Productive failure in learning from generation and invention activities

Manu Kapur · Nikol Rummel

This special issue comes amidst persistent and recently reignited debates about how and when to structure learning and problem-solving activities (Kirschner et al. 2006; Tobias and Duffy 2009). At the heart of the work reported in this special issue lies the incommensurability between learning and performance; that is, conditions that maximize performance in the short term *may not necessarily* be the ones that maximize learning in the long term (Schmidt and Bjork 1992). Two possibilities for designing instruction emerge.

First is the possibility of designing conditions that maximize performance in the short term and that also maximize learning in the long term. Let us call such design efforts designing for *productive success*. Productive success research examines different instructional designs for structuring learning and problem-solving activities with the goal to achieve both improved performance on the learning tasks and sustainable learning. In fact, immediate success is often thought of as a proxy for long-term learning gains—as students who can perform well in the short run, are more likely to perform better on delayed assessments. Indeed, the majority of research in the cognitive and learning sciences speaks to this paradigm; and rightly so, because understanding conditions under which learning and problem-solving activities can lead to productive success is an important line of research. For example, cognitive load theory proposes direct instruction in the form of well-designed worked examples with the goal to avoid cognitively overloading learners, thus enabling learners to succeed in performing well on the learning tasks and to learn more (Sweller 2010). Constructivist approaches that fall into the genre of *guided inquiry* involve scaffolded activities initially to engender learning, with a gradual fading of scaffolding as learners gain expertise (Puntambekar and Hübscher 2005; Schmidt et al. 2007).

M. Kapur (✉)
Program of Curriculum, Teaching and Learning, Learning Sciences Laboratory, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616, Singapore
e-mail: manu.kapur@nie.edu.sg

N. Rummel
Institute of Educational Research, Ruhr-Universität Bochum, Universitätsstraße 150, GA 1/62, 44801 Bochum, Germany
e-mail: nikol.rummel@rub.de

However, accumulating evidence from a broad set of interventions suggests that immediate and delayed performances are, in fact, not as correlated as is often believed (Needham and Begg 1991; Richland et al. 2009; Strand-Cary and Klahr 2008; VanLehn et al. 2003). Thus, the second possibility, and one that concerns this special issue, is one of designing conditions that may not maximize performance in the short term but in fact maximize learning in the longer term. One such design effort is called *productive failure* (Kapur 2008; Kapur and Kinzer 2009). Designing for productive failure involves two phases: a generation (or invention) phase followed by a consolidation (or instruction) phase. The generation phase affords opportunities for students to generate and explore the affordances and constraints of multiple representations and solution methods (RSMs) to novel, complex problems. The consolidation phase affords opportunities for comparing and contrasting, organizing, and assembling the relevant student-generated RSMs into canonical RSMs (Kapur and Bielaczyc 2012).

In contrast to productive success, the concept of productive failure (PF), much as it is intuitively compelling, remains largely underdetermined and under-researched (Clifford 1984; Schmidt and Bjork 1992). Building on the seminal work of Schmidt and Bjork (1992) on "desirable difficulties", as well as Schwartz and colleagues' work on preparation for future learning (Schwartz and Bransford 1998) and inventing to prepare for learning (Schwartz and Martin 2004), there is now a growing body of work that examines conditions under which PF designs can be just as or perhaps even more effective than conditions designed for productive success (see all of the papers in this special issue).

However, work on PF is still in its infancy and raises many significant questions. This special issue is a first effort at consolidating the current state of research on the role of PF in learning and problem solving. This editorial is followed by four empirical studies carried out by independent research groups in Singapore, Germany, Canada, and the USA, and a discussant piece by Professor Allan Collins. Together, these studies extend earlier work on PF by demonstrating the efficacy of PF in different learning contexts and domains, with different samples and age groups. Moreover, they shed light on the role of factors that support student generation and exploration of RSMs such as collaborative role scripts (Westermann and Rummel 2012), metacognitive prompts (Roll et al. 2012), and ability-based group composition (Wiedmann et al. 2012).

Constituent papers

The issue starts off with Kapur's study with ninth-grade mathematics students on learning the concept of variance. Consistent with his prior work (e.g., Kapur 2010, 2011), Kapur (2012) found that PF students, who initially attempted to solve problems on the novel concept in small groups, generated a diversity of formulations for variance but were unsuccessful in developing the canonical formulation. Other than providing affective support and encouragement to PF students to persist in generation, no other support was provided. After the generation phase, the teacher compared and contrasted student-generated solutions and taught the canonical concept. In contrast, students who initially received direct instruction (DI) solved practice problems successfully by relying on the canonical formulation of variance taught to them. On the posttest however, PF students significantly outperformed DI students on conceptual understanding and transfer without compromising procedural fluency. Interestingly, Kapur also replicated the finding that the diversity of student-generated RSMs is significantly correlated with how much students learn from PF. A direct implication of this finding is to examine ways of supporting

students in RSM generation that go beyond affective support that may even further increase the efficacy of PF. Indeed, the rest of the papers in this special issue contribute precisely to this question.

Westermann and Rummel (2012) supported PF students during their initial problem-solving and RSM generation with a role-script called Think Ask Understand (TAU). In a four-week, in vivo experiment with 76 university students, Westermann and Rummel compared TAU to a direct instruction (DI) condition. Their study was conducted in a re-learning situation on four topics of mathematical analysis. Participating students wanted to repeat the topics for upcoming examinations. Westermann and Rummel found that students in the TAU condition outperformed students in the DI condition in all weeks but the first. Process data further indicated that students collaborated fruitfully in accordance with the role script and increasingly internalized the script. The results suggest that the more students were familiarized with TAU, the better their learning outcomes became. The improved collaboration may thus have paved the way for increased learning from the subsequent instruction. Importantly, these findings call into question whether all support must be delayed. The primary issue may not be whether or not to provide support, but rather when to provide support and which type of support to employ. Moreover, this study provides evidence that delaying instruction can also promote learning in relearning situations and at the university level. Finally, the results suggest that implementing and investigating learning conditions over a longer period to familiarize learners with the new method may maximize the effect on learning.

Roll et al. (2012) conceptualized support for undergraduate students' RSM generation in the form of metacognitive scaffolding. Students enrolled in a first-year physics lab course engaged in a PF activity on an advanced topic in data analysis, namely, the uncertainty in the slope of best-fitted lines. Roll and colleagues identified exploratory analysis, self-explanation, peer interaction, and evaluation as key strategies that students fail to engage in, and designed metacognitive supports in the form of question, self-explanation, and peer-interaction prompts to support students during the process of generating measures for uncertainty in the slope of a best-fit line. Students in the unguided invention (UI) condition received conventional invention activities where they were asked to invent methods for calculating uncertainties in best-fitting lines; students in the guided invention (GI) condition received metacognitive scaffolding in the form of the described prompts. GI students invented methods that included more conceptual features and ranked the given datasets more accurately than those in the UI condition, although the quality of their mathematical expressions was not improved. At the learning strategy level, GI students showed more and better instances of unprompted self-explanations and they revised their methods more frequently—even on components of the task that were not supported by the metacognitive prompts—than the UI students. These results suggest that process guidance in the form of metacognitive scaffolding augments the inherent benefits of invention activities and can lead to gains at both domain and learning strategy levels.

Finally, Westermann and Rummel (2012) investigated how the effectiveness of formula invention activities in math may be mediated by composition of the small groups in terms of their members' mathematical ability. In two studies, small groups of undergraduate students engaged in a variance learning task based on the one used in the Kapur (2012) study. Results suggested that groups may need at least one member with high math ability to take advantage of the invention learning setting. Groups consisting of both high and low math ability members generated a broader range of RSMs during the invention task than than all-low ability and all-high ability groups, and this related to better uptake of the canonical formula when it was presented after the invention task.

Table 1 Overview of empirical studies in this special issue

Authors	Focus/research question	Participants	Domain/learning content	Status	Setting
Kapur	Replicate PF and the relevance of RSM diversity	9th-grade students	Statistics: concept of variance	New concept	Classroom-based
Westermann and Rummel	Process guidance for PF in the form of role script	University students (Masters level)	Mathematics: analysis	Relearning	Optional tutorials following previous lecture
Roll et al.	Process guidance for PF in the form of metacognitive scaffolding	University students (undergraduate)	Statistics: uncertainty in the slope of best-fitting line	New concept	Laboratory-based
Wiedmann et al.	Effect of math ability/group composition	University students (undergraduate)	Statistics: concept of variance	New concept	Classroom-based, laboratory

Table 1 gives a compact overview of the studies in this special issue highlighting their commonalities and differences. As evident from the table, the studies cover a broad array of participants from 9th-graders to university students at the Master's level. The content domain of all studies was mathematics and statistics. With the exception of the Westermann and Rummel study, which examined PF in a re-learning situation, all studies examine PF in the context of initial learning of a new concept.

Themes across the four papers

Even though the constituent papers report findings from studies that differ in certain ways as discussed above, three themes that emerge from them are noteworthy. The first deals with the role of (unsuccessful) RSM generation in learning. The second deals with how generation tasks are designed specifically to activate prior knowledge. And the third deals with the nature of support provided during the generation phase. We discuss each in turn.

An important theme across the four studies concerns the role of generation attempts in learning. The central claim here is that while processes of RSM generation may not lead to canonical solutions, they can still be germane for learning. While not all the studies explicitly define or conceptualize *failure*, there is a common commitment to designing tasks and activity structures that afford students opportunities to generate RSMs without expecting them to be able to develop the canonical RSMs to the given problems. In other words, success in developing the canonical solution to the problems is not considered a necessary condition for learning. Furthermore, none of the studies makes the claim that failure is a necessary, indispensable condition for learning. It would be interesting to compare these findings with VanLehn et al. (2003), who found that learning depends on students' *perception* of failure, rather than actual failure.

A second common theme concerns the design of the generation task. The goal of the RSM generation is to activate prior knowledge, whether in the initial learning phase of learners trying to learn a new concept or in a re-learning situation where learners are required to re-activate knowledge they should have acquired before. We design generation tasks based on the assumption that while students may not have canonical knowledge about

the targeted concept, this does not mean that they do not have or could not generate initial ideas about that concept. The challenge is to design problem-solving contexts that can activate students' thinking about the concept even if they have not formally learned it yet. Not all generation tasks may be useful – but the ones used in these studies were, and that may be because they were chosen to represent topics for which students could use other knowledge and skills. Researchers attempted to find a novel concept that students did not know its canonical formula, but they would possess relevant skills and knowledge from other areas that they could apply to the RSM generation process.

Finally, a third theme emerging from these papers is that all the studies of this special issue show that support of various kinds can help students take advantage of the affordances of learning by generation and invention. Kapur provides affective support, Westermann and Rummel provide role-play support for student interaction, and Roll and colleagues provide metacognitive support. The support provided in these three studies is mainly content-general in that it does not provide specific mathematical content to students during the generation phase. Wiedmann and colleagues, on the other hand, conceptualize peer support by high-ability group members as a possible support mechanism for students during the generation phase. Taken together, the present studies underline an important distinction between providing support and direct instruction. All of the mentioned support measures were designed to help students, but without giving or telling them the targeted content knowledge directly. In other words, support structures are designed to keep students from unproductive failure experiences, while leaving a central portion of student activity unstructured, thus allowing for productive failure.

Conclusion

In consolidating what we know about the mechanisms effective in PF and when advocating more research on PF, it is neither our intention nor our proposition that one must always design learning by delaying instruction. Instead, what is being proposed here is that as a field we stand to gain more if we engage in research that seeks to understand both productive failure as well as productive success. Such a dual focus may be expected to advance the field in ways that neither single focus alone can (Kapur and Rummel 2009). As these lines of inquiry push back against and inform each other, we will generate not only better understandings of productive failure and productive success, but also better understandings of the other two design possibilities; that is, conditions under which designs lead to: (a) *unproductive success*—an illusion of performance without learning, which may happen when a learner is able to successfully solve problems without a proper or full understanding of the underlying conceptual basis, and (b) *unproductive* failure—which may happen when a learner is neither able to solve problems successfully nor demonstrate an understanding of the underlying conceptual basis as a result of his or her attempts at problem-solving. Most importantly, this line of research can help decouple learning from success, and better understand what mechanisms and features of instruction yield learning.

References

Clifford, M. M. (1984). Thoughts on a theory of constructive failure. *Educational Psychologist, 19*(2), 108–120.
Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379–424.

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523–550.

Kapur, M. (2011). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science, 39*(4), 561–579.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*. doi: 10.1007/s11251-012-9209-6.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences, 21*(1), 45–83.

Kapur, M., & Kinzer, C. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Collaborative Learning, 4*, 21–46.

Kapur, M., & Rummel, N. (2009). The assistance dilemma in CSCL. In A. Dimitracopoulou, C. O'Malley, D. Suthers, & P. Reimann (Eds.), *Computer Supported Collaborative Learning practices-CSCL2009 community events proceedings* (Vol. 2, pp. 37–42). New Brunswick: International Society of the Learning Sciences.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work. *Educational Psychologist, 41*(2), 75–86.

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory and Cognition, 19*(6), 543–557.

Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in complex learning environments: What have we gained and what have we missed. *Educational Psychologist, 40*(1), 1–12.

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied, 15*(3), 243–257. doi:10.1037/a0016496.

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

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, 42*(2), 91–97.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–522.

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

Strand-Cary, M., & Klahr, D. (2008). Developing elementary science skills: Instructional effectiveness and path independence. *Cognitive Development, 23*(4), 488–511. doi:10.1016/j.cogdev.2008.09.005.

Sweller, J. (2010). What human cognitive architecture tells us about constructivism. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure* (pp. 127–143). New York: Routledge.

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

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction, 2*(3), 209–249.

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

Wiedmann, M., Leach, R. C., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science*. doi:10.1007/s11251-012-9204-y.