Inventing Prepares Learning Motivationally, but a Worked-out Solution Enhances Learning Outcomes

Inga Glogger (glogger@psychologie.uni-freiburg.de)
Department of Educational and Developmental Psychology, University of Freiburg, Engelbergerstraße 41
D-79085 Freiburg, Germany

Julian Kappich (kappich@psychologie.uni-freiburg.de)
Department of Educational and Developmental Psychology, University of Freiburg, Engelbergerstraße 41
D-79085 Freiburg, Germany

Rolf Schwonke (schwonke@psychologie.uni-freiburg.de)
Department of Educational and Developmental Psychology, University of Freiburg, Engelbergerstraße 41
D-79085 Freiburg, Germany

Lars Holzäpfel (lars.holzaepfel@ph-freiburg.de)
Institute for Mathematical Education, University of Education, Kunzenweg 21
D-79117 Freiburg, Germany

Matthias Nückles (nueckles@ezw.uni-freiburg.de)
Department for Empirical Instruction and School Research, University of Freiburg, Rempartstraße 11
D-79098 Freiburg, Germany

Alexander Renkl (renkl@psychologie.uni-freiburg.de)
Department of Educational and Developmental Psychology, University of Freiburg, Engelbergerstraße 41
D-79085 Freiburg, Germany

Abstract

Solving an open problem as proposed by inventing and productive failure approaches has been shown to prepare learners effectively for subsequent direct instruction even though invented solutions are often suboptimal for the given problems. Inventing can make the learners aware of knowledge gaps (cognitive) and more curious about and interested in the learning contents (motivational effects). However, working on the same problem with a given (optimal) solution helps avoid misconceptions and disorganized knowledge, while providing useful basic knowledge. Therefore, a given solution could be more effective. In an experiment ($N = 42$), we tested to what extent working on an open problem (inventing) versus a solution prepares student teachers for learning strategy evaluation. The inventing group invented criteria to evaluate learning strategies while the worked solution group studied the same problem in a solved, worked-out version. We found differential effects: inventing enhanced knowledge-gap experience, curiosity, and interest. However, studying the worked-out solution enhanced learning outcomes.

Keywords: instruction; invention activities; worked examples; teacher education; learning-strategy assessment.

Introduction

In order to prepare learners for a new topic and to raise their attention as well as curiosity, teachers often address interesting problems in the beginning before they directly instruct learners about the topic. Similarly, there are experimentally tried-and-tested problem-oriented approaches (Schmidt, De Volder, De Grave, Moust, & Patel, 1989) such as inventing problem solutions (Schwartz, Chase, Oppezzo, & Chin, 2011; Schwartz & Martin, 2004) or productive failure at initial problems (Kapur, 2010). These approaches aim at preparing learners for subsequent direct instruction (preparation for future learning, Schwartz & Martin, 2004). However, when preparing a lesson, should a teacher really put such initial problems up for "inventing" or "productive failure" before implementing direct instruction? Is it not more productive to immediately begin with tried-and-tested forms of direct instruction such as example-based learning in order to avoid wasting time when students search for problem solutions that are very hard to find (see Sweller, Kirschner, & Clark, 2007)?

When starting immediately with methods of direct instruction, a problem might arise: learners often process directly presented information only superficially (Berthold & Renkl, 2010), leading to little knowledge acquisition and transfer. Problem-oriented introductions such as invention activities can prepare learners to more deeply process directly presented information. For example, Schwartz and Martin (2004) had learners invent formulas describing four different distributions of pitches around a target. Later, the learners were taught the concept of mean deviation. Schwartz and Martin assumed that inventing creates preparedness for future learning by generating "early forms of knowledge" (p. 132). These early forms of knowledge can then be used to easily assimilate further knowledge.

Invention activities can appear problematic because learners might not generate canonical or even false solutions. According to the IKEA effect – the increased valuation of self-made products (Norton, Mochon & Ariely, 2012) –, these own suboptimal solutions can be valued higher than expert ones. Similarly, research on the continued influence effect (Johnson & Seifert, 1994) suggests that learners tend to stay with their own suboptimal solution instead of taking up the directly instructed canonical one. However, research on productive failure (e.g., Kapur, 2010, 2012) shows that initial problem-solving activities can be effective even though invented solutions to problems are often suboptimal or even false (see Schmidt et al., 1989, for similar findings). In addition, larger numbers of suboptimal solutions were followed by higher learning outcomes (Kapur, 2012). Difficulties as well as the production of suboptimal solutions can be seen as productive because they cause impasses making the learners realize that certain solutions do not work for all cases. Furthermore, research on impasse-driven learning has shown that instructional explanations are more effective when given in the context of such an impasse (Sánchez, García-Rodicio, & Acuña, 2009; VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003). If prior knowledge is not sufficient to solve the inventing task and an impasse is reached, a perceived "vacuum" can help to see more clearly the "information needs" and "knowledge gaps to be filled", which can lead to a better focus on the most relevant contents in a subsequent learning phase.

Besides the more cognitive effects of creating a form of prior knowledge and an experience of knowledge gaps, problem-oriented instruction can influence motivation. Enhancing motivation can foster deep processing, understanding, and transfer (Belenky & Nokes-Malach, 2012; Entwistle & Ramsden, 1983; Pintrich, 2000; Pugh & Bergin, 2006). Schmidt et al. (1989) discussed an epistemic curiosity (i.e., motivation to strive for knowledge) which could explain higher learning outcomes in the problem-based condition of their experiment. Interest can be enhanced because "people like to produce things" (Schwartz & Martin, 2004, p. 171; diSessa, Hammer, Sherin, & Kolpakowski, 1991; Norman & Schmidt, 1992). Enhancing learner motivation is argued to be a major advantage of problem-oriented learning in general, but there is little research bearing directly on motivational issues (Hmelo-Silver, 2004). Most studies do not assess learners' perceived knowledge gaps, either.

Some researchers criticize the postulated effects of such forms of problem-oriented learning (Mayer, 2004). Sweller et al. (2007) as well as Kirschner, Sweller, and Clark (2006) assume that the problem-oriented activities and especially failure within these activities are unproductive. "Not only is unguided instruction normally less effective; there is also evidence that it may have negative results when learners acquire misconceptions or incomplete or disorganized knowledge" (Kirschner et al., 2006, p. 84). They criticize that many studies favoring problem-oriented learning did not employ an adequate control group. Such a control group would have to engage in the same topic as the experimental group for the same timespan (see Sweller et al., 2007).

Against the background of these different positions on the value of inventing, we designed an experiment with a control group that is adequate and "strong" in the sense of implementing "good" direct instructional procedures, but using the exact same problem as in an inventing group. However, the problem was presented in a worked-out version. Boths group engaged in their problem for the same amount of time. Specifically, we tested the following hypotheses and asked the following research questions:

(1) The inventing activity leads to more experience of knowledge-gaps, more epistemic curiosity, and more interest than working through the worked-out solution. (2) The more knowledge gaps participants perceive, the higher their focus on the most relevant contents in a subsequent learning phase. (3) Does the inventing activity lead to superior learning outcomes when compared to a "strong" control group? (4) Is failure in the inventing activity productive (Kapur, 2010, 2012), that is, (4a) how is the appropriateness of solutions to the invention problem related to learning outcomes? and (4b) how is the number of (different) suboptimal invented solutions related to learning outcomes?

Method

Participants and Design

As participants, forty-two German student teachers (sex: 12 female, 30 male; $M_{\text{age}} = 22.74$, $SD = 3.44$) were randomly assigned to two conditions: "inventing" ($n = 21$; 13 female, $M_{\text{age}} = 22.05$; $SD = 2.92$) and "worked solution" ($n = 21$; 17 female; $M_{\text{age}} = 23.43$, $SD = 3.83$). As learning domain we used the assessment of learning strategies in learning journals written by high school students. By writing learning journals, for example, as homework after biology or mathematics classes, high school students are encouraged to apply learning strategies. For example, they can develop their own thoughts based on the new learning contents (elaboration strategy). Ideally, they can do this in a detailed way and on their own, for example, "I realized that nothing can grow without mitosis, not even myself! Because (…)." Such an elaboration can be evaluated as high in quality (see Glogger, Schwonke, Holzäpfel, Nückles, & Renkl, 2012).

The inventing group invented criteria to evaluate the quality of learning strategies applied in learning journals. First, *all* participants received a short introduction (134 words) about learning journals and the quality of learning strategies in general. The instruction to the subsequent activity (for both groups) read as follows: "On pages B and C, you will find four extracts from learning journals. Each extract shows a variation of the same elaboration strategy (developing own thoughts) in a way a student in a biology class (dealing with the topic mitosis) could have realized it." The extracts looked similar in length, but differed systematically in two quality criteria "detailed elaboration vs. wordy, but shallow elaboration" and "self-made vs. copied from the lesson". Participants in the *inventing group* were prompted to contrast the four extracts, rate the quality of each one on a 3-point scale (low, medium, or high), make notes on discerning aspects, and generalize from these aspects to generic evaluation criteria for the learning strategy elaboration. The student teachers had to write down their criteria in a box labeled "my criteria." They were also instructed to check whether or not the final criteria really work to discern all extracts (cf. Roll, Holmes, Day, & Bonn, 2012). In contrast, participants in the *solution condition* neither had to rate the extracts nor to invent the evaluation criteria. Instead, they were asked to carefully study the same problem that was worked out by a (fictitious) experienced teacher. That is, the canonical criteria were written in the "my criteria" box, the quality of the four extracts was rated, and short notes about discerning aspects were written down. In summary, the two groups had the exact same work sheets with four extracts. However, the inventing group had to generate a solution to the problem how to evaluate the quality of learning strategies whereas the worked-solution group was given the solution, namely the criteria. That is, the inventing group had to generate core learning principles by contrasting cases, whereas the solution group worked through the contrasted cases with the given principles. The two criteria (principles) were explained in the subsequent learning phase (inter alia), that is, the information on the criteria was redundant for the solution group. Both groups were given the same amount of time (15 minutes) for their preparation activity. Participants were compensated with 15 Euros for the average 85 minutes duration of the study.

Materials

Pretest and Demographic Questionnaire. A web-based pretest assessed participants' topic-specific prior knowledge. Participants received up to five points for the four open questions (α = .83, e.g., "Which learning strategies can students apply by writing a learning journal?"). Two independent raters scored 25% of the pretests (*ICC* = .87) and of all following data with open format including posttest. A demographic questionnaire assessed sex, age, number of semesters in teacher education, experience with learning journals, and computer skills.

Process Variables. Questionnaires assessed the participants' experience of knowledge gaps, epistemic curiosity, and interest by items with 6-point rating scales (6: *absolutely true*). Experience of knowledge gaps was assessed with nine items (α = .89; e.g., "My knowledge was insufficient to complete the task"). Epistemic curiosity was assessed with 10 items (α = .85), based on the "Melbourne Curiosity Inventory - State Form" (Naylor, 1981) and adapted to the present context (e.g., "I feel curious about how to evaluate learning strategies"). Topic-specific interest (Schiefele & Krapp, 1996) was measured with six items (α = .72; e.g., "Learning how to evaluate learning strategies is entertaining."). In addition, we rated the appropriateness (quality) of the invented solutions (i.e., in the inventing group only) on a 6-point scale ranging from 1 (*not at all appropriate*) to 6 (*absolutely appropriate*, *ICC* = .82); and we counted solutions, operationalized by the number of (different) criteria invented by participants.

Computer-Based Learning Environment and Learning Time. After the experimentally varied preparation activity and the questionnaires, participants worked individually in a computer-based learning environment (CBLE). The CBLE explained several sub-categories of elaboration strategies, how they improve comprehension, and how they can be identified in learning journals. A subsequent unit explained the quality criteria of elaboration strategies, using various (new) examples (i.e., not used in the preparation phase). Learners could navigate freely. The focus on the most relevant learning contents was operationalized as the time learners spent in the quality criteria unit. The duration spent in this unit and in the environment as a whole was logged by the software.

Posttest. A posttest consisting of seven tasks measured learning outcomes as application of quality criteria on students' learning strategies (α = .75). Each task consisted of a short extract from a learning journal, representing one sub-category of an elaboration strategy. This sub-strategy was labeled. All extracts were new so that the tasks required transfer (content transfer, Barnett & Ceci, 2002). Participants were asked to rate the quality of the strategy (low, medium, or high) and to explain their rating by applying the previously learned criteria. Answers were rated on a 6-point scale ranging from 1 (*no conceptual understanding*) to 6 (*very clear conceptual understanding*; SOLO taxonomy by Biggs & Collis, 1982; *ICC* = .93).

Procedure

We required participants to work on the web-based pretest four days before the experiment in order to avoid knowledge activation effects. On the day of the experiment, participants first filled out the demographic questionnaire. Next, they worked individually on the task that prepared the following learning phase (inventing vs. worked-out solution) for 15 minutes. Subsequently, questionnaires assessed participants' experience of knowledge gaps, epistemic curiosity, and interest. The participants then worked on the CBLE without time limits (20 minutes on average) as we were interested in potential effects of the two conditions on the learning time spent in the environment. After the learning phase, interest was reassessed. Finally, participants worked on the posttest.

Results

A significance level of .05 was used for all analyses. We used *d* as an effect-size measure with values between .20 and .50 classified as small, values between .50 and .80 as medium, and values > .80 as large (Cohen, 1988). We did not find any significant differences between the groups in prior knowledge (inventing: *M* = 2.05 [41 % correct], *SD* =

Table 1: Means (standard deviations in parentheses) of process variables, and the posttest in the experimental groups, and test statistics.

	Knowledge-gap^a	Epistemic curiosity	Interest^b	Learning time^c	Posttest
Inventing	3.80 (0.78)	4.54 (0.63)	5.05 (0.58)	8.02 (2.06)	3.20 (0.80)
Worked solution	2.75 (1.09)	4.15 (0.78)	4.64 (0.52)	10.52 (3.00)	3.80 (0.84)
$t(40)$	3.61	1.80	2.37	-3.10	-2.39
p	<.001^d	.039^d	.012^d	.004^e	.022^e
d	1.14	0.57	0.75	-0.98	-0.75

Note. All 6-point scales: Knowledge-gap experience, epistemic curiosity, and interest: from 1 (*not true at all*) to 6 (*absolutely true*); Posttest: from 1 (*no conceptual understanding*) to 6 (*very clear conceptual understanding*). ^{a} Knowledge-gap experience. ^{b} Interest after the learning phase. ^{c} Learning time (in minutes) in the most relevant unit of the CBLE on quality criteria. ^{d} one-tailed. ^{e} two-tailed.

1.44, solution: $M = 2.37$ [47.4 % correct], $SD = 1.29$, $t(40) = .75$, $p = .459$) or in demographic variables (sex, age, number of semesters, experience with writing learning journals, and computer skills; all p's > .05).

Table 1 presents the means and standard deviations of knowledge-gap experience, epistemic curiosity, interest, learning time, and the posttest for the two experimental groups. Directly after the preparation task, the participants of the inventing condition stated higher knowledge-gap experience (large effect), and higher epistemic curiosity (medium effect), than participants of the solution condition, confirming hypothesis 1. Even after the learning phase, they stated higher interest in learning about the assessment of learning strategies in learning journals (medium effect).

Table 2 shows correlation coefficients between all process variables and dependent variables. Regarding hypothesis 2, perceived knowledge gaps did not significantly correlate with learning time (simple correlation, Table 2). However, when controlling for condition (and prior knowledge) we found a significant partial correlation, $r(36)_{\text{part}} = .34$, $p = .035$, medium effect. Thus, the more knowledge gaps participants perceived (independent from their condition and prior knowledge), the more time they spent in the most relevant learning unit about quality criteria indeed.

Surprisingly, even though the inventing group experienced more knowledge gaps (correlating with learning time), higher epistemic curiosity, and higher interest, this group did not achieve better learning outcomes (research question 3, see Table 1). Participants of the solution condition even outperformed the inventing condition (medium effect). Controlling for prior knowledge, the effect remained stable, $F(1,38) = 4.95$, $p = .032$, $d = -0.72$. For exploratory purposes, we searched for variables explaining this effect. 'Learning time' is the only variable that correlated significantly with learning outcomes (Table 2), even if controlled for condition, $r(38)_{\text{part}} = .55$, $p < .001$, large effect. Against this background, we analyzed whether the effect of conditions on learning outcomes is mediated by learning time. We tested this mediation effect with a set of related multiple regression equations, following a products-of-coefficients strategy (MacKinnon, 2008). In this approach, there are essentially two assumptions to be met in order to speak of a mediated effect. First, an independent variable must significantly affect a mediating variable (path a). Second, the mediating variable must significantly affect a dependent variable (path b). The significance of the effect can be tested according to Sobel (1982).

The independent variable 'condition' did in fact significantly affect the mediating variable 'learning time' (path a): The worked solution group spent more time learning in the CBLE (quality criteria unit, $b = -1.09$, $SE = 0.41$, $b^* = -.39$, $p = .010$; controlled for prior knowledge). The learning time predicted learning outcomes significantly (path b: $b = .210$, $SE = 0.0404$, $b^* = .625$, $p < .001$; controlled for prior knowledge). Learning time significantly mediated the effect of conditions on learning outcomes (Sobel test = -2.37, $p = .018$ [two-tailed], $LCL = -0.418404$, $HCL = -0.04032$).

Referring to research question 4a, the appropriateness of participants' invented solutions ($M = 2.8$, $SD = 1.31$) correlated substantially with learning outcomes, $r(16)_{\text{part}} = .52$, $p = .028$, (large effect, controlled for prior knowledge. The more a participant failed to invent an appropriate solution to the inventing problem, the more this participant failed in the posttest as well.

Table 2: Intercorrelations of pretest, process variables, and posttest.

	Knowledge-gap	Epistemic curiosity	Interest pre	Interest post	Learning time^a	Posttest
Pretest	.18	.40*	.19	-.01	.30$^+$	.25
Knowledge-gap experience	—	.52***	.33*	.24	.10^b	-.10
Epistemic curiosity		—	.61***	.48**	.09	-.20
Interest pre			—	.53***	.01	-.11
Interest post				—	-.13	-.29$^+$
Learning time^a					—	.66***

Note. $N = 44$. ^{a} Learning time (in minutes) in the most relevant CBLE unit on quality criteria. ^{b} Partial correlation differs and is given in the text with condition controlled. $^+p < .10$. $^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Finally, hypothesis 4b was rejected. The number of suboptimal invented solutions ($M = 4.1$, $SD = 1.52$) did not correlate significantly with learning outcomes, $r(16)_{\text{part}} = .28$, $p = .135$ (one-tailed, controlled for prior knowledge).

Discussion

In line with the literature on problem-oriented learning (e.g., Hmelo-Silver, 2004), we found that an inventing activity had positive motivational effects (see also Belenky & Nokes-Malach, 2012). Learners are more curious about and interested in the target learning domain. Learners also become aware of knowledge gaps to be filled. The more knowledge gaps learners perceived, the higher their focus on the most relevant learning contents. The motivational and knowledge-gap effects can be seen as a preparation for learning. However, they did not lead to higher learning outcomes in the inventing group. In contrast to the inventing and productive-failure literature, the worked solution group achieved better learning outcomes, mediated by learning time.

The results are in line with Sweller et al.'s (2007) and Kirschner et al.'s (2006) argument that positive results in problem-oriented learning studies could be an effect of "weak" control conditions or different time-on-task during the experimental variation. The worked-out solution condition resembles a worked-example condition. Possibly, a worked-example effect (Renkl, 2011; Sweller, 2006) had the worked solution group outperform the inventing group, even though self-explanations were not prompted, which is usually sensible in order to exploit the potential of example-based learning (cf. Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Renkl, 2011). In the present case, the preparation activity of working through a worked-out solution of a problem obviously prepared learners to learn, possibly by enhancing basic knowledge about quality criteria. This well-organized basic knowledge might have facilitated working intensively with the instructional explanations and explained strategy examples of different quality in the CBLE. That is, cognitive mechanisms such as deeper elaboration and spontaneous application of the learned concepts to presented examples during the learning phase could account for the results. Findings of a subsequent think-aloud study indicate such mechanisms. The cognitive mechanisms could have predominated motivational mechanisms in this study. Alternatively (or additionally), if the worked-solution functioned as a worked example (a model of a good solution), self-efficacy could have been enhanced. Self-efficacy can enhance effort and persistence (Schunk, 1990) which could explain the short learning time in the inventing group despite enhanced curiosity and interest.

One could interpret a speed-accuracy tradeoff: learning outcomes as well as learning time is higher in worked example. However, in the context of self-regulated learning in a CBLE, where diverse learning paths are provided and working through it can be more or less thorough, enhancing learning time can be advantageous. To put it differently, inventing might have constrained learning time, because learners did not want to deal thoroughly with the learning contents (see also below). Also note that we did not find any differences between groups in efficacy (learning outcome per minute spent in the quality part, $p = .751$; in the CBLE, $p = .975$).

Another evidence for the claim that some "correct" basic knowledge of quality criteria facilitated future learning in the CBLE can be seen in the highly positive correlation between the level of appropriateness of the inventing solutions and the learning outcome. Failure was not productive in the present case. The number of suboptimal invented solutions was not significantly related to learning outcomes. These findings contradict Kapur's (2010, 2012) approach of productive failure. Additionally, findings about the continued influence effect (Johnson & Seifert, 1994) and the IKEA effect (Norton et al., 2012) suggest that the inventing group could have clung to their initial suboptimal solution (quality criteria) even though the canonical criteria were explained and exemplified in the CBLE. Holding on to one's own solution ideas and partly neglecting canonical explanations could be another reason why the inventing group spent less time with learning in the CBLE. If they partly held on to their own solutions, the role of a transition phase, which is a usual part of a productive-failure procedure, could be of major importance: The teacher leads a discussion about students' own (suboptimal) solutions towards the canonical one.

Usually, productive failure and inventing include a collaborative learning setting. Students work on preparatory open problems either in groups (Kapur, 2012; Schmidt et al., 1989; Schwartz & Martin, 2004) or in pairs (Schwartz et al., 2011; Westermann & Rummel, 2012). The collaborative setting has not been explicitly discussed as an "active ingredient" of the preparatory activities. However, invention activities might only be effective in collaborative learning settings. This is a point to consider about the present study in which participants worked only individually.

Thus, further studies on problem-oriented learning settings such as inventing should investigate whether collaborative work during preparatory activities is a crucial ingredient of such instructional approaches. It can also be worthwhile to look at the learning processes that follow inventing and inventing-enhanced motivational states (e.g., applying basic knowledge from the preparatory activity, holding on to own solutions) in order to explain why inventing did not enhance learning outcomes when compared to working through a worked-out solution. More generally, it is important to use strong control groups in future research. In doing so, research should not simply investigate which instructional approach is better or worse. Instead, the results of the present study show that future research should achieve a differential analysis of cognitive and motivational effects of the various approaches.

Acknowledgments

The project "Fostering diagnosis skills of teachers: Identifying and fostering learning strategies through learning journals" was part of the research program "Bildungsforschung" (educational research) of the Baden-Württemberg Stiftung. We would like to thank our research assistants Corinna Fleischer, Andrea Ohst, Duygu Pektas.

References

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

Berthold, K., & Renkl, A. (2010). How to foster active processing of explanations in instructional communication. *Educational Psychology Review*, *22*, 25-40.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO Taxonomy*. New York: Academic Press.

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*, 145–182.

diSessa, A. A., Hammer, D., Sherin, B. L., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, *10*, 117–160.

Entwistle, N. J., & Ramsden, P. (1983). *Understanding student learning*. London: Croom Helm.

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology*, *104*, 452–468.

Hmelo-Silver, C. E. (2004) Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*, 235-266.

Johnson, H. M., & Seifert, C. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1420-1436.

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, *38*, 523–550.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*, 651–672.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist*, *41*, 75-86.

MacKinnon, D. P. (2008). *Introduction into statistical mediation analysis*. New York: Erlbaum.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist*, *59*, 14–19.

Norman, G. R., & Schmidt, H. G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic Medicine*, *67*, 557-565.

Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, *22*, 453-460.

Pintrich, N. E. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation*. Orlando, FL: Academic Press.

Pugh, K. J., & Bergin, D. (2006). Motivational influences on transfer. *Educational Psychologist*, *41*, 147-160.

Renkl, A. (2011). Instruction based on examples. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction*. New York, NY: Routledge.

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluating metacognitive scaffolding in guided invention activities. *Instructional Science*, *40*, 691–710.

Sánchez, E., García-Rodicio, H., & Acuña, S. R. (2009). Are instructional explanations more effective in the context of an impasse? *Instructional Science*, *37*, 537–563.

Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and Individual Differences*, *8*, 141-160.

Schmidt, H. G., De Volder, M. L., De Grave, W. S., Moust, J. H. C., & Patel, V. L. (1989). Exploratory models in the processing of science texts: The role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology*, *81*, 610-619

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, 25, 71–86.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*. 759-775.

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

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290-312.

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction*, *16*, 165–169.

Sweller, J., Kirschner, P., & Clark, R. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist*, *42*, 115-121.

Van Lehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, *21*, 209-249.

Westermann, K., & Rummel, N. (2012). Delaying instruction: Evidence from a study in a university relearning setting. *Instructional Science*, *40*, 673–689.