects of generative strategies on students' processing of reference information to learn from errors. Future research can also explore how to measure the quality of students' internal feedback by examining their products of generative learning activities. Generative activities such as self-explaining or drawing not only promote deep learning but also enable students to externalize their thinking in a verbal or visual format, making it possible to examine the quality of students' internal feedback.

The role of motivation and emotion in learning from errors

Research on learning from self-generated errors should consider both students' cognitive/metacognitive processes and emotional and motivational processes to gain a comprehensive picture of the phenomenon (Tulis et al., 2016). In our model, motivation and emotion are parts of SRL knowledge and can influence and be influenced by learning from errors, but these relations have only limited, and thus currently insufficient, empirical evidence from prior studies. For example, limited empirical studies in memory research measured students' motivation and emotion and reported mixed evidence on whether generating errors indeed benefits learning through increased motivation and positive emotions (e.g., Griffiths & Higham, 2018; Seabrooke, Mitchell, et al., 2019). There are also few findings in PS-I research showing the positive effects of invention activities on students' curiosity and goal orientation (e.g., Loibl & Rummel, 2014a; Sinha et al., 2021). Only self-assessment research found relatively strong evidence that students' motivational beliefs of self-efficacy can be improved by self-assessment with instructional support (Panadero et al., 2017).

Overall, the existing empirical evidence suggests students' motivation and emotion play an important role in the process of learning from errors, yet more research is needed to establish specific mechanisms. Future work should not only consider moderating and/or mediating the effects of motivational beliefs and emotional states on learning from errors but also how different experiences of learning from errors impact students' motivational beliefs and emotional states. For example, it is unclear how students' motivation and emotion might influence the quality of their responses in the GE phase, which can influence the subsequent DCE phase. Future research should also explore additional motivational constructs (e.g., students' self-expectancy and values, students' mindset, etc.) as well as emotional constructs (e.g., negative emotions related to errors and failures) in the context of learning from errors.

More importantly, negative feelings are likely common when students experience errors or failures in learning, but students may not be able to manage negative emotions well without support. Researchers have emphasized the significance of a learning environment or classroom climate that embraces errors/failures and provides metacognitive, motivational, and emotional support (Hattie & Timperley, 2007; Henry et al., 2019; Kapur & Bielaczyc, 2012; Keith & Frese, 2005). A thorough understanding of students' motivation and emotion in learning from errors will help future research develop and test interventions for overcoming the negative effects of errors and to optimize learning from errors. Our model can guide future research to systematically examine the negative effects of errors and relevant interventions by manipulating the design of learning tasks, reference information, and learning activities in the GE and DCE phase.[4]

Developmental differences and other individual differences in learning from errors

Participants in past research on learning from errors ranged widely from primary school students to undergraduates. However, different ages of participants are not evenly distributed within different lines of research. Specifically, participants in most memory research were undergraduate students, whereas participants in most self-assessment research were K-12 students. Only PS-I research has a relatively balanced mix of students from both age ranges. The lack of K-12 participants in memory research and the lack of adult participants in self-assessment research make it difficult to generalize findings from the two lines of research. Therefore, it is crucial for future research on learning from errors in improving retention of knowledge and in self-assessment to sample from a population of different ages. Meanwhile, future research may consider systematically investigating the moderating effect of developmental differences on learning from errors. Although we believe the two general principles to support learning from errors should work for students of different ages, some implementation details might need to be adjusted for students with developmental differences.

Besides developmental differences, the direct and indirect influences of many other individual differences on learning from errors are unclear. For example, students' motivational beliefs develop in different learning and socio-cultural contexts (Rosenzweig et al., 2021). As a result, students from different learning communities, racial groups, genders, and cultural environments can have different beliefs about errors and react differently to making errors. Thus, these individual differences will also have effects on students' learning from errors. In sum, future research should consider the influences of students' developmental differences as well as many other individual differences on learning from errors.

Conclusion

Making errors is a part of the learning process and provides opportunities to improve learning and self-regulation. The present review proposed a theoretical model to describe the two major phases of learning from self-generated errors: The Generating Errors (GE) phase contributes to learning by activating students' deficient prior knowledge and the Detecting and Correcting Errors (DCE) phase contributes to learning through high-quality internal feedback. Corresponding to the two phases, our model suggests two general principles to support learning from errors: (1) The GE phase should align with learning goals and support students to generate responses semantically associated with target content knowledge, and (2) the DCE phase should help students self-explain to generate high-quality internal feedback when processing and comparing with reference information. Depending on specific learning goals, prior studies provided empirical evidence for various

[4]Although our model focused on errors that happen naturally due to students' deficient knowledge, deliberately generating errors when students know the correct answers might be a way to overcome negative effects of errors on emotion and motivation (Wong & Lim, 2022).

implementations of the principles. Yet there are more unknowns than knowns about how to utilize errors generated by learners to optimize learning and self-regulation. Our model provides an outline for future research to systematically investigate the GE and DCE phases to uncover more specific design features and mechanisms that explain and support students' learning from self-generated errors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by grants from the National Science Foundation [1955348 and 2055117], a National Academy of Education/ Spencer Foundation Postdoctoral Fellowship awarded to Logan Fiorella and Graduate School Dean's Award for Social Science from University of Georgia awarded to Qian Zhang

ORCID

Qian Zhang http://orcid.org/0000-0001-9268-9930
Logan Fiorella http://orcid.org/0000-0002-1672-325X

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712

Anderson, F. T., & McDaniel, M. A. (2021). Restudying with the quiz in hand: When correct-answer feedback is no better than minimal feedback. *Journal of Applied Research in Memory and Cognition*, *10*(2), 278–288. https://doi.org/10.1016/j.jarmac.2020.10.004

Anderson, L. W., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Andrade, H. (2009). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. Andrade & G. J. Cizek (Eds.), *Handbook of Formative Assessment*. Routledge.

Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research*, *97*(1), 21–30. https://doi.org/10.1080/00220670309596625

Andrade, H. L. (2018). Feedback in the context of self-assessment. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 376–408). Cambridge University Press.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, e00087. https://doi.org/10.3389/feduc.2019.00087

Andrade, H. L., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education*, *17*, 199–214.

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement*, *27*(2), 3–13. https://doi.org/10.1111/j.1745-3992.2008.00118.x

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, *48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Arnold, K. M., & McDermott, K. B. (2013a). Free recall enhances subsequent learning. *Psychonomic Bulletin & Review*, *20*(3), 507–513. https://doi.org/10.3758/s13423-012-0370-3

Arnold, K. M., & McDermott, K. B. (2013b). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology*, *39*(3), 940–945. https://doi.org/10.1037/a0029199

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt, Rinehart and Winston.

Baars, M., Vink, S., van Gog, T., de Bruin, A., & Paas, F. (2014). Effects of training self-assessment and using assessment standards on retrospective and prospective monitoring of problem solving. *Learning and Instruction*, *33*, 92–107. https://doi.org/10.1016/j.learninstruc.2014.04.004

Becker, N., Noyes, K., & Cooper, M. (2016). Characterizing students' mechanistic reasoning about london dispersion forces. *Journal of Chemical Education*, *93*(10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences*, *21*(3), 399–432. https://doi.org/10.1080/10508406.2011.651232

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444.

Bobek, E., & Tversky, B. (2016). Creating visual explanations improves learning. *Cognitive Research*, *1*, 27.

Brown, G., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367–394). SAGE.

Brown, G. T., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research*, *2*, 22–30.

Butler, A. C., Karpicke, J. D., & Roediger, H. L. III (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology*, *13*(4), 273–281.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. https://doi.org/10.3102/00346543065003245

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 69–84. https://doi.org/10.1007/s11409-006-6894-z

Carless, D. (2019). Feedback loops and the longer-term: Towards feedback spirals. *Assessment & Evaluation in Higher Education*, *44*(5), 705–714. https://doi.org/10.1080/02602938.2018.1531108

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology*, *35*(6), 1563–1569.

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology*, *23*(6), 760–771. https://doi.org/10.1002/acp.1507

Carpenter, S. K., Rahman, S., & Perkins, K. (2018). The effects of prequestions on classroom learning. *Journal of Experimental Psychology*, *24*(1), 34–42.

Carpenter, S. K., & Toftness, A. R. (2017). The effect of prequestions on learning from video presentations. *Journal of Applied Research in Memory and Cognition*, *6*(1), 104–109. https://doi.org/10.1016/j.jarmac.2016.07.014

Chi, M. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in Instructional Psychology*, *5*, 161–238.

Chin, D. B., Chi, M., & Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: Inventing a general solution versus compare and contrast. *Instructional Science*, *44*(2), 177–195. https://doi.org/10.1007/s11251-016-9374-0

Clark, C. M., Bjork, E. L., & Bjork, R. A. (2021). On the role of generation rules in moderating the beneficial effects of errorful generation. *Zeitschrift Für Psychologie, 229*(2), 120–130. https://doi.org/10.1027/2151-2604/a000442

Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, *22*(9), 1041–1053. https://doi.org/10.1080/095006900416901

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology*, *113*(4), 552–568.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34*(3), 169–189. https://doi.org/10.1207/s15326985ep3403_3

Enders, N., Gaschler, R., & Kubik, V. (2021). Online quizzes with closed questions in formal assessment: How elaborate feedback can promote learning. *Psychology Learning & Teaching*, *20*(1), 91–106. https://doi.org/10.1177/1475725720971205

Fazio, L. K., Huelser, B. J., Johnson, A., & Marsh, E. J. (2010). Receiving right/wrong feedback: Consequences for learning. *Memory*, *18*(3), 335–350.

Fiechter, J. L., & Benjamin, A. S. (2019). Techniques for scaffolding retrieval practice: The costs and benefits of adaptive versus diminishing cues. *Psychonomic Bulletin & Review*, *26*(5), 1666–1674.

Finn, B., & Metcalfe, J. (2010). Scaffolding feedback to maximize long-term error correction. *Memory & Cognition*, *38*(7), 951–961.

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. Cambridge University Press.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 393–408. https://doi.org/10.1037/0022-0663.95.2.393

Glogger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., & Renkl, A. (2015). Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learning and Instruction*, *39*, 72–87. https://doi.org/10.1016/j.learninstruc.2015.05.001

Griffiths, L., & Higham, P. A. (2018). Beyond hypercorrection: Remembering corrective feedback for low-confidence errors. *Memory*, *26*(2), 201–218. https://doi.org/10.1080/09658211.2017.1344249

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, *40*(4), 505–513. https://doi.org/10.3758/s13421-011-0174-0

Hafner, J., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education*, *25*(12), 1509–1528. https://doi.org/10.1080/0950069022000038268

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology*, *39*, 290–296.

Henry, M. A., Shorter, S., Charkoudian, L., Heemstra, J. M., & Corwin, L. A. (2019). Fail is not a four-letter word: A theoretical framework for exploring undergraduate students' approaches to academic challenge and responses to failure in stem learning environments. *CBE—Life Sciences Education*, *18*(1), ar11. https://doi.org/10.1187/cbe.18-06-0108

Hilaire, K. J. S., Carpenter, S. K., & Jennings, J. M. (2019). Using prequestions to enhance learning from reading passages: The roles of question type and structure building ability. *Memory*, *27*(9), 1204–1213. https://doi.org/10.1080/09658211.2019.1641209

Holmes, N. G., Day, J., Park, A. H. K., Bonn, D. A., & Roll, I. (2014). Making the failure more productive: Scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. *Instructional Science*, *42*(4), 523–538. https://doi.org/10.1007/s11251-013-9300-7

Horiguchi, T., Imai, I., Toumoto, T., & Hirashima, T. (2014). Error-based simulation for error-awareness in learning mechanics: An evaluation. *Journal of Educational Technology & Society*, *17*, 1–13.

Hoyos, C., & Gentner, D. (2017). Generating explanations via analogical comparison. *Psychonomic Bulletin & Review*, *24*(5), 1364–1374. https://doi.org/10.3758/s13423-017-1289-5

Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition*, *40*(4), 514–527. https://doi.org/10.3758/s13421-011-0167-z

Jacobson, M. J., Goldwater, M., Markauskaite, L., Lai, P. K., Kapur, M., Roberts, G., & Hilton, C. (2020). Schema abstraction with productive failure and analogical comparison: Learning designs for far across domain transfer. *Learning and Instruction*, *65*, 101222. https://doi.org/10.1016/j.learninstruc.2019.101222

Jansen, R. S., van Leeuwen, A., Janssen, J., Jak, S., & Kester, L. (2019). Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis. *Educational Research Review*, *28*, 100292. https://doi.org/10.1016/j.edurev.2019.100292

Jax, J., Ahn, J. N., & Lin-Siegler, X. (2019). Using contrasting cases to improve self-assessment in physics learning. *Educational Psychology*, *0*, 1–24.

Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*(4–5), 528–558. https://doi.org/10.1080/09541440601056620

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, *38*(6), 523–550. https://doi.org/10.1007/s11251-009-9093-x

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*(4), 651–672. https://doi.org/10.1007/s11251-012-9209-6

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science, 38*(5), 1008–1022. https://doi.org/10.1111/cogs.12107

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, *51*(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, *21*(1), 45–83. https://doi.org/10.1080/10508406.2011.591717

Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *The Journal of Applied Psychology*, *90*(4), 677–691. https://doi.org/10.1037/0021-9010.90.4.677

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology*, *35*(4), 989–998.

Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology*, *41*, 283–294.

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 65, pp. 183–215). Academic Press.

Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, *22*(2), 121–132. https://doi.org/10.1016/j.learninstruc.2011.08.004

Lamnina, M., & Chase, C. C. (2019). Developing a thirst for knowledge: How uncertainty in the classroom influences curiosity, affect, learning, and transfer. *Contemporary Educational Psychology*, *59*, 101785. https://doi.org/10.1016/j.cedpsych.2019.101785

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100. https://doi.org/10.1111/j.1551-6708.1987.tb00863.x

Larsen, D. P., Butler, A. C., Lawson, A. L., & Roediger, H. L. (2013). The importance of seeing the patient: Test-enhanced learning with standardized patients and written tests improves clinical application of knowledge. *Advances in Health Sciences Education*, *18*(3), 409–425.

Leggett, J. M. I., Burt, J. S., & Carroll, A. (2019). Retrieval practice can improve classroom review despite low practice test performance. *Applied Cognitive Psychology*, *33*(5), 759–770. https://doi.org/10.1002/acp.3517

Little, J. L., & Bjork, E. L. (2016). Multiple-choice pretesting potentiates learning of related information. *Memory & Cognition*, *44*(7), 1085–1101.

Loehr, A., Fyfe, E., & Rittle-Johnson, B. (2014). Wait for it … delaying instruction improves mathematics problem solving: A classroom study. *The Journal of Problem Solving*, *7*(1), 1166. https://doi.org/10.7771/1932-6246.1166

Loibl, K., & Leuders, T. (2018). Errors during exploration and consolidation — The effectiveness of productive failure as sequentially guided discovery learning. *Journal für Mathematik-Didaktik*, *39*(1), 69–96. https://doi.org/10.1007/s13138-018-0130-7

Loibl, K., & Leuders, T. (2019). How to make failure productive: Fostering learning from errors through elaboration prompts. *Learning and Instruction*, *62*, 1–10. https://doi.org/10.1016/j.learninstruc.2019.03.002

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., & Rummel, N. (2014a). Knowing what you don't know makes failure productive. *Learning and Instruction*, *34*, 74–85. https://doi.org/10.1016/j.learninstruc.2014.08.004

Loibl, K., & Rummel, N. (2014b). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science*, *42*(3), 305–326. https://doi.org/10.1007/s11251-013-9282-5

Marsh, E. J., & Eliseev, E. D. (2019). Correcting student errors and misconceptions. In *The Cambridge handbook of cognition and education* (pp. 437–459). Cambridge University Press.

Mason, B., & Bruning, R. (2001). Providing feedback in computer-based instruction: What the research tells us. *Center for Instructional Innovation*, 15.

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning* (Second Edition). Cambridge University Press.

Mayer, R. E., & Fiorella, L. (Eds.). (2022). *The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press.

Memis, E. K., & Seven, S. (2015). Effects of an SWH approach and self-evaluation on sixth grade students' learning and retention of an electricity unit. *International Journal of Progressive Education*, *11*, 32–49.

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*, 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Metcalfe, J., & Huelser, B. J. (2020). Learning from errors is attributable to episodic recollection rather than semantic mediation. *Neuropsychologia*, *138*, 107296.

Metcalfe, J., & Kornell, N. (2007). Principles of cognitive science in education: The effects of generation, errors, and feedback. *Psychonomic Bulletin & Review*, *14*(2), 225–229. https://doi.org/10.3758/bf03194056

Metcalfe, J., & Miele, D. B. (2014). Hypercorrection of high confidence errors: Prior testing both enhances delayed performance and blocks the return of the errors. *Journal of Applied Research in Memory and Cognition*, *3*(3), 189–197. https://doi.org/10.1016/j.jarmac.2014.04.001

Metcalfe, J., & Xu, J. (2018). Learning from one's own errors and those of others. *Psychonomic Bulletin & Review*, *25*(1), 402–408. https://doi.org/10.3758/s13423-017-1287-7

Mory, E. H. (2004). Feedback research revisited. In *Handbook of research on educational communications and technology* (Vol. 2, pp. 745–783). Citeseer.

Nicol, D. (2020). The power of internal feedback: Exploiting natural comparison processes. *Assessment & Evaluation in Higher Education*, *46*(5), 1–23. https://doi.org/10.1080/02602938.2020.1823314

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, *103*(2), 241–262. https://doi.org/10.1037/0033-295X.103.2.241

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, *39*(3), 125–132. https://doi.org/10.1016/j.stueduc.2013.04.001

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

Panadero, E., Lipnevich, A., & Broadbent, J. (2019). Turning self-assessment into self-feedback. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.), *The impact of feedback in higher education: Improving assessment outcomes for learners* (pp. 147–163). Springer International Publishing.

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education*, *21*, 133–148.

Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*, *22*(6), 806–813. https://doi.org/10.1016/j.lindif.2012.04.007

Piaget, J. (1952). *The origins of intelligence in children*. W W Norton & Co.

Potts, R., Davies, G., & Shanks, D. R. (2019). The benefit of generating errors during learning: What is the locus of the effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(6), 1023–1041. https://doi.org/10.1037/xlm0000637

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology*, *143*(2), 644–667. https://doi.org/10.1037/a0033194

Raaijmakers, S. F., Baars, M., Schaap, L., Paas, F., van Merriënboer, J., & van Gog, T. (2018). Training self-regulated learning skills with video modeling examples: Do task-selection skills transfer? *Instructional Science*, *46*(2), 273–290. https://doi.org/10.1007/s11251-017-9434-0

Raković, M., Bernacki, M. L., Greene, J. A., Plumley, R. D., Hogan, K. A., Gates, K. M., & Panter, A. T. (2022). Examining the critical role of evaluation and adaptation in self-regulated learning. *Contemporary Educational Psychology*, *68*, 102027. https://doi.org/10.1016/j.cedpsych.2021.102027

Reason, J. (1990). *Human error* (pp. xv, 302). Cambridge University Press.

Reindl, M., Tulis, M., & Dresel, M. (2020). Profiles of emotional and motivational self-regulation following errors: Associations with learning. *Learning and Individual Differences*, *77*, 101806. https://doi.org/10.1016/j.lindif.2019.101806

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology. Applied*, *15*(3), 243–257.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362. https://doi.org/10.1037/0022-0663.93.2.346

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210.

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluating metacognitive scaffolding in guided invention activities. *Instructional Science, 40*(4), 691–710. https://doi.org/10.1007/s11251-012-9208-7

Rosenzweig, E. Q., Wigfield, A., & Eccles, J. S. (2021). Beyond utility value interventions: The why, when, and how for next steps in expectancy-value intervention research. *Educational Psychologist*, *57*(1), 11–30. https://doi.org/10.1080/00461520.2021.1984242

Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002). Student self-evaluation in grade 5–6 mathematics effects on problem- solving achievement. *Educational Assessment*, *8*(1), 43–58. https://doi.org/10.1207/S15326977EA0801_03

Ross, J. A., & Starling, M. (2008). Self-assessment in a technology-supported environment: The case of grade 9 geography. *Assessment in Education*, *15*, 183–199.

Ryan, S. A. C., & Stieff, M. (2019). Drawing for assessing learning outcomes in chemistry. *Journal of Chemical Education*, *96*(9), 1813–1820. https://doi.org/10.1021/acs.jchemed.9b00361

Sadler, P. M., & Good, E. (2006). The impact of self-and peer-grading on student learning. *Educational Assessment*, *11*(1), 1–31. https://doi.org/10.1207/s15326977ea1101_1

Safadi, R. (2017a). Self-diagnosis as a tool for supporting students' conceptual understanding and achievements in physics: The case of 8th-graders studying force and motion. *Physics Education*, *52*(1), 014002. https://doi.org/10.1088/1361-6552/52/1/014002

Safadi, R. (2017b). Designing efficient self-diagnosis activities in the physics classroom. *The Physics Teacher*, *55*(9), 575–579. https://doi.org/10.1119/1.5011839

Safadi, R. (2018). Knowledge-integration processes and learning outcomes associated with a self-diagnosis activity: The case of 5th-graders studying simple fractions. *International Journal of Science and Mathematics Education*, *16*(5), 929–948. https://doi.org/10.1007/s10763-017-9798-2

Safadi, R., & Saadi, S. (2021). Learning from self-diagnosis activities when contrasting students' own solutions with worked examples: The case of 10th graders studying geometric optics. *Research in Science Education*, *51*(2), 523–546. https://doi.org/10.1007/s11165-018-9806-8

Sana, F., Yan, V. X., Clark, C. M., Bjork, E. L., & Bjork, R. A. (2021). Improving conceptual learning via pretests. *Journal of Experimental Psychology*, *27*(2), 228–236. https://doi.org/10.1037/xap0000322

Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of Educational Psychology*, *109*(8), 1049–1066. https://doi.org/10.1037/edu0000190

Schalk, L., Schumacher, R., Barth, A., & Stern, E. (2018). When problem-solving followed by instruction is superior to the traditional tell-and-practice sequence. *Journal of Educational Psychology*, *110*(4), 596–610. https://doi.org/10.1037/edu0000234

Schunk, D. H., & Ertmer, P. A. (1999). Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences. *Journal of Educational Psychology*, *91*(2), 251–260. https://doi.org/10.1037/0022-0663.91.2.251

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

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129–184. https://doi.org/10.1207/s1532690xci2202_1

Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language*, *104*, 70–82. https://doi.org/10.1016/j.jml.2018.10.001

Seabrooke, T., Mitchell, C. J., Wills, A. J., & Hollins, T. J. (2021). Pretesting boosts recognition, but not cued recall, of targets from unrelated word pairs. *Psychonomic Bulletin & Review*, *28*(1), 268–273. https://doi.org/10.3758/s13423-020-01810-y

Seabrooke, T., Mitchell, C. J., Wills, A. J., Waters, J. L., & Hollins, T. J. (2019). Selective effects of errorful generation on recognition memory: The role of motivation and surprise. *Memory*, *27*(9), 1250–1262. https://doi.org/10.1080/09658211.2019.1647247

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Simpson, A., Maltese, A. V., Anderson, A., & Sung, E. (2020). Failures, errors, and mistakes: A systematic review of the literature. In E. Vanderheiden & C.-H. Mayer (Eds.), *Mistakes, errors and failures across cultures: Navigating potentials* (pp. 347–362). Springer International Publishing.

Sinha, T., Kapur, M., West, R., Catasta, M., Hauswirth, M., & Trninic, D. (2021). Differential benefits of explicit failure-driven and success-driven scaffolding in problem-solving prior to instruction. *Journal of Educational Psychology*, *113*(3), 530–555. https://doi.org/10.1037/edu0000483

Skinner, B. F. (1958). Teaching machines. *Science*, *128*(3330), 969–977. https://doi.org/10.1126/science.128.3330.969

Smith, T. A., & Kimball, D. R. (2010). Learning from feedback: Spacing and the delay-retention effect. *Journal of Experimental Psychology*, *36*(1), 80–95.

Terrace, H. S. (1963). Discrimination learning with and without "errors". *Journal of the Experimental Analysis of Behavior*, *6*, 1–27. https://doi.org/10.1901/jeab.1963.6-1

Toftness, A. R., Carpenter, S. K., Lauber, S., & Mickes, L. (2018). The limited effects of prequestions on learning from authentic lecture videos. *Journal of Applied Research in Memory and Cognition*, *7*(3), 370–378. https://doi.org/10.1016/j.jarmac.2018.06.003

Tulis, M., Steuer, G., & Dresel, M. (2016). Learning from errors: A model of individual processes. *Frontline Learning Research*, *4*(4), 12–26. https://doi.org/10.14786/flr.v4i2.168

Tulis, M., Steuer, G., & Dresel, M. (2018). Positive beliefs about errors as an important element of adaptive individual dealing with errors during academic learning. *Educational Psychology*, *38*(2), 139–158. https://doi.org/10.1080/01443410.2017.1384536

Uner, O., & Roediger, H. L. (2018). The effect of question placement on learning from textbook chapters. *Journal of Applied Research in Memory and Cognition*, *7*(1), 116–122. https://doi.org/10.1016/j.jarmac.2017.09.002

Valanides, N., Efthymiou, I., & Angeli, C. (2013). Interplay of internal and external representations: Students' drawings and textual explanations about shadow phenomena. *Journal of Visual Literacy*, *32*(2), 67–84. https://doi.org/10.1080/23796529.2013.11674710

Van der Kleij, F. M., Adie, L. E., & Cumming, J. J. (2019). A meta-review of the student role in feedback. *International Journal of Educational Research*, *98*, 303–323. https://doi.org/10.1016/j.ijer.2019.09.005

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In *Learning issues for intelligent tutoring systems* (pp. 19–41). Springer.

Van Meter, P., & Riley, S. E. (1999). Students' writing and drawing: What is gained with different representational formats? *Yearbook-National Reading Conference*, *48*, 146–156.

Vojdanoska, M., Cranney, J., & Newell, B. R. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology*, *24*(8), 1183–1195. https://doi.org/10.1002/acp.1630

Wang, Y., & Sperling, R. A. (2020). Characteristics of effective self-regulated learning interventions in mathematics classrooms: A systematic review. *Frontiers in Education*, *5*, e00058. https://doi.org/10.3389/feduc.2020.00058

Westermann, K., & Rummel, N. (2012). Delaying instruction: Evidence from a study in a university relearning setting. *Instructional Science*, *40*(4), 673–689. https://doi.org/10.1007/s11251-012-9207-8

Wong, S. S. H., & Lim, S. W. H. (2019). Prevention–permission–promotion: A review of approaches to errors in learning. *Educational Psychologist*, *54*(1), 1–19. https://doi.org/10.1080/00461520.2018.1501693

Wong, S. S. H., & Lim, S. W. H. (2022). The derring effect: Deliberate errors enhance learning. *Journal of Experimental Psychology*, *151*(1), 25–40. https://doi.org/10.1037/xge0001072

Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, *42*(8), 1247–1262. https://doi.org/10.1080/02602938.2016.1260091

Zamora, Á., Súarez, J. M., & Ardura, D. (2018a). A model of the role of error detection and self-regulation in academic performance. *The Journal of Educational Research*, *111*(5), 595–602. https://doi.org/10.1080/00220671.2017.1349072

Zamora, Á., Suárez, J. M., & Ardura, D. (2018b). Error detection and self-assessment as mechanisms to promote self-regulation of learning among secondary education students. *The Journal of Educational Research*, *111*(2), 175–185. https://doi.org/10.1080/00220671.2016.1225657

Zawadzka, K., & Hanczakowski, M. (2019). Two routes to memory benefits of guessing. *Journal of Experimental Psychology*, *45*(10), 1748–1760.

Zhang, Q., & Fiorella, L. (2021). Learning by drawing: When is it worth the time and effort? *Contemporary Educational Psychology*, *66*, 101990. https://doi.org/10.1016/j.cedpsych.2021.101990

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation* (pp. 13–39). Elsevier.