



Generative Learning: Which Strategies for What Age?

Garvin Brod^{1,2} 

Accepted: 6 September 2020 / Published online: 23 October 2020
© The Author(s) 2020

Abstract

Generative learning strategies are intended to improve students' learning by prompting them to actively make sense of the material to be learned. But are they effective for all students? This review provides an overview of six popular generative learning strategies: concept mapping, explaining, predicting, questioning, testing, and drawing. Its main purpose is to review for what ages the effectiveness of these strategies has been demonstrated and whether there are indications of age-related differences in their effectiveness. The description of each strategy covers (1) how it is supposed to work, (2) the evidence on its effectiveness in different age groups, and (3) if there are age-related differences in its effectiveness. It is found that while all six generative learning strategies reviewed have proven effective for university students, evidence is mixed for younger students. Whereas some strategies (practice testing, predicting) seem to be effective already in lower-elementary-school children, others (drawing, questioning) seem to be largely ineffective until secondary school. The review closes with a call for research on the cognitive and metacognitive prerequisites of generative learning that can explain these differences.

Keywords Generative learning strategies · Active learning · Constructive learning · Developmental differences · Children · Effective learning techniques

Teaching an 8-year-old is different from teaching a 16-year-old. Although this is a trivial statement for educators, the implications are surprisingly often ignored in educational research. Should teachers use different strategies depending on learners' age? Excellent reviews and meta-analyses have compared different learning strategies with the goal of finding the ones most effective for all learners (e.g., Dunlosky et al. 2013; Fiorella and Mayer 2015, 2016; Hattie et al. 1996; Yiping et al. 2001). But are the strategies thus identified equally effective in all age groups? This is an open question because the majority of studies included in these

✉ Garvin Brod
garvin.brod@dipf.de

¹ DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany

² Department of Psychology, Goethe University, Frankfurt am Main, Germany

reviews were performed with university students, and there has been little systematic research regarding age-related differences in strategies' effectiveness. The current review is intended to shed light on the question of whether there are age-related differences in the effectiveness of a particular group of strategies—generative learning strategies (GLSs).

GLSs are grounded in the constructivist view of learning, which posits that learning is an active construction process that is based on an individual's prior knowledge (see Bonwell and Eison 1991; von Glasersfeld 1983; Wittrock 1985). For the purpose of this review, GLSs are defined as activities that prompt learners to produce something meaningful that goes beyond the information provided by an instructor. In doing so, learners have to activate prior knowledge and link it to the provided information, which is assumed to foster integration of new information into existing knowledge structures. This definition of GLSs is supposed to differentiate GLSs from other popular learning strategies that also require activities by the learners but do not require the generation of additional content (e.g., highlighting, paraphrasing). Note that this distinction is highly similar to the one between active and constructive learning activities proposed by Chi (2009). GLSs as defined in this review can, thus, be considered as instantiations of constructive learning activities.

The chosen definition of GLSs has several implications for which strategies are considered for this review. The emphasis on the production of meaningful content that goes beyond the provided information but that does not necessarily have to be invented by the learner implies that generating answers to test questions (i.e., practice testing) counts as a GLS. This holds true independent of whether the learner is able to provide the correct answer. There are learning strategies for which it is not easy to tell whether learners necessarily produce something that is both meaningful and goes beyond the provided information, however. One example is the strategy of summarizing a text. Summarizing a text would qualify as a GLS if learners enrich the provided information in the text with additional content, but it would not qualify when it entails only paraphrasing or condensing the given information. Another example is enacting. Asking learners to manipulate models to represent a story or to perform gestures that enrich the provided information clearly requires the translation of verbal information into actions. Whether these actions go beyond the provided information can be hard to tell, however. In the case of gestures, the production of iconic gestures might qualify, whereas indexical or symbolic gestures would not. Because of these classificatory difficulties and resultant problems in inferring and comparing underlying cognitive processes, these strategies are not considered for the current review.

The main purpose of the current article is to review for what ages the effectiveness of various GLSs has been demonstrated and whether there are indications of age-related differences in their effectiveness. I will review studies that tested school-age students on one of the following six strategies: concept mapping, explaining, predicting, questioning, testing, and drawing. Note that this review is restricted to GLSs in which students have to generate individually; collaborative generative activities are not discussed. This choice does not imply that collaborative generation is necessarily less effective than individual generation. Rather, it reflects the focus of this review on developmental differences, which can be expected to be more clearly identifiable for activities that are localized within individual students as compared with those that depend heavily on student–student or teacher–student interactions. This review further does not include studies performed with children younger than school age. This is done to enhance comparability between studies because many GLSs require basic language abilities and because many of the studies discussed were performed in regular classrooms.

To lay the ground for the discussion of age-related differences, I begin by providing a concise summary of developmental psychological research on the use of learning strategies

from childhood through emerging adulthood. This summary shows why it is plausible to assume that there are developmental differences in the effectiveness of GLSs. The bulk of this article then reviews six popular GLSs and discusses research testing their effectiveness in improving students' acquisition of declarative knowledge.¹ The review of each strategy covers (1) how it is supposed to work, (2) the evidence on its effectiveness in different age groups, and (3) if there are developmental differences in its effectiveness. In a nutshell, it is found that, while all six generative learning strategies reviewed have proven effective for university students, evidence is mixed for younger students (see Table 1). In particular for elementary-school children, the techniques seem to differ strongly in their effectiveness, but there is a lack of age-comparative studies that can explain these differences. I conclude this review by discussing potential reasons for these differences between strategies and with a call for research that tests these ideas.

A Brief Developmental-Science Perspective on Learning Strategies

Research on children's strategy use has a rich history, and children's increasing use of learning strategies as they grow older is considered a key driving force underlying the observed age-related improvements in learning and memory performance (e.g., Bjorklund et al. 2008; Flavell 1970; Pressley and Hilden 2006). Numerous studies have demonstrated that the efficiency with which elaboration and organization strategies can be employed increases substantially across childhood and continues to increase well into adolescence (for overviews, see Bjorklund et al. 2008; Schneider 2015). Many elementary-school children struggle in using learning strategies effectively and efficiently, and this is particularly the case for those strategies that capitalize on activating prior knowledge. The most prominent difficulty that children face is *production deficiency*, that is, difficulty in spontaneously applying an appropriate learning strategy in the first place. In addition, although prompting children to use a particular strategy is often successful, not all children—and particularly not the younger ones—have the cognitive prerequisites to profit from it, even if they receive extensive training. This phenomenon has been dubbed *mediation deficiency*. A third difficulty that has been demonstrated in children is called *utilization deficiency*: Even when young children can spontaneously apply an appropriate learning strategy, it does not help them as much as it helps older children or adults. Collectively, these deficiencies in children's strategy use have been shown to play an important role in the observed age-related improvement in learning performance (see Bjorklund et al. 2008; Schneider 2015).

The age-related increase in ability to use and profit from learning strategies has been linked to three major developmental processes:

1. **Increases in knowledge:** At least during the first two decades of life, learners' age and their amount of world knowledge are closely correlated (e.g., Li et al. 2004). The increase in knowledge during development has been shown to contribute to the increase in use of learning strategies and in their effectiveness across childhood and adolescence (see Bjorklund 1987; Chi and Ceci 1987). In addition to having effects that are directly related

¹ The effects of GLSs on the acquisition of procedural knowledge, albeit also of high importance, are beyond the scope of this review. The declarative-knowledge acquisition targeted in this review can be further considered as enriching rather than as involving conceptual change (see Chi 2008).

Table 1 Evidence on the effectiveness of generative learning strategies in different age groups

	University students	Secondary-school students	Fourth/fifth-grade students	Below fourth-grade students
Concept mapping	Favorable	Favorable	Favorable	Insufficient
Explaining	Favorable	Favorable	Mixed	Mixed
Predicting	Favorable	Favorable	Favorable	Favorable
Questioning	Favorable	Mixed	Mixed	Unfavorable
Testing	Favorable	Favorable	Favorable	Favorable
Drawing	Favorable	Favorable	Unfavorable	Unfavorable

to particular learning strategies, increasing world knowledge has strong indirect effects. Greater world knowledge facilitates elaboration and organization of to-be-learned material because it provides a richer semantic network, which enables learners to better relate new material to known concepts (Schneider 1993). Bjorklund (1987) pointed to an additional mechanism by which increases in world knowledge facilitate learning: They free up cognitive capacities, which can then be invested in elaborating the new information or in applying more elaborate learning strategies. Although knowledge strongly impacts the use and effectiveness of learning strategies, developmental studies that have carefully controlled for age-related differences in prior knowledge have found that it is not the sole source of age-related differences in learning performance, but that increases in cognitive capacities play a role as well (e.g., Brod et al. 2017; Hasselhorn 1990).

- Increases in cognitive capacities: The increase in use of learning strategies over middle childhood and adolescence coincides with an increase in working memory and inhibitory capacities, which are closely linked to the ongoing maturation of the prefrontal cortex (Best and Miller 2010; Schneider 2015). Working memory and inhibition together facilitate efficient shifting between tasks or mental states (Best and Miller 2010). These three abilities, collectively termed executive functions (Miyake et al. 2000), underlie the ability to reason as well as to test and revise mental models, which are crucial for the acquisition of complex concepts (Bascandziev et al. 2018; Brod et al. 2019). Children's deficits in executive functions can therefore be expected to impinge upon the effectiveness of all GLSs, as the construction of relations between to-be-learned information and prior knowledge, as well as the integration of the new information, requires at least basic reasoning abilities (Zaitchik et al. 2014).
- Increases in metacognitive abilities: Children's ability to accurately represent and regulate their current cognitive activities has strong effects on how effectively they can make use of the learning strategies about which they have knowledge (for a review, see Schneider 2010). These metacognitive skills are called monitoring and control (or regulation). They follow a late-maturing developmental trajectory similar to that of executive functions, and depend on but are not fully determined by them (Roebbers 2017). Although increases in knowledge and cognitive capacity may be necessary for applying learning strategies in the first place, understanding how a particular strategy can lead to higher learning success may be necessary for its effective use. Using learning strategies effectively also entails accurate monitoring and, if necessary, adaptation of the strategy.

To conclude, developmental research indicates that there are substantial age-related increases in the use of learning strategies, such as GLSs, that capitalize on activation of prior knowledge.

Many elementary-school children struggle to use learning strategies effectively and efficiently even if they receive specific training on the strategies. Although this picture looks bleak, one can look at the problem from a different angle: Children's immature strategy use suggests that any GLS that works for elementary-school children should have an even stronger beneficial effect for them than for secondary-school children and adults, who spontaneously use learning strategies anyhow. However, important questions remain. Are some GLSs more effective for children than others are? Are there GLSs that do not work at all until a particular level of cognitive capacities is reached? Can GLSs currently implemented in the classroom be modified in order to work better for elementary-school children? In the following section, I try to answer these questions by reviewing the literature on popular GLSs and taking into account the developmental-science perspective. A particular emphasis is thus laid on studies that tested these techniques with children of various ages.

Review of Techniques

I review six popular GLSs that are used to improve students' learning: generating concept maps, generating explanations, generating predictions, generating questions, generating answers (i.e., practice testing), and generating drawings. Each review consists of three sections: The first section briefly describes how the technique is commonly implemented and how it is supposed to improve students' learning. The second section reviews the available evidence regarding the technique's effectiveness in improving learning in various age groups. To provide a balanced evaluation of the evidence base, I relied on existing meta-analyses whenever possible. The third section reviews the literature concerning age-related differences in the strategy's effectiveness. While this review is clearly influenced by the evaluation provided in the second section, the third section particularly reviews studies that tested learners of various ages. It, thus, aims at evaluating evidence for an effectiveness trajectory of a particular strategy.

Generating Concept Maps

How Is it Supposed to Work?

Concept maps depict the hierarchy of and relations among concepts (for a toy example of a concept map, see Fig. 1). This technique is based on Ausubel's (1960) assimilation theory of cognitive learning and was developed by Novak and colleagues as an instructional tool to represent knowledge structures and to promote meaningful learning (e.g., Novak 1990). According to Ausubel, meaningful learning can be achieved by anchoring new ideas or concepts with previously acquired knowledge. Concept maps thus serve the goal of activating relevant prior knowledge and thereby providing an "optimal anchorage for the learning material" (Ausubel 1960, p. 271).

Concept maps come in various forms and can be used in various ways (for an overview, see O'Donnell et al. 2002). They can range from being entirely pre-generated (typically by the teacher) to being entirely student generated. Because the focus of this review is on generation by students, it is limited to studies in which students at least partly generated the maps themselves; that is, they either modified a map they were given or generated the map entirely by themselves. Furthermore, in line with the constructivist view of learning, meta-analyses suggest that self-generating a concept

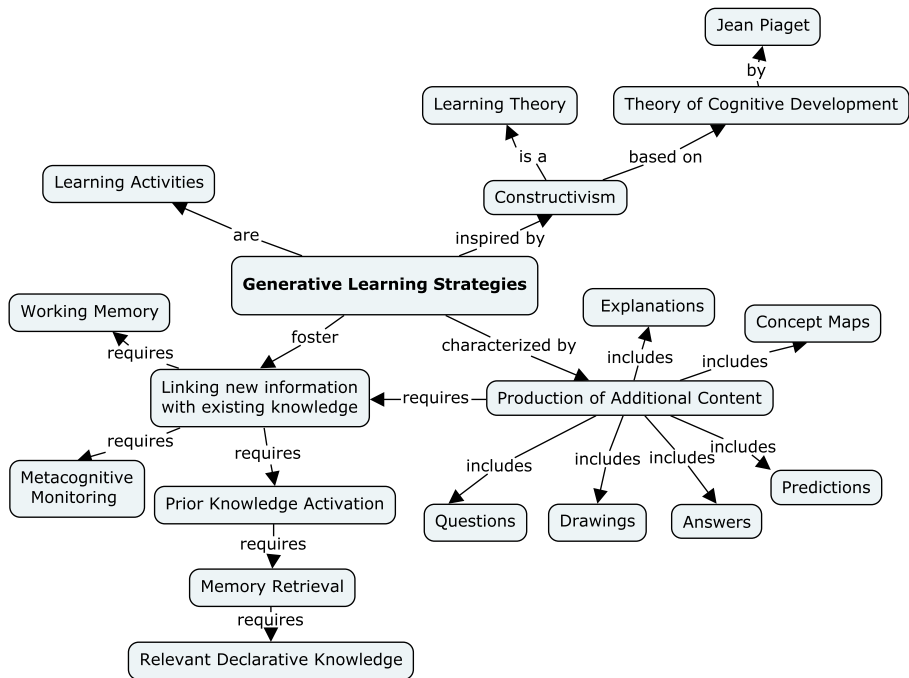


Fig. 1 Example of a simple concept map about generative learning strategies

map is more effective than studying a provided concept map (Nesbit and Adesope 2006; Schroeder et al. 2018; but see Horton et al. 1993). The implementation of this technique also varies in when it is employed. In the tradition of Ausubel (1960), concept maps are often used as *advance organizers*, intended to activate relevant prior knowledge so as to prime students for learning new information. Additional uses, not targeted by this review, are as an activity-closing summary and as an evaluation to test students' knowledge.

For What Ages Has Its Effectiveness Been Demonstrated?

Several reviews and meta-analyses indicate that generating concept maps is generally beneficial for learning (e.g., Horton et al. 1993; Nesbit and Adesope 2006; Novak 1990). The most recent and comprehensive meta-analysis by Schroeder et al. (2018), which clearly distinguished self-generated and provided concept maps, revealed that the mean size of the effect of self-generated concept mapping on students' achievement scores was 0.72 SDs. As can be expected, this beneficial effect was greater for studies that compared generating concept maps to passive activities such as hearing a lecture (1.05 SDs) than for studies that included active control conditions such as summarizing (0.48 SDs). The meta-analysis also looked at whether children's grade (a proxy for age) influenced the effectiveness of having them generate concept maps. Results indicated consistent beneficial effects across studies performed with 4th–8th graders² (0.68 SDs), 9th–12th graders (0.74 SDs), and college students (0.73 SDs).

² A closer examination of these studies, treating studies with 4th and 5th graders and studies with 6th–8th graders separately, indicated no differences between these groups either.

The youngest children in the studies included in the meta-analyses were fourth-graders. What about children below fourth grade? Novak (1990, p. 37) writes that, based on his experience, concept maps are successful only in children from grade 4 onwards, but no empirical data were provided to bolster this claim. Furthermore, generating concept maps that contain written language is, of course, constrained by children's ability to write, which makes this technique unlikely to work in young children. In summary, results indicate that there is good evidence that letting students generate concept maps is generally effective from grade 4 onwards, while there is a clear lack of evidence regarding its effectiveness in younger students.

Are There Age-Related Differences in Its Effectiveness?

Concept maps can be implemented in various ways and with varying instructional support, which may impact their effectiveness in different age groups. Direct evidence for this idea has been provided in a study comparing concept mapping among college and high-school students, which revealed a map coherence \times age group interaction (Gurlitt and Renkl 2008). University students profited more when they received a low-coherence map that required them to create and label lines and thus called for more self-organization of the material being learned, whereas high-school students profited more when they received a high-coherence map that required only labeling existing lines, and thus provided a highly structured learning environment and more support for the learner. Differences in prior knowledge between the groups were not assessed in the present study, but are likely to have contributed to the observed interaction. Nevertheless, the study strongly indicates that learners of different ages need different amounts of support for concept-map generation to be maximally effective.

Further support for this conclusion was provided in a study with late-elementary-school children that varied the level of support provided during concept mapping (Karpicke et al. 2014). Children in the no-support condition (experiment 1) had to generate the entire concept map by themselves, which led to poor concept maps and did not benefit their performance on a later test in which they had to retrieve key concepts of the target material. Children who only had to fill in parts of an existing concept map (experiments 2 and 3) did significantly better than children in the no-support condition and children who did not work on concept maps at all.

To conclude, existing evidence suggests that—among students grade 4 and older—concept-map generation can be similarly effective in students of different ages. However, demands on prior knowledge, working memory, and self-regulation during mapping vary strongly depending on the complexity of the concept and on the support given by pre-existing structures. Initial evidence suggests that a suboptimal implementation of the technique for a particular age group can render it ineffective, and that younger learners generally need more support than older learners. There is a clear need for further studies that systematically vary the amount of information to be generated within age groups and then compare the results across different age groups.

Generating Explanations

How Is It Supposed to Work?

Researchers have suggested that asking students to generate explanations during learning activates relevant prior knowledge and facilitates integration and organization of new information (Chi et al. 1994; Pressley et al. 1992; Renkl et al. 1998). This strategy further encourages elaboration of the new information, such as processing similarities and differences among elements of the to-be-learned content (Dunlosky et al. 2013). Such processing has been

shown to be beneficial for memory (Hunt and Lamb 2001). Furthermore, asking students to provide explanations is supposed to prompt them to generate inferences that go beyond the given information and to revise their mental models (Chi 2000). Generating explanations, thus, requires reasoning abilities in that additional information has to be inferred or deduced on the basis of prior knowledge.

The strategy of student-generated explanations can be implemented in various formats. Most research on this strategy has focused either on elaborative interrogation or on self-explanation. Although both involve answering questions during learning, elaborative interrogation refers to explaining something to other people and focuses on “Why?” questions that are directly related to understanding the phenomenon at hand (see Pressley et al. 1992). Self-explanation refers to explaining something to oneself and includes a much wider array of questions, ranging from simple “Why?” questions, as used in elaborative interrogation, to metacognitive questions about the learning process (Dunlosky et al. 2013). Typically, self-explanation is performed silently, such as during silent reading (e.g., Chi et al. 1989; McNamara 2004). Because the focus of this review is on GLSs, I discuss only those self-explanation studies that used questions intended to prompt reflection on the to-be-learned content. In this case, self-explanation requires cognitive mechanisms similar to those required by elaborative interrogation, which is why I consider the two together. However, it deserves mentioning that explaining something to other people may be more demanding than explaining something to oneself (Pressley et al. 1992).

For What Ages Has Its Effectiveness Been Demonstrated?

Beneficial effects of both elaborative interrogation and self-explanation have been reported for various student groups and under various learning conditions (for an overview, see Dunlosky et al. 2013). A recent meta-analysis on self-explanation (Bisra et al. 2018) that did not include unpublished studies revealed a mean effect size of 0.55 SDs when conditions in which participants were instructed to explain content to themselves were compared with conditions in which they were not instructed to do this. Focusing on the domain of mathematics instruction, a recent meta-analysis of the literature on self-explanation prompts revealed a mean effect size of 0.33 SDs for short-term improvement in conceptual knowledge (Rittle-Johnson et al. 2017). However, less is known about the persistence of these effects, and only few studies equated time on task (cf. Rittle-Johnson et al. 2017). When considering only studies in which time on task was equated, the meta-analysis by Bisra et al. (2018) found an effect size of 0.41 SDs.

Bisra et al. (2018) reported a slightly reduced effect size of this strategy for elementary-school students (0.48 SDs, based on 10 studies) and high-school students (0.43 SDs, based on 13 studies) compared with university students (0.61 SDs). Most of the studies included in this comparison did not control for time on task, however, which indicates that effect sizes are overestimated. In the meta-analysis by Rittle-Johnson and colleagues (2017) on mathematics instruction, none of the studies that were performed with elementary-school children revealed lasting beneficial effects on conceptual knowledge. In summary, while asking students to generate explanations has proven beneficial across a wide variety of learning situations, evidence is mixed as to how much elementary-school children benefit from it.

Are There Age-Related Differences in Its Effectiveness?

Because generating an explanation for novel phenomena requires at least some reasoning abilities, it could be expected to disadvantage younger students. In short, it is not before around

age 9 that children are consistently able to generate inferences based on deep structural similarity (Gentner and Toupin 1986), and this inability should hamper their ability to generate explanations based on underlying properties of to-be-learned material. However, laboratory experiments with children as young as 3 years of age (e.g., Legare and Lombrozo 2014) have demonstrated a beneficial effect of this strategy on conceptual understanding. Furthermore, in classroom studies with elementary-school children, beneficial effects of explanation have repeatedly been observed (e.g., Pine and Messer 2000; for overviews, see Dunlosky et al. 2013; Pressley et al. 1992).

Turning to studies that included a wider age range, Rittle-Johnson (2006) compared effects of self-explanation and direct instruction in a sample of third-, fourth-, and fifth-graders and found comparable improvements in conceptual knowledge for the two instructional methods. In a sample of second-, third-, and fourth-graders, Mceldoon et al. (2013) compared a self-explanation condition with two conditions in which students had to solve additional problems instead of generating explanations. In one of these comparison conditions, time on task was equated by letting children solve additional practice problems, and in the other comparison condition, it was not. Whereas the self-explanation condition yielded larger gains in conceptual knowledge than the non-time-equated control condition, there was no difference in gains between the self-explanation condition and the time-equated control condition. It deserves mentioning that solving additional practice problems can be considered a generative activity as well and, thus, constitutes a rigorous control condition. One could, thus, interpret this finding as providing some evidence suggesting that generating explanations can be effective relative to non-generative activities already in elementary school children.

Turning to research on elaborative interrogation, a study in fourth- to eighth-graders (Wood et al. 1990) reported beneficial effects of elaborating on “Why?” questions related to a text relative to control conditions in which children just read the text or were provided with elaborative answers. While the study found a strong link between students’ benefit from elaborative interrogation and the quality of their generated answer, which is likely higher in older students, it did not analyze potential age-related differences. Studies that investigated younger children did not find benefits of elaborative interrogation (Miller and Pressley 1989; Wood et al. 1993), which led Pressley et al. (1992, p. 103) to conclude that the beneficial effects of elaborative interrogation likely increase from early childhood to adulthood, and that this might be due to age-related increases in the extent and accessibility of prior knowledge. In conclusion, the available literature suggests an age-related increase in the benefit from generating explanations. This explanation remains clearly speculative, however, because of the lack of research directly comparing results obtained with the same task in different age groups.

Generating Predictions

How Is It Supposed to Work?

In this technique, teachers ask learners to generate a prediction (often called hypothesizing in science classes) about a specific fact or outcome before providing them with the to-be-learned information. Generating a prediction requires accessing prior knowledge and connecting it to the new information being learned (Schmidt et al. 1989). Furthermore, generating a prediction may stimulate curiosity for the correct answer (Brod and Breitwieser 2019; Potts et al. 2019) and if the correct answer is different from the prediction, learners experience surprise (Brod

et al. 2018). Both curiosity and surprise are epistemic emotions that are supposed to lead to increased attention to the to-be-learned information, which strengthens learning (D’Mello et al. 2014). Whereas generating a prediction requires retrieving prior knowledge, learning from a prediction requires processing feedback on that prediction (Vaughn and Rawson 2012). Monitoring feedback on a prediction has been shown to require at least basic executive-function abilities (Brod et al. 2019).

For What Ages Has Its Effectiveness Been Demonstrated?

Asking students to generate predictions has been successful in classroom studies that have investigated ways to improve students’ learning in various fields of study, including learning from text (Fielding et al. 1990; Palinscar and Brown 1984), physics (Champagne et al. 1982; Inagaki and Hatano 1977; Liew and Treagust 1995), and biology (Schmidt et al. 1989). In addition, asking students to generate predictions is a component of teaching with audience response systems, which have been shown to enhance students’ engagement in university lectures and learning from them (e.g., Crouch et al. 2004). However, these classroom studies have often used generating predictions along with other tasks that were intended to foster learning, which makes it difficult to tell how much of the observed benefit was uniquely due to predicting. A recent laboratory study that examined the specific effects of generating predictions among university students revealed that this strategy led to greater learning of geography than did another generative activity (Brod et al. 2018). The lack of a control condition that did not engage in generative activities means that the present study does not allow an assessment of the effectiveness of generating predictions relative to more passive learning, however. This heterogeneity in studies might further explain why there is no meta-analysis available that would allow drawing conclusions regarding the general effectiveness of generating predictions.

In contrast to other GLSs, however, many of the studies on the effects of generating predictions have been performed with children and adolescents. Inagaki and Hatano (1977) taught fourth-graders about the conservation law in physics and observed beneficial effects of letting them generate predictions first. Similar beneficial effects were revealed in studies on text comprehension (Fielding et al. 1990; Palinscar and Brown 1984), which were conducted with third-graders and seventh-graders, respectively. Brod and colleagues (Brod et al. 2019) tested third- to fifth-graders on a belief-revision task and found that a prediction-generation condition improved belief revision compared with a no-generation control condition. Similarly, a facts-learning study with second-graders revealed memory benefits of generating a prediction before seeing the correct answer relative to control conditions without the prior prediction (Marsh et al. 2012). In summary, asking students to generate predictions has proven beneficial across a wide variety of ages and learning situations relative to passive control conditions. However, the lack of meta-analyses means that the robustness of the effect as well as the role of potential moderators such as prior knowledge or learning material is currently unclear.

Are There Age-Related Differences in Its Effectiveness?

Metcalf and Kornell (2007) compared sixth-graders and college students on a definition-learning task in which children had to generate (mostly incorrect) definitions before seeing the correct one. Similarly for both age groups, this prediction condition yielded better learning than control conditions in which the correct definition was presented at the same time as the word or in which no correct definition was presented (i.e., no feedback on the correctness of the prediction). A recent study (Breitwieser and Brod 2020) compared fourth- and fifth-graders

and university students' learning of trivia facts under two conditions; participants had to generate either a prediction about the fact or an example related to the fact. Although these two GLSs were similarly effective for the university students, the children clearly gained more from generating predictions than from generating examples. The present study's findings suggest that, for children, the benefits of generating predictions extend beyond prior knowledge activation. However, the lack of a control group that did not engage in generative activities means that the present study does not allow an assessment of differences between the age groups in the benefit of generating predictions relative to a passive control condition.

While the results of these studies offer initial evidence against age-related differences in the effectiveness of generating predictions, they also suggest that processing of feedback on a prediction is crucial for the success of this method. There is a wealth of studies suggesting that feedback processing increases across childhood and that this increase is related to developing executive functions (Crone et al. 2004; Zelazo et al. 1996). Support for this conjecture comes from a study in which preschool to third-grade children either had to study complete word pairs or they had to guess the second word of the pair based on the first (Cameiro et al. 2018). It was shown that the benefit of guessing over studying increased with age, and that this was partly due to a decrease in interference from wrongly guessed words. With increasing age, children were better able to use corrective feedback and inhibit their initial guess. In summary, while generating predictions has been shown to be effective already in lower elementary grades, it is tempting to speculate that there still is an age-related increase in effectiveness because of increases in feedback processing. However, the currently available evidence is insufficient to draw such a conclusion.

Generating Questions

How Is It Supposed to Work?

Asking a good question requires accessing and elaborating relevant prior knowledge, and a key instructional goal of this technique is to help learners identify gaps in their knowledge (for overviews, see Graesser et al. 1992; King 1992). This technique can be implemented either in an interactive context, such as in reciprocal teaching (Palinscar and Brown 1984) or guided questioning of peers (King 1994), or in the form of questioning oneself (Wong 1985). For the purpose of this review, only the latter form of questioning will be considered. It has been suggested that asking learners to generate questions on the taught content also guides their attention to the main ideas and helps them and the instructor monitor their current state of understanding (Palinscar and Brown 1984). The quality of the questions asked are both a viable indicator of learners' knowledge as well as of their evaluation of their current state of understanding (Graesser and Olde 2003). Thus, in order to generate a good question, metacognitive abilities are needed. By teaching students to generate questions, it is argued, one is also teaching them a self-regulatory cognitive strategy that helps them to learn by themselves (Garcia and Pearson 1990; Scardamalia and Bereiter 1985).

For What Ages Has Its Effectiveness Been Demonstrated?

The overwhelming majority of studies on question generation have been carried out in the domain of text comprehension. A review of studies that examined the effects of generating questions in this domain (Wong 1985) indicated that generating questions had a beneficial

effect if students were properly instructed on how to do it beforehand and if they were given sufficient time to do so. A meta-analysis of intervention studies in which students were first taught to generate questions either during or after reading a text (Rosenshine et al. 1996) revealed a positive overall effect on comprehension tested at the end of the intervention (0.35 SDs for standardized tests and 0.82 SDs for non-standardized tests). In their meta-analysis, Rosenshine et al. (1996) qualitatively reviewed the effects at different grade levels of the students who participated in the studies. This review indicated that college students consistently showed positive effects of question-generation training (but see Hoogerheide et al. 2019), but only one out of four studies that tested third-graders showed such an effect. Results were mixed for fourth- through ninth-graders. In line with this meta-analysis, a large study including a question training manipulation in third-grade science and math units found no benefit of the question training for learning outcomes (Souvignier and Kronenberger 2007). These results suggest that third-graders are unable to profit from generating questions even if they have been intensively trained to do so.

For older children, the mixed evidence indicates that generating questions might work provided sufficient instructional scaffolding is provided by the teachers. An exemplary study with fourth- and fifth-graders (King 1994) found that providing them with generic questions or question stems led to beneficial effects of questioning, but unguided questioning did not. It seems uncertain, though, whether the guided conditions truly involved question generation given that the learners were provided with generic questions that they only had to apply. In an exemplary study with ninth-graders, training them to pose questions for themselves during a classroom lecture proved effective in promoting their comprehension of and memory for the lecture (King 1991).

In summary, there is sufficient evidence to conclude that generating questions can be an effective learning strategy in university students, while it is likely not effective in younger elementary school students. For the intermediate ages, evidence is mixed, suggesting that generating questions might work if sufficient instructional scaffolding is provided by the teachers.

Are There Age-Related Differences in Its Effectiveness?

The evidence discussed in the last section already strongly suggests the existence of age-related differences in the effectiveness of generating questions. One study has examined age-related differences in the effectiveness of student-generated questions directly (Denner and Rickards 1987). It compared text comprehension in fifth-, eighth-, and eleventh-graders and found that self-generated questions were less effective than provided questions and not more effective than rereading in promoting comprehension across all age groups. However, there was an increase in effectiveness of self-generated questions relative to rereading as well as in the number of conceptual questions with age, which the authors attribute to older students' better metacognitive abilities. This interpretation seems plausible given that only among the eleventh-graders did the majority of students direct their questions toward the main ideas in the text. Younger students mainly asked simple factual questions that were often targeting less relevant details in the text. In summary, findings of the present study suggest that the unguided generation of questions by students will result in strong age-related differences. For this method to be effective, especially without intensive training or provision of question stems, high levels of metacognitive abilities and prior knowledge seem necessary, which presents a challenge for children.

Generating Answers (Practice Testing)

How Is It Supposed to Work?

Answering test questions about previously learned material is known by many names—most prominently, *self-testing*, *practice testing*, and *retrieval practice*. The focus on *practice* makes clear that this method qualifies as a GLS because it is not intended as a one-shot summative assessment of learners' knowledge but rather involves repeated activity to enhance their retention of the material (Fiorella and Mayer 2016). According to the rationale for this strategy, attempting to retrieve target information involves memory search processes that activate related prior knowledge (Carpenter 2009). The co-activated knowledge gets bound to the target information, resulting in an elaborated memory trace that facilitates later retrieval of the target because more cues can be used to guide memory search. Attempting to retrieve target information can be effective even when the attempt fails because it helps to structure existing knowledge, thereby facilitating retrieval in the future. Retrieval failures may allow learners to identify ineffective cue–target connections and to shift to more effective ones (Pyc and Rawson 2012). As is generating predictions, self-testing is most effective when corrective feedback is provided (Kang et al. 2007), which means that it requires capability for at least basic feedback monitoring.

For What Ages Has Its Effectiveness Been Demonstrated?

In the past two decades, a plethora of studies have found a strong beneficial effect of practice testing on learning (also called the *testing effect*, for reviews, see Rawson and Dunlosky 2012; Roediger and Butler 2011). A meta-analysis (Rowland 2014) that took into account unpublished studies revealed a moderately strong effect size (0.50 SDs) for studies that compared testing with a restudy control condition. Another meta-analysis that also included no-activity control conditions revealed slightly larger effect sizes (i.e., between 0.60 SDs and 0.70 SDs; Adesope and Trevisan 2017). The meta-analyses further suggested that testing is similarly beneficial when performed in the classroom or in the laboratory, as well as for learning verbal and nonverbal material.

Although most of the studies on the effects of practice testing have been conducted with university students, several studies have involved children as well (for a review, see Fazio and Marsh 2019). The meta-analysis by Adesope and Trevisan (2017) found negligible differences in effect size between studies conducted in elementary, secondary, and postsecondary students, indicating that testing can be successful across a wide age range. Moreover, the beneficial effects of practice testing seem to be independent of learners' prior knowledge (Adesope and Trevisan 2017; Carroll et al. 2007). In summary, the existing evidence indicates that practice testing is an effective strategy in students of all ages.

Are There Age-Related Differences in Its Effectiveness?

Turning to studies that compared different age groups, Lipowski et al. (2014) tested first- and third-graders and found robust beneficial effects of testing compared with restudying. While the benefit of testing compared with restudying was stronger in third-graders than in first-graders, the interaction did not reach significance. A significant interaction between age and condition (testing, restudying) was found in a study that compared lower elementary, upper

elementary, and college students on the “forward testing effect” (Aslan and Bäuml 2016). In this variant of the common testing effect, testing is found to enhance learning of subsequent information. While this was the case in college students and, to a lesser extent, in upper elementary students, lower elementary students did not profit from testing. The authors attribute this age-related increase to younger children’s difficulties in combating interference from the previously studied information, which is related to the development of executive functions. In summary, while most studies indicate that testing is an effective strategy early on, recent results point to an additional age-related increase in effectiveness. Future research involving different age groups is certainly necessary to bolster this claim, however.

Generating Drawings

How Is It Supposed to Work?

Asking learners to draw an illustration that looks like or corresponds to a studied concept has been investigated mainly in the context of learning from instructional texts (but see Ploetzner and Fillisch 2017). To draw an illustration, learners have to translate the verbal information into a picture that represents spatial relationships among the elements mentioned in the text (Alesandrini 1984; Schwaborn et al. 2010). Drawings are different from concept maps in that the latter do not physically resemble the concepts they depict (Van Meter and Garner 2005).

Grounded in Mayer’s *generative theory of textbook design* (Mayer et al. 1995), Van Meter and Garner (2005) proposed that this method requires learners to engage in three cognitive processes: selecting the relevant information from the text, organizing the selected information to build up an internal verbal representation, and constructing an internal nonverbal representation that corresponds to the verbal one. Activating relevant prior knowledge, such as suitable nonverbal representations, and connecting it with the new information is crucial during all three steps. In addition, the task involves analogical reasoning in that a good drawing has to be structurally similar to the text, which means that learners have to constantly compare the correspondence between the text and their drawing. A specific feature of drawings is the integration of the verbal and nonverbal representational domains. The translation from the verbal to the nonverbal representation involves metacognitive processes such as monitoring and regulation, as learners have to go back and forth between the two representations (Schwaborn et al. 2010; Van Meter and Garner 2005).

For What Ages Has Its Effectiveness Been Demonstrated?

Research has by and large revealed positive effects of learner-generated drawing, but findings have been inconsistent (for reviews, see Alesandrini 1984; Leutner and Schmeck 2014; Van Meter and Garner 2005). A recent review/meta-analysis (Fiorella and Zhang 2018) that included only published studies revealed that drawing was clearly more effective than just reading a text for both text comprehension (0.46 SDs) and transfer performance (0.75 SDs). However, when compared to provided illustrations, drawing was only advantageous when instructional support was very high. This conclusion is supported by a study in which a drawing-only condition was compared with a reading-only condition as well as to two drawing-plus-support conditions (Van Meter 2001). Participants in the condition with the most support, who were prompted to compare their drawing with a provided illustration, generated more accurate drawings, acquired more conceptual knowledge, and engaged in more

beneficial self-monitoring than did participants in the other conditions. In line with this finding, drawing accuracy and self-monitoring processes have repeatedly been found to mediate the beneficial effects of learner-generated drawing (Van Meter and Garner 2005).

The effects of learner-generated drawings have been examined across a wide age range. The aforementioned reviews, however, did not discuss potential age-related differences in the effectiveness of this method. The summary tables of empirical studies on learner-generated drawings provided by Van Meter and Garner (2005) and Fiorella and Zhang (2018) reveal no clear pattern regarding age-related differences either. They rather indicate mixed findings regardless of learners' age, which again points to the importance of examining how drawing was implemented in the different studies. For example, a study with first-graders revealed a benefit relative to provided drawings (Lesgold et al. 1975), but in the present study children only had to assemble cutouts, which changes the nature of the learning activity. A study with fourth- and fifth-grade children that provided only little instructional support (Rasco et al. 1975, Experiment 3) found that children profited less from drawing than from being provided with illustrations.

Studies with slightly older children revealed more promising results, however. For example, eighth-grade biology students who were provided with background parts for their drawings showed better text comprehension than students who only read the text or who read the text and received the complete drawings (Schmeck et al. 2014). In summary, while there is sufficient evidence to conclude that generating drawings can be an effective learning strategy in university students, it is likely not effective in elementary school students unless instructional support is extremely high, thus requiring only limited self-generation.

Are There Age-Related Differences in Its Effectiveness?

While the evidence discussed in the last section already strongly suggests the existence of age-related differences, two studies directly compared the effectiveness of generating drawings between different age groups. One (Van Meter et al. 2006) compared fourth- and sixth-graders in a design similar to that of Van Meter (2001), including a read-only condition as well as three drawing conditions that varied in instructional support. Gains in conceptual knowledge were assessed on a problem-solving task. Only the sixth-graders benefited from drawing, and this benefit was enhanced in the high-support condition, in which students were prompted to compare their drawing with a provided illustration. Contrary to the prediction that the fourth-graders would benefit more from additional support than the sixth-graders would, they did not benefit from drawing even in the high-support condition. This outcome could not be explained by differences in prior knowledge or general comprehension ability, which were comparable between the two age groups. The authors speculated that the provided support was not high enough for fourth-graders even in the high-support condition or that fourth-graders would need additional practice in the technique. The second study comparing the effectiveness of learner-generated drawing between age groups (Van Essen and Hamaker 1990) compared first-, second-, and fifth-graders in an intervention design in which children were instructed to generate drawings to help them solve arithmetic word problems. It was found that, despite extensive practice and instructional support, first- and second-graders' problem solving did not profit from drawing at all, whereas fifth-graders' did. The lack of an active control group makes it difficult, however, to assess how much of this effect in fifth-graders is due to drawing itself and how much of it is due to other effects of the intervention.

To conclude, the available evidence suggests strong age-related differences in the ability to profit from drawing. While elementary-school children struggle to profit from drawing at all, secondary school children can profit only if instructional support is high. Because generating drawings that correspond to a studied concept requires good analogical reasoning and monitoring abilities, this technique's effectiveness can be expected to exhibit a slowly ascending age trajectory that reaches its peak only in early adulthood.

Overall Summary

The first and foremost goal of this review was to evaluate the evidence for the effectiveness of six popular GLSs in improving declarative learning in different age groups. All six techniques reviewed generally work well for university students, which should make them promising for younger students as well. The success of these techniques with university students is unsurprising, however, given that most of the techniques were initially developed and tested at universities. Do they work equally well for younger students? Table 1 is a first attempt toward a general summary of each strategy's level of effectiveness for different age groups. In keeping with the fact that many of the studies assessing them were performed in the classroom, the table is organized by grade level instead of age. Grade levels are combined on the basis of the grade distribution of the available studies, many of which involved upper-elementary-school children (grades 4–5), which is why they can be considered separately.

It bears mention that this assessment for the different age groups is tentative because it often rests on very few studies. Moreover, the conclusions are based predominantly on published studies. Because studies finding no benefit of a particular strategy are less likely to be published (the well-known file-drawer problem; Rosenthal 1979), the table likely overestimates effectiveness. In addition, for several GLSs, very few studies involving children below fourth grade were found. A pragmatic reason for this might be that it is difficult to use written assessments with children in this age range. A conceptual reason might be that GLSs require a great deal of self-regulatory and especially inhibitory abilities, which are relatively slow to mature.

An overall pattern that can be observed is that there is stronger evidence for the effectiveness of GLSs in older students, and that the evidence for their effectiveness in elementary-school children is quite variable. For example, whereas practice testing and predicting seem to be effective already in lower-elementary-school children, generating drawings seems to be largely ineffective until secondary school. These findings also speak to the second goal of this review, which was to examine whether there are age-related differences in the effectiveness of GLSs and whether these differ between strategies. However, the limited amount of age-comparative studies precludes any strong conclusions regarding age-related differences and mechanisms thereof. Therefore, in the final sections of this review, I will discuss theoretical reasons that speak for an age-related increase in effectiveness and differences therein between strategies along with caveats against premature conclusions. This discussion will also open up ways to alleviate some of the difficulties that particularly children face in using GLSs successfully.

Discussion

Why are all six strategies effective for university students, but only some of them for elementary-school children? The second section summarized developmental research indicating that the effectiveness of learning strategies increases substantially across the first two

decades of life. Reasons for this increase have been found in the ongoing development of children's world knowledge as well as of their cognitive and metacognitive capacities. Together, these psychological constructs have been described as key prerequisites for good information processing and have been claimed to underlie age-related increases in effective knowledge acquisition (Pressley et al. 1989). Results of this line of research can explain the overall increase in effectiveness of GLSs with increasing age of the learner. The observed greater effectiveness of GLSs during late-elementary/early-secondary school dovetails with the observation that it is during these grades (i.e., around age 10) that one can observe the emergence of abstract self-reflection in children, which enables an efficient use of knowledge-based learning strategies (e.g., Hasselhorn 1995). Findings are also in line with a recent study that showed that the advantage from giving learners active control over their learning increases across middle childhood (Ruggeri et al. 2019). What this line of research cannot explain straightforwardly, however, are the differences in developmental trajectories between strategies.

To explain differences in trajectories between strategies, it is necessary to consider the strategies' different mechanisms and, consequently, prerequisites. These characteristics have been reviewed in the first section on each GLS ("How is it supposed to work?"). Although all GLSs are supposed to derive some of their effectiveness from activation of prior knowledge, they all have additional and specific assumed mechanisms and prerequisites that—in combination—are different from those of the other strategies. Whereas some strategies seem to require basic reasoning capacities, others rely on good metacognitive monitoring. This observation suggests that differences between the strategies' developmental trajectories may be related to differences between the developmental trajectories of the strategies' prerequisites.

Drawing such a conclusion on the basis of a qualitative inspection of differences in assumed but rarely tested mechanisms seems bold, however. What is needed to bolster this conclusion are empirical studies that establish a link between age-related differences in the effectiveness of particular strategies and developmental trajectories of their prerequisites. Ideally, these studies would be longitudinal in order to allow an analysis of couplings between trajectories (i.e., an increase in one of a strategy's prerequisites should precede an increase in the strategy's effectiveness). To the best of my knowledge, no such study exists.

A less ambitious approach would be to cross-sectionally compare two different GLSs or two different implementations of one GLS between age groups and to examine whether age-related differences in effectiveness are related to differences in prerequisites. A recent study that took just this approach compared the effectiveness of two GLSs in late-elementary-school children and university students (Breitwieser and Brod 2020). Participants were given a learning task in which they generated either predictions or examples before being presented with the correct information (i.e., numerical facts). They were also given tasks to measure their reasoning ability, which has been suggested to be a prerequisite for generating helpful examples. Overall, results revealed that whereas the university students were similarly successful in learning facts in the two GLS conditions, the elementary-school children were more successful when they were asked to generate a prediction than when they were asked to generate an example. The magnitude of this difference in the children was correlated with their reasoning abilities such that the more adult-like their responses (i.e., the more similar their learning performance in the two conditions), the higher their reasoning scores. In summary, these results suggest that different GLSs can be differentially effective for elementary-school children and that this difference is related to the developmental status of the strategies' prerequisites.

To conclude, it seems likely that differences between GLSs in their effectiveness for elementary-school children are related to differences in the developmental trajectories of the strategies' particular prerequisites. Children generally have lower cognitive and metacognitive capacities than adolescents and adults, and this matters more for some GLSs than for others. This point can be illustrated with the GLSs reviewed here, by comparing those strategies that seem to work best for children (i.e., generating predictions and answers) with those strategies that are least effective for children (i.e., generating questions and drawings). Whereas the former strategies require mainly effortful retrieval of prior knowledge, the latter strategies require the construction of diagnostic questions or drawings that can later serve as explicit mediators to facilitate retrieval. It seems that higher levels of prior knowledge and metacognitive monitoring of its relation to the to-be-learned information are necessary for constructing diagnostic questions and drawings than for providing guesses in the form of predictions or answers. Given the lack of studies that have compared mechanisms underlying GLSs, however, these speculations clearly are tentative.

Outlook and Call for Research

This review faced two major difficulties, which also limit the conclusions that can be drawn from it. First, as mentioned previously, only a small number of studies have compared the effectiveness of GLSs for different age groups. Such studies can provide the strongest evidence for age-related differences in the strategies' effectiveness. Second, even fewer studies have investigated the cognitive and metacognitive processes involved in generative learning. Although researchers working on GLSs agree that activation of prior knowledge plays a key role in making generative learning effective, this idea has rarely been tested directly (e.g., by experimentally manipulating the amount of prior knowledge a learner possesses or is able to activate). Similarly, other prerequisites have been assumed on the basis of particular strategies' characteristics but have not been rigorously tested. Such research would be beneficial not only for understanding age-related differences in the strategies' effectiveness, but also for selecting the best strategy for each individual learner and for improving the general effectiveness of GLSs. This section elaborates on these, in my view, most exciting future directions.

Going beyond Age and toward an Individualized Selection of Strategies

This review suggests that it is important for educators to consider age or grade level (which are typically closely correlated) when they want to select the best learning strategies for their students. However, age and grade are not explanatory variables. Eight-year-olds struggle with certain GLSs not because they are 8-year-olds but because they typically lack certain prerequisites that are not fully mature yet. Furthermore, even in elementary-school children, age is clearly only a crude proxy for the amount of world knowledge as well as cognitive and metacognitive capacities that a learner possesses, given that there are substantial individual differences between children of the same age. Age can provide only a rough heuristic for deciding which strategy to use for a particular learner based on typical age trajectories of the strategy's prerequisites. On a more general note, this means that the answer to the question of what is the most effective learning strategy will differ for different learners. A distant goal of this line of research should, thus, be to move toward selecting the best GLS for each learner individually on the basis of that person's levels of learning prerequisites.

Making Generative Learning Work Better for Children

The current review indicates that some GLSs clearly work better than others for elementary-school children. However, it also indicates that the way in which a particular strategy is implemented exerts a strong influence on its effectiveness. For example, two studies (Gurlitt and Renkl 2008; Karpicke et al. 2014) have shown that providing elementary-school and high-school learners with parts of a concept map instead of asking them to generate the full map strongly benefits their learning. Another study (Van Meter et al. 2006) found that sixth-graders' gains in conceptual knowledge were greater when they were prompted to constantly monitor how the drawing they were creating compared with a provided illustration than when they were not given an illustration or when they received an illustration but were not prompted to compare it with their own drawing. In summary, these studies suggest that younger learners—or, indeed, learners with less strong learning prerequisites—need more support during a generative-learning task. In these particular studies, they needed to be provided with some of the material that older learners had to generate themselves, and in the study by Van Meter et al. (2006), they also needed metacognitive prompts that likely compensated for their lower monitoring skills.

Providing learners with parts of the material for a generative-learning task frees up cognitive capacities that can be invested in executing the task and retrieving relevant prior knowledge. Research on the development of learning strategies indicates that this can offset some of the impediments that children typically face in using these strategies (Bjorklund 1987). Providing learners with metacognitive prompts likely serves a similar function, such that the prompt helps them to focus on performing the intended learning activities instead of some other, secondary activities.

In summary, some of the difficulties children face in using GLSs can likely be alleviated by providing extra support and guidance, such as in the form of relevant material or metacognitive prompts. This will work only up to a certain limit, however. If children do not need to engage in effortful retrieval because they are provided with all of the material, the strategy is not generative anymore and unlikely to help their learning.

Conclusion

GLSs are appealing because of their theoretical grounding and evidence of their effectiveness for university students. This review has shown that the techniques differ strongly in their effectiveness for elementary-school children. Some GLSs seem to work already for first-graders, whereas others do not seem to work well before late high school. Research on the development of learning and learning strategies sheds light on how GLSs work and for whom. It has identified prerequisites for optimal use of learning strategies and charted their developmental trajectories. Additional research is clearly necessary, however, to understand how GLSs differ in their mechanisms and, thereby, in their prerequisites. Taking the development of prerequisites into account will help educators to select the best learning strategy for an individual student and to decide which kind of support will be most helpful.

Availability of Data and Material/Code Availability Not applicable.

Funding Open Access funding enabled and organized by Projekt DEAL. G.B. was supported by a Jacobs Foundation Research Fellowship.

Compliance with Ethical Standards

Conflicts of Interest/Competing Interests The author has no conflict of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adesope, O. O., & Trevisan, D. A. (2017). Rethinking the use of tests: a meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659–701. <https://doi.org/10.3102/0034654316689306>.
- Alesandrini, K. L. (1984). Pictures and adult learning. *Instructional Science*, 13(1), 63–77. <https://doi.org/10.1007/BF00051841>.
- Aslan, A., & Bäuml, K. H. T. (2016). Testing enhances subsequent learning in older but not in younger elementary school children. *Developmental Science*, 19(6), 992–998. <https://doi.org/10.1111/desc.12340>.
- Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 51(5), 267–272. <https://doi.org/10.1037/h0046669>.
- Bascandziev, I., Tardiff, N., Zaitchik, D., & Carey, S. (2018). The role of domain-general cognitive resources in children's construction of a vitalist theory of biology. *Cognitive Psychology*, 104(March), 1–28. <https://doi.org/10.1016/j.cogpsych.2018.03.002>.
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>.
- Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: a meta-analysis. *Educational Psychology Review*, 30(3), 703–725. <https://doi.org/10.1007/s10648-018-9434-x>.
- Bjorklund, D. F. (1987). How age changes in knowledge base contribute to the development of children's memory: an interpretive review. *Developmental Review*, 7(2), 93–130. [https://doi.org/10.1016/0273-2297\(87\)90007-4](https://doi.org/10.1016/0273-2297(87)90007-4).
- Bjorklund, D. F., Dukes, C., & Brown, R. D. (2008). The development of memory strategies. In M. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood: second edition* (2nd ed., pp. 157–188). London: Psychology Press. <https://doi.org/10.4324/9780203934654>.
- Bonwell, C. C., & Eison, J. A. (1991). Active learning: creating excitement in the classroom. 1991 ASHE-ERIC Higher Education Reports. In *ERIC Clearinghouse on Higher Education* (pp. 1–121). Washington, DC: The George Washington University. ED340272.
- Breitwieser, J., & Brod, G. (2020). Cognitive prerequisites for generative learning: why some learning strategies are more effective than others. *Child Development. Advance Online Publication*. <https://doi.org/10.1111/cdev.13393>.
- Brod, G., & Breitwieser, J. (2019). Lighting the wick in the candle of learning: generating a prediction stimulates curiosity. *Npj Science of Learning*, 4(1), 18. <https://doi.org/10.1038/s41539-019-0056-1>.
- Brod, G., Lindenberger, U., & Shing, Y. L. (2017). Neural activation patterns during retrieval of schema-related memories: differences and commonalities between children and adults. *Developmental Science*, 20(6), e12475. <https://doi.org/10.1111/desc.12475>.
- Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: the element of surprise. *Learning and Instruction*, 55, 22–31. <https://doi.org/10.1016/j.learninstruc.2018.01.013>.
- Brod, G., Breitwieser, J., Hasselhorn, M., & Bunge, S. A. (2019). Being proven wrong elicits learning in children – but only in those with higher executive function skills. *Developmental Science*, 23(3), e12916. <https://doi.org/10.1111/desc.12916>.
- Carneiro, P., Lapa, A., & Finn, B. (2018). The effect of unsuccessful retrieval on children's subsequent learning. *Journal of Experimental Child Psychology*, 166, 400–420. <https://doi.org/10.1016/j.jecp.2017.09.010>.

- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1563–1569. <https://doi.org/10.1037/a0017021>.
- Carroll, M., Campbell-Ratcliff, J., Murnane, H., & Perfect, T. (2007). Retrieval-induced forgetting in educational contexts: monitoring, expertise, text integration, and test format. *European Journal of Cognitive Psychology*, 19(4–5), 580–606. <https://doi.org/10.1080/09541440701326071>.
- Champagne, A. B., Klopfer, L. E., & Gunstone, R. F. (1982). Cognitive research and the design of science instruction. *Educational Psychologist*, 17(1), 31–53. <https://doi.org/10.1080/00461528209529242>.
- Chi, M.T.H. (2000). Self-explaining expository texts: the dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah: Lawrence Erlbaum Associates. <https://doi.org/10.1177/1046878102238607>, 33, 4.
- Chi, M. T. H. (2008). Three types of conceptual change: belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Hillsdale: Erlbaum.
- Chi, M. T. H. (2009). Active-constructive-interactive: a conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>.
- Chi, M. T. H., & Ceci, S. J. (1987). Content knowledge: its role, representation, and restructuring in memory development. *Advances in Child Development and Behavior*, 20(C), 91–142. [https://doi.org/10.1016/S0065-2407\(08\)60401-2](https://doi.org/10.1016/S0065-2407(08)60401-2).
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182. [https://doi.org/10.1016/0364-0213\(89\)90002-5](https://doi.org/10.1016/0364-0213(89)90002-5).
- Chi, M. T. H., de Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanation improves understanding. *Cognitive Science*, 18(3), 439–477. https://doi.org/10.1207/s15516709cog1803_3.
- Crone, E. A., Richard Jennings, J., & Van Der Molen, M. W. (2004). Developmental change in feedback processing as reflected by phasic heart rate changes. *Developmental Psychology*, 40(6), 1228–1238. <https://doi.org/10.1037/0012-1649.40.6.1228>.
- Crouch, C., Fagen, A. P., Callan, J. P., & Mazur, E. (2004). Classroom demonstrations: learning tools or entertainment? *American Journal of Physics*, 72(6), 835–838. <https://doi.org/10.1119/1.1707018>.
- D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, 29, 153–170. <https://doi.org/10.1016/j.learninstruc.2012.05.003>.
- Denner, P. R., & Rickards, J. P. (1987). A developmental comparison of the effects of provided and generated questions on text recall. *Contemporary Educational Psychology*, 12(2), 135–146. [https://doi.org/10.1016/S0361-476X\(87\)80047-4](https://doi.org/10.1016/S0361-476X(87)80047-4).
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>.
- Fazio, L. K., & Marsh, E. J. (2019). Retrieval-based learning in children. *Current Directions in Psychological Science*, 28(2), 111–116. <https://doi.org/10.1177/0963721418806673>.
- Fielding, L. G., Anderson, R. C., & Pearson, P. D. (1990). How discussion questions influence children's story understanding. *Center for the Study of Reading Technical Report No. 490*. University of Illinois at Urbana-Champaign.
- Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity*. New York: 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>.
- Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review*, 30(3), 1115–1137. <https://doi.org/10.1007/s10648-018-9444-8>.
- Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (5th ed., pp. 181–211). New York: Academic Press. [https://doi.org/10.1016/S0065-2407\(08\)60467-X](https://doi.org/10.1016/S0065-2407(08)60467-X).
- Garcia, G. E., & Pearson, P. D. (1990). Modifying reading instruction to maximize its effectiveness for all students. *Technical Report No. 489*. University of Illinois at Urbana-Champaign.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, 10(3), 277–300. [https://doi.org/10.1016/S0364-0213\(86\)80019-2](https://doi.org/10.1016/S0364-0213(86)80019-2).
- Graesser, A. C., & Olde, B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology*, 95(3), 524–536. <https://doi.org/10.1037/0022-0663.95.3.524>.
- Graesser, A. C., Person, N., & Huber, J. (1992). Mechanisms that generate questions. In T. W. Lauer, E. Peacock, & A. C. Graesser (Eds.), *Questions and information systems* (pp. 167–187). Hillsdale: Lawrence Erlbaum.
- Gurlitt, J., & Renkl, A. (2008). Are high-coherent concept maps better for prior knowledge activation? Differential effects of concept mapping tasks on high school vs. university students. *Journal of Computer Assisted Learning*, 24(5), 407–419. <https://doi.org/10.1111/j.1365-2729.2008.00277.x>.

- Hasselhorn, M. (1990). The emergence of strategic knowledge activation in categorical clustering during retrieval. *Journal of Experimental Child Psychology*, 50(1), 59–80. [https://doi.org/10.1016/0022-0965\(90\)90032-4](https://doi.org/10.1016/0022-0965(90)90032-4).
- Hasselhorn, M. (1995). Beyond production deficiency and utilization inefficiency: mechanisms of the emergence of strategic categorization in episodic memory tasks. In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: issues in growth and development* (pp. 141–159). Hillsdale: Erlbaum.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: a meta-analysis. *Review of Educational Research*, 66(2), 99–136. <https://doi.org/10.3102/00346543066002099>.
- Hoogerheide, V., Staal, J., Schaap, L., & van Gog, T. (2019). Effects of study intention and generating multiple choice questions on expository text retention. *Learning and Instruction*, 60, 191–198. <https://doi.org/10.1016/j.learninstruc.2017.12.006>.
- Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95–111. <https://doi.org/10.1002/sce.3730770107>.
- Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1359–1366. <https://doi.org/10.1037/0278-7393.27.6.1359>.
- Inagaki, K., & Hatano, G. (1977). Amplification of cognitive motivation and its effects on epistemic observation. *American Educational Research Journal*, 14(4), 485–491. <https://doi.org/10.3102/00028312014004485>.
- 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>.
- Karpicke, J. D., Blunt, J. R., Smith, M. A., & Karpicke, S. S. (2014). Retrieval-based learning: the need for guided retrieval in elementary school children. *Journal of Applied Research in Memory and Cognition*, 3(3), 198–206. <https://doi.org/10.1016/j.jarmac.2014.07.008>.
- King, A. (1991). Improving lecture comprehension: effects of a metacognitive strategy. *Applied Cognitive Psychology*, 5(4), 331–346. <https://doi.org/10.1002/acp.2350050404>.
- King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist*, 27(1), 111–126. https://doi.org/10.1207/s15326985Sep2701_8.
- King, A. (1994). Guiding knowledge construction in the classroom: effects of teaching children how to question and how to explain. *American Educational Research Journal*, 31(2), 338–368. <https://doi.org/10.3102/00028312031002338>.
- Legare, C. H., & Lombrozo, T. (2014). Selective effects of explanation on learning during early childhood. *Journal of Experimental Child Psychology*, 126, 198–212. <https://doi.org/10.1016/j.jecp.2014.03.001>.
- Lesgold, A. M., Levin, J. R., Shimron, J., & Guttman, J. (1975). Pictures and young children's learning from oral prose. *Journal of Educational Psychology*, 67(5), 636–642. <https://doi.org/10.1037/0022-0663.67.5.636>.
- Leutner, D., & Schmeck, A. (2014). The generative drawing principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (second ed., pp. 433–448). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.022>.
- Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science*, 15(3), 155–163. <https://doi.org/10.1111/j.0956-7976.2004.01503003.x>.
- Liew, C. W., & Treagust, D. F. (1995). A predict-observe-explain teaching sequence for learning about students' understanding of heat and expansion of liquids. *Australian Science Teachers Journal*, 41(1), 68–71.
- Lipowski, S. L., Pyc, M. A., Dunlosky, J., & Rawson, K. A. (2014). Establishing and explaining the testing effect in free recall for young children. *Developmental Psychology*, 50(4), 994–1000. <https://doi.org/10.1037/a0035202>.
- Marsh, E. J., Fazio, L. K., & Goswick, A. E. (2012). Memorial consequences of testing school-aged children. *Memory*, 20(8), 899–906. <https://doi.org/10.1080/09658211.2012.708757>.
- Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development*, 43(1), 31–41. <https://doi.org/10.1007/BF02300480>.
- Mceldoon, K. L., Durkin, K. L., & Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional practice. *British Journal of Educational Psychology*, 83(4), 615–632. <https://doi.org/10.1111/j.2044-8279.2012.02083.x>.
- McNamara, D. S. (2004). SERT: self-explanation reading training. *Discourse Processes*, 38(1), 1–30. https://doi.org/10.1207/s15326950dp3801_1.
- Metcalfe, J., & Kornell, N. (2007). Principles of cognitive science in education: the effects of generation, errors, and feedback. *Psychonomic Bulletin and Review*, 14(2), 225–229. <https://doi.org/10.3758/BF03194056>.

- Miller, G. E., & Pressley, M. (1989). Picture versus question elaboration on young children's learning of sentences containing high- and low-probability content. *Journal of Experimental Child Psychology*, 48(3), 431–450. [https://doi.org/10.1016/0022-0965\(89\)90051-9](https://doi.org/10.1016/0022-0965(89)90051-9).
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>.
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: a meta-analysis. *Review of Educational Research*, 76(3), 413–448.
- Novak, J. D. (1990). Concept mapping: a useful tool for science education. *Journal of Research in Science Teaching*, 27(10), 937–949. <https://doi.org/10.1002/tea.3660271003>.
- O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review*, 14(1), 71–86. <https://doi.org/10.1023/A:1013132527007>.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117–175. https://doi.org/10.1207/s1532690xcio102_1.
- Pine, K. J., & Messer, D. J. (2000). The effect of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction*, 18(1), 35–51. https://doi.org/10.1207/S1532690XCI1801_02.
- Ploetzner, R., & Fillisch, B. (2017). Not the silver bullet: learner-generated drawings make it difficult to understand broader spatiotemporal structures in complex animations. *Learning and Instruction*, 47, 13–24. <https://doi.org/10.1016/j.learninstruc.2016.10.002>.
- 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>.
- Pressley, M., & Hilden, K. (2006). Cognitive strategies. In *Handbook of Child Psychology*. <https://doi.org/10.1002/9780470147658.chpsy0212>.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: what it is and how education can promote it. *International Journal of Educational Research*, 13(8), 857–867. [https://doi.org/10.1016/0883-0355\(89\)90069-4](https://doi.org/10.1016/0883-0355(89)90069-4).
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27(1), 91–109. <https://doi.org/10.1207/s15326985ep2701>.
- Pyc, M. A., & Rawson, K. A. (2012). Why is test-restudy practice beneficial for memory? An evaluation of the mediator shift hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 737–746. <https://doi.org/10.1037/a0026166>.
- Rasco, R. W., Tennyson, R. D., & Boutwell, R. C. (1975). Imagery instructions and drawings in learning prose. *Journal of Educational Psychology*, 67(2), 188–192. <https://doi.org/10.1037/h0077014>.
- Rawson, K. A., & Dunlosky, J. (2012). When is practice testing most effective for improving the durability and efficiency of student learning? *Educational Psychology Review*, 24(3), 419–435. <https://doi.org/10.1007/s10648-012-9203-1>.
- Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: the effects of example variability and elicited self-explanations. *Contemporary Educational Psychology*, 23(23), 90–108. <https://doi.org/10.1006/ceps.1997.0959>.
- Rittle-Johnson, B. (2006). Promoting transfer: effects of self-explanation and direct instruction. *Child Development*, 77(1), 1–15. <https://doi.org/10.1111/j.1467-8624.2006.00852.x>.
- Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: a meta-analysis and instructional design principles. *ZDM Mathematics Education*, 49(4), 599–611. <https://doi.org/10.1007/s11858-017-0834-z>.
- Roebers, C. M. (2017). Executive function and metacognition: towards a unifying framework of cognitive self-regulation. *Developmental Review*, 45, 31–51. <https://doi.org/10.1016/j.dr.2017.04.001>.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: a review of the intervention studies. *Review of Educational Research*, 66(2), 181–221. <https://doi.org/10.3102/00346543066002181>.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>.
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1463. <https://doi.org/10.1037/a0037559>.

- Ruggeri, A., Markant, D. B., Gureckis, T. M., Bretzke, M., & Xu, F. (2019). Memory enhancements from active control of learning emerge across development. *Cognition*, 186(June 2018), 82–94. <https://doi.org/10.1016/j.cognition.2019.01.010>.
- Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills, Research and open questions* (Vol. 2, pp. 563–577). Hillsdale: Erlbaum.
- Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, 39(4), 275–286. <https://doi.org/10.1016/j.cedpsych.2014.07.003>.
- Schmidt, H. G., De Volder, M. L., De Grave, W. S., Moust, J. H. C., & Patel, V. L. (1989). Explanatory models in the processing of science text: the role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology*, 81(4), 610–619. <https://doi.org/10.1037/0022-0663.81.4.610>.
- Schneider, W. (1993). Domain-specific knowledge and memory performance in children. *Educational Psychology Review*, 5(3), 257–273.
- Schneider, W. (2010). Metacognition and memory development in childhood and adolescence. In H. S. Water & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 54–81). New York: Guilford Press.
- Schneider, W. (2015). *Memory development from early childhood through emerging adulthood*. Cham: Springer. <https://doi.org/10.1007/978-3-319-09611-7>.
- Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: a meta-analysis. *Educational Psychology Review*, 30(2), 431–455. <https://doi.org/10.1007/s10648-017-9403-9>.
- Schwaborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology*, 102(4), 872–879. <https://doi.org/10.1037/a0019640>.
- Souvignier, E., & Kronenberger, J. (2007). Cooperative learning in third graders' jigsaw groups for mathematics and science with and without questioning training. *The British Journal of Educational Psychology*, 77(Pt 4), 755–771. <https://doi.org/10.1348/000709906X173297>.
- Van Essen, G., & Hamaker, C. (1990). Using self-generated drawings to solve arithmetic word problems. *Journal of Educational Research*, 83(6), 301–312.
- Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, 93(1), 129–140. <https://doi.org/10.1037/0022-0663.93.1.129>.
- Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: literature review and synthesis. *Educational Psychology Review*, 17(4), 285–325. <https://doi.org/10.1007/s10648-005-8136-3>.
- Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, 31(2), 142–166. <https://doi.org/10.1016/j.cedpsych.2005.04.001>.
- Vaughn, K. E., & Rawson, K. A. (2012). When is guessing incorrectly better than studying for enhancing memory? *Psychonomic Bulletin and Review*, 19(5), 899–905. <https://doi.org/10.3758/s13423-012-0276-0>.
- von Glasersfeld, E. (1983). Learning as constructive activity. In J. C. Bergeron & N. Hescovics (Eds.), *Proceedings of the 5th Annual Meeting of the North American Group of Psychology in Mathematics Education* (Vol. 1, pp. 41–101). Montreal.
- Witrock, M. C. (1985). Teaching learners generative strategies for enhancing reading comprehension. *Theory Into Practice*, 24(2), 123–126. <https://doi.org/10.1080/00405848509543158>.
- Wong, B. Y. L. (1985). Self-questioning instructional research: a review. *Review of Educational Research*, 55(2), 227–268. <https://doi.org/10.3102/00346543055002227>.
- Wood, E., Pressley, M., & Winne, P. H. (1990). Elaborative interrogation effects on children's learning of factual content. *Journal of Educational Psychology*, 82(4), 741–748. <https://doi.org/10.1037/0022-0663.82.4.741>.
- Wood, E., Miller, G., Symons, S., Canough, T., & Yedlicka, J. (1993). Effects of elaborative interrogation on young learners' recall of facts. *The Elementary School Journal*, 94(2), 245–254. <https://doi.org/10.1086/461764>.
- Yiping, L., Abrami, P. C., & D'Apollonia, S. (2001). Small group and individual learning with technology: a meta-analysis. *Review of Educational Research*, 71(3), 449–521. <https://doi.org/10.3102/00346543071003449>.
- Zaitchik, D., Iqbal, Y., & Carey, S. (2014). The effect of executive function on biological reasoning in young children: an individual differences study. *Child Development*, 85(1), 160–175. <https://doi.org/10.1111/cdev.12145>.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, 11(1), 37–63. [https://doi.org/10.1016/S0885-2014\(96\)90027-1](https://doi.org/10.1016/S0885-2014(96)90027-1).