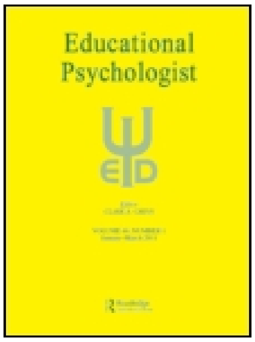

Routledge
Taylor & Francis Group

Educational Psychologist

ISSN: 0046-1520 (Print) 1532-6985 (Online) Journal homepage: http://www.tandfonline.com/loi/hedp20

Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning

Manu Kapur

EDUCATIONAL PSYCHOLOGIST, *51*(2), 289–299, 2016

ISSN: 0046-1520 print / 1532-6985 online
DOI: 10.1080/00461520.2016.1155457

Routledge
Taylor & Francis Group

Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning

Manu Kapur

Department of Psychological Studies
The Hong Kong Institute of Education

Learning and performance are not always commensurable. Conditions that maximize performance in the initial learning may not maximize learning in the longer term. I exploit this incommensurability to theoretically and empirically interrogate four possibilities for design: productive success, productive failure, unproductive success, and unproductive failure. Instead of only looking at extreme comparisons between discovery learning and direct instruction, an analysis of the four design possibilities suggests a vast design space in between the two extremes that may be more productive for learning than the extremes. I show that even though direct instruction can be conceived as a productive success compared to discovery learning, theoretical and empirical analyses suggests that it may well be an unproductive success compared with examples of productive failure and productive success. Implications for theory and the design of instruction are discussed.

Incommensurability between performance and learning lies at the core of the argument I advance in this article. R. A. Schmidt and Bjork's (1992) seminal and highly influential review of psychological research on verbal and motor learning suggested that experimental manipulations that may hinder performance in the shorter term can actually be productive for learning in the longer term. They advanced the notion of introducing "desirable difficulties" during the initial learning to afford learners opportunities to engage in processes that are germane for learning even if they result in a performance dip. Examples of desirable difficulties include increasing the complexity of/variability in the task, unguided problem solving, or reducing or delaying feedback during the initial learning. They concluded that conditions that maximize performance in the initial learning may not necessarily be the ones that maximize learning in the longer term. Conversely, conditions that adversely affect performance initially may result in better learning in the longer term.

Exploiting the incommensurability between learning and performance results in four possibilities for design: productive success, productive failure, unproductive success, and unproductive failure.

First among these four design possibilities is designing conditions that maximize performance in the shorter term and maximize learning in the longer term. I refer to such design efforts as *productive success*. Productive success involves structuring problem-solving and learning activities with the goal of achieving both improved performance on problem solving and sustainable learning. For example, constructivist approaches that fall into the genres of *problem-based learning* (PBL) and *guided inquiry* involve scaffolded problem-solving activities initially to engender learning, with a gradual fading of the scaffolds as learners gain expertise (Puntambekar & Hübscher, 2005; H. G. Schmidt, Loyens, Van Gog, & Paas, 2007).

Second is the possibility of designing conditions that may not maximize performance in the shorter term but in fact maximize learning in the longer term. I refer to such design efforts as *productive failure* (Kapur, 2008). Productive failure engages students in solving problems requiring concepts they have yet to learn, followed by consolidation and instruction on the targeted concept. By failure, I simply mean that students will typically not be able to generate or discover the correct solution(s) by themselves. However, to the extent that students are able to use their prior knowledge

Correspondence should be addressed to Manu Kapur, Department of Psychological Studies, The Hong Kong Institute of Education, 10 Lo Ping Road, Tai Po, New Territories, Hong Kong. E-mail: mkapur@ied.edu.hk

to generate suboptimal or even incorrect solutions to the problem, the process can be productive in preparing them to learn better from the subsequent instruction that follows (Kapur & Bielaczyc, 2012; Schwartz & Martin, 2004).

Third is the possibility of designing conditions that may maximize performance in the shorter term without maximizing learning in the longer term. In other words, there is an illusion of learning in initial high performance. I refer to such design efforts as *unproductive success*. For example, teaching methods that rely largely on drill-and-practice or rote memorization would fall into this category, for it is possible for students to show high performance on memory tasks or carrying out problem-solving procedures without a commensurable understanding of what it is that they are doing. A classic example comes from Miller and Gildea's (1987) work on vocabulary learning. They described how children who learned the meaning and use of words mainly from dictionary definitions are often not able transfer it appropriately to practice. For example, even though they may be able to state the meaning of the word *correlate*, how they use the word in practice (e.g., *Me and my parents correlate, because without them I wouldn't be here*) may be completely meaningless.

Finally, there is the possibility of designing conditions that maximize neither performance nor learning in the short or long terms. I refer to such design efforts as *unproductive failure*. A well-studied example of unproductive failure is pure discovery learning, where students are expected to learn (or discover) the targeted concepts by engaging in solving problems without any guidance or support whatsoever. Throughout this article, I use "discovery learning" and "unguided problem solving" interchangeably.

The purpose of this article is to interrogate the four design possibilities, compare and contrast them with each other, and derive implications for the design of initial learning. I start with the obvious low-hanging fruit in unproductive failure. I use research on the ineffectiveness of pure discovery learning to illustrate unproductive failure and the case against it. I then turn my attention to productive failure and productive success, and the comparison between the two. I conclude by situating my findings in the broader research literature, as well as deriving implications for the long-standing instructivist-constructivist debate on the design of learning (Kirschner, Sweller, & Clark, 2006; Tobias & Duffy, 2009).

UNPRODUCTIVE FAILURE

There is consensus among instructivists (e.g., Kirschner, Sweller, & Clark, 2006) and constructivists (e.g., Hmelo-Silver, Duncan, & Chinn, 2007) that unguided problem solving or discovery learning often leads to dismal learning outcomes. For example, instructivists bring to bear substantive empirical evidence against unguided or minimally guided instruction to claim that there is little efficacy in having learners solve problems that target novel concepts and that learners should receive direct instruction on the concepts before any problem solving (for fuller reviews, see Kirschner, Sweller, & Clark, 2006; Mayer, 2004).

Perhaps this view is best captured by Sweller (2009): "What can conceivably be gained by leaving the learner to search for a solution when the search is usually very time consuming, may result in a suboptimal solution, or even no solution at all?" (p. 128). This view is grounded in cognitive load theory (CLT; Sweller, 1988) and is supported by a large body of evidence that has compared some form of heavily guided direct instruction (e.g., through well-designed worked examples) favorably with unguided or minimally guided discovery learning. This led Kirschner, Sweller, and Clark (2006) to argue that "controlled experiments almost uniformly indicate that when dealing with novel information, learners should be explicitly shown what to do and how to do it" (p. 79). Direct instruction typically involves the use of instruction on the targeted concepts followed by or coupled with the use of well-designed worked examples to illustrate and explain the targeted concepts, before independent problem solving (Kirschner, Sweller, & Clark, 2006).

It is not surprising that learners do not learn as much from unguided problem solving when compared with a heavily guided direct instruction through worked examples. From the perspective of CLT, when students do not have the knowledge to solve a problem, they often search the problem space for solutions by engaging in resource intensive processes such as trial and error or means-ends analysis, which burden the limited working memory capacity. Because all conscious processing happens in the working memory, working memory is less available for learning new concepts and procedures if it is mainly occupied with such a search of the problem space (Kirschner, Sweller, & Clark, 2006; Tuovinen & Sweller, 1999). By showing the learner exactly what to do and how to do it, direct instruction reduces this load on the working memory, thereby facilitating the development of correct domain knowledge and procedures (Klahr & Nigam, 2004; Sweller & Chandler, 1991). Compared with unguided problem solving, where students may not even be able to discover correct knowledge and procedures on their own at all (Klahr & Nigam, 2004), direct instruction affords students the opportunities to attend to and acquire the correct procedures and knowledge while reducing the probability of encoding of errors and misconceptions (Sweller & Chandler, 1991).

Theoretically and empirically, there is no doubt about the superiority of direct instruction over unguided problem solving. Based on the comparison against unguided problem solving, one could even judge direct instruction as a productive success. However, I argue to hold off on the judgment until we have more closely interrogated the following fallacious inferences that usually get derived from the comparison:

1. No conceivable benefit can be derived from unguided problem solving.
2. Because unguided problem solving is the problem, maximal guidance is the solution.
3. CLT is against the use of unguided problem solving when learning something new.

I discuss each in turn.

No Conceivable Benefit Can Be Derived From Unguided Problem Solving

That direct instruction trumps unguided problem solving does not logically imply there is little efficacy in having learners solve problems that target concepts they have not learned yet. To determine if there such an efficacy, a stricter comparison is needed in which one compares direct instruction with an approach where students first engage in unguided problem solving on their own, followed by appropriate instruction.

I use Klahr and Nigam's (2004) study as a case in point because it is often cited as a stellar example of the effectiveness of direct instruction over discovery learning. Klahr and Nigam first conducted a baseline assessment to see if students knew the control of variables strategy (CVS) principle by getting them to design experiments on their own. The success rate on the baseline assessment was expectedly very low. Students who were subsequently assigned to the discovery learning condition simply continued to design these experiments without any instruction on CVS or any feedback. However, for students in the direct instruction condition, the instructor modeled and contrasted the design of both confounded and unconfounded experiments with appropriate instructional facilitation and explanation to make them attend to critical features of why CVS helps isolate the effects of a factor whereas confounded experiments do not.

It is unsurprising that direct instruction was found to be more effective than discovery learning on posttest measures. However, one can also argue that the baseline assessment in Klahr and Nigam's (2004) study seems to function very much like a discovery learning phase where students generated their own solutions (in this case, experiments) to solve a problem that targets a concept (in this case, CVS) they had not learned yet. If so, the very effects that Klahr and Nigam attribute to direct instruction *alone* seem more appropriately attributed to a pure discovery learning phase (their baseline assessment) followed by direct instruction. Therefore, much as Klahr and Nigam set out to show that there is little efficacy in students exploring and solving problems requiring concepts they have not learned yet, their findings can be reinterpreted to support precisely the opposing contention that such discovery learning can in fact be efficacious provided some form of instruction that build upon it follows.

Because Unguided Problem Solving Is the Problem, Maximal Guidance Is the Solution

The superiority of direct instruction over unguided problem solving does not mean that a rejection of one extreme (no guidance at all) implies an automatic adoption of the other (heavy or maximal guidance) as the most effective solution. Perhaps there exist other ways of designing guidance in the initial learning that are just as or even more effective than the maximally guided direct instruction.

In a strong critique of discovery learning, Mayer (2004) reviewed studies that compared unguided discovery with guided discovery and direct instruction. Although the review clearly showed that unguided discovery was the least effective, it also showed that direct instruction was not the most effective, and that it was in fact guided discovery that outperformed both unguided discovery and direct instruction. For example, Kittel (1957) compared learning of logic problems (e.g., finding the rule for excluding a word from a group of words) through pure discovery, guided discovery (where a hint was given), and direct instruction (where the rule was told first), and Kittel found guided discovery to be the best for retention and transfer. Gagne and Brown's (1961) study echoed similar findings in the learning of math where students had to learn how to derive formulas to sum a number series. There are other methods, too, such as PBL, the effectiveness of which over direct instruction has been demonstrated, though these effects are not always consistent (e.g., see Hmelo-Silver, Duncan, & Chinn, 2007). Research on productive failure (Kapur, 2008, 2010, 2014) and preparation for future learning (Schwartz & Bransford, 1998; Schwartz & Martin, 2004) also represent teaching methods that are more effective than direct instruction.

The point is simple: Just because a complete lack of guidance (in unguided problem solving) is the problem does not logically mean that maximal provision of guidance is the most effective solution; one does not have to swing from one extreme to the other. It is possible to design teaching methods that are better than both completely unguided problem solving and heavily guided direct instruction. If so, direct instruction may come across as an example of a productive success when compared with discovery learning, but that may not hold true when compared with other designs. I take this up later in the article.

CLT Is Against the Use of Unguided Problem Solving When Learning Something New

According to CLT, the processing of novel information depends upon an interaction between a limited working memory capacity and relevant information stored in the long-term memory. Kirschner, Sweller, and Clark (2006) argued that "any instructional theory that ignores the limits of working memory when dealing with novel information or ignores the disappearance of those limits when dealing with familiar information is unlikely to be effective" (p. 77).

If what a learner already knows—prior knowledge—about a concept is a critical determinant of either limiting or expanding the working memory capacity as conceptualized by CLT, then a commitment to CLT entails a commitment to understanding whether and the extent to which the targeted concept is novel to the learner. If one assumes that learners do not have any prior knowledge of the targeted concept, one is constrained to work within the limiting aspects of the working memory, which is what the proponents of direct instruction largely seem to have done (e.g., Carroll, 1994; Paas, 1992; Sweller & Chandler, 1991).

Note that there are CLT studies that take into account learners' prior knowledge. For example, Kalyuga, Chandler, Sweller, and Tuovinen (2001) found that the learning with worked examples benefited learners initially when they had low or no domain knowledge. As their domain knowledge increased, the effectiveness of worked examples over unguided problem solving disappeared. However, these studies maintain the recommendation that one should use direct instruction to build initial expertise, and only with the build-up domain expertise that the use of worked examples be diminished or removed in favor of unguided problem solving.

In contrast, the possibility of using unguided problem solving for learners with initially no formal knowledge of the concept has not been explored within CLT. For such learners, could we not design problem-solving tasks that can activate and elicit their prior knowledge, albeit suboptimal or even incorrect, about a concept even if they have not formally learned it yet? To the extent that we can accomplish this, it follows that by activating and working with these priors in the long-term memory, one can leverage the expandable aspects of the working memory capacity. At the very least, this is a theoretical possibility that CLT allows for, yet it remains underresearched. It seems that the proponents of CLT, by assuming learners have no prior knowledge, have theoretically underdetermined the design implications of CLT.

Against the backdrop of the three fallacious inferences, I now turn to demonstrating a case for productive failure. I show that it is possible to design problem-solving activities for eliciting students' prior knowledge through unguided problem-solving initially, and that there can be an efficacy of such unguided problem solving provided an appropriate guidance in the form of consolidation and instruction subsequently follows.

PRODUCTIVE FAILURE

Productive failure involves two phases: a problem-solving phase followed by a consolidation (or instruction) phase (for a fuller description of the design, see Kapur & Bielaczyc, 2012). The problem-solving phase affords opportunities for students to generate and explore the affordances and constraints of multiple solutions to novel, complex problems. The consolidation phase affords opportunities for comparing and contrasting, organizing, and assembling the relevant student-generated solutions into canonical solutions.

There is now a growing body of evidence that generating solutions to novel problems *prior* to instruction can help students learn better from the instruction (Kapur & Rummel, 2012). Evidence comes not only from quasi-experimental studies conducted in the real ecologies of classrooms (e.g., Kapur, 2012, 2013; Kapur & Bielaczyc, 2011; Schwartz & Bransford, 1998; Schwartz & Martin, 2004) but also from controlled experimental studies (e.g., DeCaro & Rittle-Johnson, 2012; Kapur, 2014; Loibl & Rummel, 2013, 2014; Roll, Aleven, & Koedinger, 2011; R. A. Schmidt & Bjork, 1992; Schwartz, Chase, Oppezzo, & Chin, 2011).

For example, in a study with eight-grade students, Schwartz and colleagues (2011) compared students who invented solutions with contrasting cases before receiving instruction on the concept of density with those who were instructed first and then practiced with the same cases. They found that although there was no effect on procedural knowledge, invention activities prepared students to learn the deep structure of density better, which resulted in better transfer than those who received instruction first. Likewise, DeCaro and Rittle-Johnson (2012) had second- to fourth-grade students solve unfamiliar math problems on number sentences before or after receiving instruction on number sentences. Once again, there was no difference on procedural knowledge, but students who solved problems first developed better conceptual understanding than those who first received instruction. More recently, in a randomized-controlled experiment with ninth-graders learning the concept of standard deviation, Kapur (2014) had students individually generate solutions to a novel problem before or after receiving instruction. He too found no difference on procedural knowledge, but students who engaged in problem solving prior to instruction demonstrated significantly better performance on conceptual understanding and transfer than those who engaged in problem solving after instruction.

There are several interdependent mechanisms underpinning the preparatory effects of problem solving prior to instruction. First, starting with problem solving may be better at activating and differentiating relevant prior knowledge provided students are able to use their priors to generate suboptimal or even incorrect solutions to the problem (DeCaro & Rittle-Johnson, 2012; Schwartz, Chase, Oppezzo, & Chin, 2011; Siegler, 1994). Because students can rely only on their prior knowledge to generate solutions, the nature of these solutions provides a measure of the types of knowledge that was activated and how this knowledge is relevant in relation to the targeted concept (Kapur, 2014; Loibl & Rummel, 2014; Westermann & Rummel, 2012). Second, prior knowledge activation may in turn afford more opportunities for students to (a) notice the inconsistencies in and realize the limits of their prior knowledge (DeCaro & Rittle-Johnson, 2012; Loibl & Rummel, 2014; Ohlsson, 1996) and (b) compare and contrast student-generated solutions and correct solutions during subsequent instruction, thereby helping students' to attend to and better encode critical features of the new concept (Kapur, 2014; Schwartz, Chase, Oppezzo, & Chin, 2011). Finally, besides the cognitive benefits, problem solving prior to instruction may also have affective benefits of greater learner agency, as well as engagement and motivation to learn the targeted concept (Belenky & Nokes-Malach, 2012; Clifford, 1984; Hiebert & Grouws, 2007).

These studies provide robust evidence that there is an efficacy of unguided problem solving, but only if some form of consolidation and instruction follows. Note also that the greater efficacy of productive failure over direct instruction lies not in the development of procedural knowledge but in conceptual knowledge and transfer.

A synthesis across several productive failure studies suggests key design features for its benefits to be realized: (a) The initial problem-solving task should be challenging enough to engage the learner in the exploration, but not so challenging that the learner gives up; (b) it must admit multiple solutions, strategies, and representations, that is, afford sufficient problem and solution spaces for exploration; (c) the problem should activate learner's prior knowledge—formal as well as intuitive—to solve the problem; and (d) a teacher or an expert should build upon the student-generated solutions by comparing and contrasting them with the correct solution, thereby directing attention to and aiding encoding of the critical features of the targeted concept. When designing for productive failure in the real ecologies of classroom, it further helps if appropriate disciplinary norms and expectations for problem solving and learning are set and reinforced (for a fuller explication of design principles of productive failure, see Kapur & Bielaczyc, 2012).

Next I turn to more evidence, this time from productive success that can result in better learning than direct instruction. Unlike productive failure, where students do not necessarily generate or learn the correct solutions in the initial problem-solving phase, productive success guides the initial problem-solving process in ways that leads to both successful problem solving and learning.

PRODUCTIVE SUCCESS

The difference between productive failure and productive success is a subtle but an important one. The goal for productive failure is a preparation for learning from subsequent instruction. Thus, it does not matter if students do not achieve successful problem-solving performance initially. In contrast, the goal for productive success is to learn through a successful problem-solving activity itself. Because students do not know the concepts to solve these problems, the problem-solving process, unlike in productive failure, is heavily guided to achieve both problem-solving success and learning (Hmelo-Silver, Duncan, & Chinn, 2007).

An established and well-researched example of productive success is PBL. PBL, as the name suggests, situates student learning in collaborative and authentic problem-solving tasks and activities. Through PBL, students learn the targeted domain knowledge, as well as collaboration, communication, and reflection skills (Hmelo-Silver, Duncan, & Chinn, 2007). PBL is supported by a range of scaffolds distributed across the materials; technological tools; and, most important, teachers and domain experts (Puntambekar & Kolodner, 2005). Teachers model behavior, guide student problem solving by asking questions, and directing attention to critical features to help the learner achieve what he or she would not be able to without the scaffolds and guidance. The intended result is both successful problem solving and learning.

PBL has been subject to numerous comparisons with traditional direct instruction. Meta-analytic studies, however, provide mixed evidence (e.g., Albanese & Mitchell, 1993; Vernon & Blake, 1993). For example, Vernon and Blake (1993) found that whereas traditional direct instruction was better for the acquisition of basic medical knowledge, PBL instruction was better in transferring that knowledge to solve problems in clinical practice. Dochy, Segers, Van den Bossche, and Gijbels (2003) meta-analysis found not only that PBL students performed better on knowledge application but also that the advantage of direct instruction over PBL on basic knowledge acquisition vanished after the 2nd year of medical school. Gijbels and colleagues (2005) echoed similar findings: PBL helped develop better conceptual understanding of the underlying principles of the domain and how these principles link the concepts. Even more interesting, their analysis suggested that "the better the capacity of an instrument for evaluating the application of knowledge by the student, the larger the ascertained effect of PBL" (p. 45). Finally, in the most recent meta-analyses of 270 comparisons, H. G. Schmidt, van der Molen, te Winkel, and Wijnen (2009) found small but positive effects on medical knowledge and diagnostic reasoning, and much stronger effects on medical skills in favor of PBL.

It is important to note that almost all the studies included in the aforementioned meta-analyses involved medical students, which limits the scope of generalization to other domains and age groups. Therefore, it is worthwhile examining the effectiveness of PBL beyond medical education, both in the schools and higher education contexts. For example, in a longitudinal quasi-experimental study with preservice teachers, Derry, Hmelo-Silver, Nagarajan, Chernobilsky, and Beitzel (2006) demonstrated consistently better transfer effects in favor of PBL over traditional instruction. In a well-controlled experimental study, Capon and Kuhn (2004) found that adult MBA students who engaged in PBL followed by a lecture on the targeted concepts demonstrated greater conceptual understanding of the targeted concepts than those who received a lecture first followed by PBL. There were no significant effects on declarative knowledge. Similar results were obtained by Kuhn and Dean (2005).

In school contexts, perhaps the most convincing piece of evidence for PBL comes from a large-scale study on anchored instruction—a type of PBL—by the Cognition and Technology Group at Vanderbilt (1992). Conducted in middle-school mathematics classrooms across 11 U.S. states, PBL students showed significantly better performance on standardized tests as well as transfer problems than their matched comparison counterparts. Further evidence from a multischool study of PBL in high school contexts comes from a study by Mergendoller, Maxwell, and Bellisimo (2006), who found better knowledge gains for PBL students than their counterparts in traditional direct instruction.

Overall, the pattern of effects seems consistent. Compared to traditional direct instruction, PBL may have a null or an even negative effect on the acquisition of basic knowledge but a positive effect on conceptual understanding and transfer. This body of evidence suggests that is possible to design instruction centered on problem-solving activities, and then supporting and guiding these activities in ways that lead to successful problem-solving and learning—in other words, productive success.

Taken together, the preceding analysis of productive failure and productive success counters the fallacious inferences outlined earlier. At the same time, it also raises two important questions:

1. Which is more effective for learning: productive failure or productive success?
2. Could direct instruction be a form of unproductive success?

Which is more effective for learning: productive failure or productive success?

Even though both productive failure and productive success can lead to productive learning outcomes, it is not clear if one is better than the other. How does productive failure fare against productive success? Answering this question would help us understand if and the extent to which problem-solving failure (or success) is a necessary condition for learning.

Ideally, a clean experimental comparison would entail comparing problem solving that initially leads to failure followed by instruction with problem solving that leads to success followed by the same instruction. However, to my knowledge, no such experimental comparison exists. That said, a couple of studies that do come close to affording such a comparison are studies by Schwartz, Chase, Oppezzo, and Chin (2011) and DeCaro and Rittle Johnson (2012) described earlier. In both the studies, there was a nontrivial proportion of students who achieved problem-solving success initially. However, neither study provides a breakdown of results comparing students who achieved problem-solving success with those who did not. Thus, the question remains an open one.

Another way to answer the question would be by scaffolding productive failure students in their initial problem solving to guide them toward a correct solution, much like one would do in productive success. Then on could compare them with the regular productive failure students who are unguided in their initial problem-solving activity.

Kapur (2011) addressed precisely this question in a quasi-experimental study where students were assigned to either productive failure or direct instruction conditions, or a third condition called the guided problem-solving condition. Whereas students in the productive failure condition did not receive any form of guidance or support during the problem-solving phase, students in the guided condition were provided with cognitive support and facilitation throughout that process. Such guidance was typically in the form of teacher clarifications, focusing attention on significant issues or parameters in the problem, question prompts that engendered student elaboration and explanations, and hints toward productive solution steps (Puntambekar & Hübscher, 2005). Findings suggested that the guidance did in fact result in problem-solving success, whereas none of the productive failure students were able to generate the correct solution. However, on the posttest after instruction, students from the productive failure condition outperformed those from both the direct instruction and guided-generation conditions on procedural knowledge, conceptual understanding, and transfer. The differences between guided-generation and direct instruction conditions were not significant.

Loibl and Rummel (2013) independently replicated this effect in a study with three similar conditions: unguided problem solving prior to instruction, guided problem solving prior to instruction in which students were supported with cognitive prompts during the problem-solving phase, and direct instruction. They found that, in spite of guidance helping students generate better quality solutions, there was no significant difference between the guided and unguided problem-solving conditions on procedural or conceptual understanding. That is, cognitive guidance during the initial problem-solving phase did not result in better learning on the posttest.

Both studies (Kapur, 2011; Loibl & Rummel, 2013) suggest that although guidance in the initial problem solving may well lead to problem-solving success or better quality of solutions, this better performance does not translate to better learning from the subsequent instruction. A plausible reason could be that both low- and high-quality solutions present opportunities to learn during the subsequent instruction, especially through a comparison and contrast between the student-generated solutions and the canonical solution. Another reason could be that problem-solving success does not guarantee that students understand how and why the solution works. Thus, guiding initial problem solving may not add to the preparatory benefits of problem solving.

This of course implies not that all guidance is unnecessary but that more research is needed to understand the types of guidance is necessary to support the initial problem solving. Therefore, the question of choosing between productive failure and productive success remains an open one, and one that future work do well to look into, for it would help unpack the boundary conditions of learning from failure (or success) in the initial problem solving.

COULD DIRECT INSTRUCTION BE A FORM OF UNPRODUCTIVE SUCCESS?

One one hand, given direct instruction's consistently superior performance against discovery learning, it would be rather harsh to categorize it as an unproductive success. On the other hand, as the preceding analyses showed, direct instruction was found wanting against productive failure and productive success. Also noted was the pattern of underperformance of direct instruction, which was not so much in terms of the acquisition of basic knowledge but more on conceptual understanding and transfer.

Add to this evidence the longer term adverse effects of direct instruction. Recall that a striking finding from Dochy, Segers, Van den Bossche, and Gijbels's (2003) meta-analysis on PBL was that the advantage of traditional direct instruction on medical students' basic knowledge acquisition disappeared after the 2nd year of college. In other words, even the positive effect of direct instruction on basic knowledge acquisition tended to vanish in the longer term. This finding was not limited to just Dochy et al.'s meta-analysis. In a synthesis of eight major meta-analyses of PBL since 1992, Strobel and van Barneveld (2009) concluded that even though PBL could result in slight underperformance on standardized tests of basic knowledge in the shorter term, it was more effective when it came to long-term retention and performance improvement.

There is further strong experimental evidence suggesting that direct instruction does not fare well over the longer term. For example, Dean and Kuhn (2006) replicated Klahr and Nigam's (2004) study on the learning of CVS, but with one exception. Instead of a short-term study involving one session of instruction and problem solving, Dean and Kuhn carried out a 10-week study to compare the relative effects of direct instruction alone, direct instruction followed by problem solving, and problem solving alone. They found that direct instruction was "neither a necessary nor sufficient condition for robust acquisition or for maintenance over time" (p. 384)—a conclusion that is consistent with the other longer term PBL studies, not only with medical students (e.g., Hmelo, 1998) but also with school students (e.g., Cognition and Technology Group at Vanderbilt, 1992) and nonmedical, higher education students (e.g., Derry et al., 2006).

Finally, consider emerging evidence that direct instruction may actually constrain search for novel solutions, which is a necessary component of inventiveness and creativity. Kapur (2014) found that, whereas productive students were able to design five to six solutions before receiving instruction on the targeted concepts, their direct instruction counterparts tended to produce only the correct solution when given the same problem to solve after instruction on the targeted concepts. This seems to suggest that although instruction may guide students to produce correct solutions, it may also create a lock-in and constrain search for new solutions. Kapur's findings are consistent with the work of Bonawitz et al. (2011), who demonstrated a similar effect on children playing with toys with versus without guidance from adults. Bonawitz et al. (2011) advanced the explanation that students tended to infer from instruction by a knowledgeable adult that all the relevant knowledge and procedures that they need to learn had already been taught during instruction. Such an inference, on one hand, increases the likelihood producing correct solutions but, on the other hand, comes at the expense of limiting exploration and search for new solutions.

What might explain the underperformance of direct instruction? The answer, in part, may simply lie in what the cognitive sciences have well-established and known for a very long time: Experts notice different things from novices. Experts tend to notice the deep structure and critical features of the domain, whereas novices attend more to the superficial features (Chase & Simon 1973; Chi, Glaser, & Farr, 1988; De Groot, 1965). It follows, then, that in starting with direct instruction, one makes the assumption that novices are prepared to notice the critical features and deep structure of the domain—an assumption that, by the very definition of a novice, is not tenable. Schwartz and colleagues (Schwartz & Bransford, 1998; Schwartz, Chase, Oppezzo, & Chin, 2011; Schwartz & Martin, 2004) have demonstrated that novices do not have the necessary prior knowledge differentiation to be able to notice and encode critical features of domain knowledge during direct instruction.

Even though direct instruction may reduce cognitive load, the benefit of a reduced cognitive of load maybe offset by the lower likelihood of novices to notice and consequently encode critical features and deep structure of the domain. Put more strongly, a reduced cognitive load may in fact be counterproductive if it allows for the noticing and encoding of features that are not critical. It is really the noticing and encoding of deep structure and critical features that makes for conceptual understanding and transfer (Bassok, 1996; Belenky & Schalk, 2014; Braithwate & Goldstone, 2015; Kaminski, Sloutsky, & Heckler, 2008, 2013).

Therefore, if the goals of learning are largely the acquisition of basic knowledge for problem solving without a commensurate understanding the concepts, let alone being able to transfer them, then direct instruction does a fairly good job. However, even this conclusion may be overly generous in the light of the longer term comparisons. Instead, if one views the acquisition of basic knowledge without deep understanding or transfer as a problem, then it is not as harsh to categorize direct instruction as an unproductive success relative to productive failure and productive success. It is important to emphasize the relative nature of the categorization; what seems productive success against one thing can seem unproductive success against another.

DISCUSSION

In this article, I interrogated the four design possibilities—unproductive failure, unproductive success, productive failure, and productive success—arising from an incommensurability between performance and learning. By examining direct instruction's CLT-grounded case against discovery learning, I argued that even though direct instruction may come across as an example of productive success in comparison with discovery learning, it could be seen as an unproductive success compared with examples of productive failure and productive success.

Several implications follow.

It is quite clear that the extremes are not very useful and that we need to abandon the dichotomy between unguided problem solving and heavily guided direct instruction. There is a large design space in between the two extremes that can be exploited to achieve optimal learning. Productive failure provides an example of how this design space can be exploited to combine the exploratory benefits of unguided problem solving and explicit instruction (Kapur & Rummel, 2012). Likewise, productive success provides an example of how students tend to learn better when their learning is situated in problems and are appropriately scaffolded during their problem solving to bring about both successful problem solving and learning (Hmelo-Silver, Duncan, & Chinn, 2007).

Furthermore, what counts as learning depends upon the context and the kinds of learning goals one commits to (Kuhn, 2007). CLT defines learning as schema acquisition resulting in a change in the long-term memory (Kirschner, Sweller, & Clark, 2006). Clearly, the definition is limited. There is more to learning than schema acquisition alone. Problem solving, inquiry, argumentation, reasoning, inventing, metacognition and self-regulation, epistemic fluency and flexibility, adaptiveness, collaboration, knowledge building, learning to learn, and so on, are all equally if not even more important goals of learning (Bereiter & Scardamalia, 2006). Therefore, the choice of instructional method needs to take into account the various types of learning goals. Conversely, instructional design prescriptions derived from a narrow definition of learning cannot be applicable more generally (Kuhn, 2007).

And this brings us nicely to the next point.

CLT needs a rethink of some its basic assumptions, without necessarily throwing out the proverbial baby with the bathwater. I focus on two assumptions: (a) cognitive load is the main mechanism of learning, and (b) higher cognitive load is monotonically bad for learning. Both assumptions are questionable. Cognitive load is but one mechanism. We also need to take into account other cognitive (e.g., attention, activation, noticing, etc.), social (e.g., collaboration, conflict, explanation and elaboration, etc.), and cultural (e.g., norms, values) mechanisms when designing for learning. Again, the choice of mechanisms embodied in a design would depend upon the learning goals.

It may also not always be true that the higher the cognitive load, the worse the learning. A higher cognitive load at times may in fact be good for driving attention and activating prior knowledge, which may in turn help students notice and learn better from instruction. There is evidence from research on productive failure suggesting that the initial problem-solving phase engenders a higher cognitive load (Kapur, 2012, 2014). Yet, as the findings described earlier show, it also leads to better learning from subsequent instruction. In other words, contrary to the predictions of CLT, cognitive load may not always be monotonically bad for learning. A certain amount of load, even high load, can be productive for learning.

Taking the variety in the types of learning goals in account together with a questioning of some of the basic assumptions of CLT will result only in a rethinking of its scope of applicability. The rethinking process has already started from within the CLT community. Kalyuga and Singh (2015) called for a reconsideration of the scope of applicability of CLT. Specifically, they argued that CLT has largely been concerned with the instructional goal of schema acquisition of basic knowledge. This, they argued, may limit the instructional recommendations arising out of CLT to context where the instructional goal is just that, and not extend to contexts where the goals that go beyond acquisition of basic knowledge schemas, such as in complex problem-based learning environments.

Finally, stepping into the long-standing instructivist–constructivist debate (Kirschner, Sweller, & Clark, 2006; Tobias & Duffy, 2009), it is perhaps worth clarifying that a commitment to a constructivist epistemology does not necessarily imply a commitment to discovery learning (Kapur, 2015). Instead, it requires a commitment to *building upon* learners' prior knowledge. However, one cannot build upon prior knowledge if one does not know what that prior knowledge is in the first place. Finding out what a learner knows and how that knowledge can be used to help them learn new concepts is central to the concept of scaffolding (Pea, 2004). Because CLT argues for heavy scaffolding right from the start, it may be worthwhile to briefly revisit the concept of scaffolding.

Scaffolding theory argues for guidance to be minimal and provided only after the learner has first been given opportunities to persist in solving a problem or a task (Wood, Bruner, & Ross, 1976). It is not surprising, then, that the notion of scaffolding originally conceived by Wood, Bruner, and Ross (1976) was eventually linked to the Vygotskian notion of the zone of proximal development (ZPD; Bruner, 1986; Vygotsky, 1978). The ZPD is defined as the "distance between the child's actual developmental level as determined by independent problem solving and the higher level of potential development as determined through problem solving under adult guidance and in collaboration with more capable peers" (Vygotsky, 1978, p. 86). Enabling the learner to bridge this gap requires the provision of support structures, which need not necessarily be in the form of a more capable person (e.g., a teacher, expert) but may also include tools, instructional facilitation, and so on (Puntambekar & Hubscher, 2005).

In both scaffolding theory and Vygotskian ZPD, one must first ascertain the limits of what learners can achieve on their own (Bruner, 1986). But it is not possible to ascertain the limits of what learners can achieve on their own without concomitantly ascertaining what it is that they cannot. Ascertaining the latter would invariably entail learners failing to solve problems or complete tasks on their own. An analysis of this failure would then provide critical information for an expert to design and administer appropriate scaffolds (Pea, 2004).

Given the centrality of scaffolding in designing for productive success, and that scaffolding, as argued earlier, necessarily entails a determination of what a learner is not able to accomplish on his or her own, it logically follows that productive success could well be conceived as a design that embodies iterative cycles of productive failure.

In the final analysis, therefore, whether one makes commitment to the instructivist or the constructivist camp, the primacy of what a learner already knows seems to be common and important to both. A direct implication for the design of instruction is to first understand the nature and limits of learners' prior knowledge, before designing appropriate guidance to build upon it. Part of the problem with CLT, as argued earlier, is that the first part of the aforementioned implication is ignored or assumed to be null. In contrast, productive failure presents a way of first engaging students in unguided problem solving to elicit what students know, especially in the failure to solve the problem, and then using this information to consolidate and assemble new knowledge.

ACKNOWLEDGMENTS

I thank my research assistants, as well as Katerine Bielaczyc, Daniel Schwartz, and Nikol Rummel, for fruitful conversations and suggestions, and I thank Clark Chinn and the reviewers for providing insightful feedback.

FUNDING

The research reported in this article was funded by grants OER50/08MK, OER03/11MK, and DEV03/14MK to the author from the Ministry of Education of Singapore.

REFERENCES

Albanese, M., & Mitchell, S. (1993). Problem-based learning: A review of the literature on its outcomes and implementation issues. *Academic Medicine, 68*, 52–81.

Bassok, M. (1996). Using content to interpret structure: Effects on analogical transfer. *Current Directions in Psychological Science, 5*(2), 54–58.

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

Belenky, D. M., & Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review, 26*(1), 27–50.

Bereiter, C., & Scardamalia, M. (2006). Education for the knowledge age: Design-centered models of teaching and instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 695–713). Mahwah, NJ: Erlbaum.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. H., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120*, 322–330.

Braithwaite, D. W., & Goldstone, R. L. (2015). Effects of variation and prior knowledge on abstract concept learning. *Cognition and Instruction, 33*, 226–256.

Bruner, J. S. (1986). *Actual minds, possible worlds*. Cambridge MA: Harvard University Press.

Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction, 22*, 61–79.

Carroll, W. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology, 86*, 360–367.
Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.
Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
Clifford, M. M. (1984). Thoughts on a theory of constructive failure. *Educational Psychologist, 19*, 108–120.
Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description and assessment data. *Educational Psychologist, 27*, 291–315.
Dean, D., & Kuhn, D. (2006). Direct instruction vs. discovery: The long view. *Science Education, 91*, 384–397.
DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology, 113*, 552–568.
De Groot, A. D. (1965). *Thought and choice in chess*. The Hague, Netherlands: Mouton.
Derry, S. J., Hmelo-Silver, C. E., Nagarajan, A., Chernobilsky, E., & Beitzel, B. (2006). Cognitive transfer revisited: Can we exploit new media to solve old problems on a large scale? *Journal of Educational Computing Research, 35*, 145–162.
Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction, 13*, 533–568.
Gagne, R. M., & Brown, L. T. (1961). Some factors in the programming of conceptual learning. *Journal of Experimental Psychology, 62*, 313–321.
Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Eds.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.
Hmelo, C. E. (1998). Problem-based learning: Effects on the early acquisition of cognitive skill in medicine. *The Journal of the Learning Sciences, 7*, 173–236.
Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*, 99–107.
Kalyuga, S., Chandler, P. A., Sweller, J. & Tuovinen, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579–588.
Kalyuga, S., & Singh, A. M. (2015). Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review*. Advance online publication. doi:10.1007/s10648-015-9352-0
Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*(5875), 454–455.
Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2013). The cost of concreteness: The effect of nonessential information on analogical transfer. *Journal of Experimental Psychology: Applied, 19*(1), 14–29.
Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*, 379–424.
Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*, 523–550.
Kapur, M. (2011). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science, 39*, 561–579.
Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science, 40*, 651–672.
Kapur, M. (2013). Comparing learning from productive failure and vicarious failure. *The Journal of the Learning Sciences*, *23*, 651–677. doi:10.1080/10508406.2013.819000
Kapur, M. (2014). Productive failure in learning math. *Cognitive Science*, *38*, 1008–1022.
Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, *1*(1), 51–65.
Kapur, M., & Bielaczyc, K. (2011). Classroom-based experiments in productive failure. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2812–2817). Austin, TX: Cognitive Science Society.
Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences, 21*(1), 45–83.
Kapur, M., & Rummel, N. (2012). Productive failure in learning and problem solving. *Instructional Science, 40*, 645–650.
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist, 41*, 75–86.
Kittel, J. E. (1957). An experimental study of the effect of external direction during learning on transfer and retention of principles. *Journal of Educational Psychology, 48*, 391–405.
Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*, 661–667.
Kuhn, D. (2007). Is direct instruction the answer to the right question? *Educational Psychologist, 42*, 109–114.
Kuhn, D., & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science, 16*, 866–870.
Loibl, K., & Rummel, N. (2013). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science, 42*, 305–326.
Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction, 34*, 74–85.
Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*, 14–19.
Mergendoller, J. R., Maxwell, N. L., & Bellisimo, Y. (2006). The effectiveness of problem-based instruction: A comparative study of instructional method and student characteristics. *Interdisciplinary Journal of Problem based Learning, 1*, 49–69.
Miller, G., & Gildea, P. (1987). How children learn words. *Scientific American, 257*, 94–99.
Ohlsson, S. (1996). Learning from performance errors. *Psychological Review, 103*, 241–262.
Paas, F. (1992). Training strategies for achieving transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*, 429–434.
Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*, 423–451.
Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist, 40*(1), 1–12.
Puntambekar, S., & Kolodner, J. L. (2005). Toward implementing distributed scaffolding: Helping students learn from design. *Journal of Research in Science Teaching, 42*, 185–217.
Roll, I., Aleven, V., & Koedinger, K. (2011). Outcomes and mechanisms of transfer in invention Activities. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2824–2829). Austin, TX: Cognitive Science Society.
Schmidt, H. G., Loyens, S. M. M., Van Gog, T., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 92*, 91–97.
Schmidt, H. G., van der Molen, H. T., te Winkel, W. W. R., & Wijnen, W. H. F. W. (2009). Constructivist, problem-based learning does work: A meta-analysis of curricular comparisons involving a single medical school. *Educational Psychologist, 44*, 227–249.
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*, 475–522.

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.

Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Sciences, 3*, 1–5.

Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-Based Learning, 3*(1), 44–58.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

Sweller, J. (2009). What human cognitive architecture tells us about constructivism. In S. Tobias and T.M. Duffy (Eds.)., *Constructivist theory applied to instruction: Success or failure*? (pp. 127–143). New York, NY: Taylor and Francis.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*, 351–362.

Tobias, S., & Duffy, T. M. (2009). *Constructivist instruction: Success or failure*. New York, NY: Routledge.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*, 334–341.

Vernon, D. T. A., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine*, *68*, 550–563.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

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

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89–100.