Title	Investigating analogical problem posing as the generative task in the *Productive Failure* design
Authors	Jun Song Huang, Rachel Lam and Manu Kapur
Source	*12th International Conference of the Learning Sciences (ICLS 2016), Singapore, 20-24 June 2016*
Published by	International Society of the Learning Sciences

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

Original citation: Huang, J. S., Lam, R., & Kapur, M. (2016). Investigating analogical problem posing as the generative task in the *Productive Failure* design. In C. –K. Looi, J. Polman, U. Cress & & P. Reimann (Eds.), *Transforming learning, empowering learners: The International Conference of the Learning Sciences (ICLS) 2016, Volume 2* (pp. 998-1001). Singapore: International Society of the Learning Sciences.

© 2016 International Society of the Learning Sciences

Archived with permission from the copyright owner.

Investigating Analogical Problem Posing as the Generative Task in the *Productive Failure* Design

Jun Song Huang, National Institute of Education, Nanyang Technological University, junsong.huang@nie.edu.sg
Rachel Lam, National Institute of Education, Nanyang Technological University, rachel.lam@nie.edu.sg
Manu Kapur, Hong Kong Institute of Education, mkapur@ied.edu.hk

Abstract: Research on *Productive Failure* and preparatory mechanisms has consistently demonstrated a positive learning effect when students generate problem solutions before receiving formal instruction. However, it has been less examined whether the effect still holds when the generative task does not involve problem solving. Using a 2x2 experimental design, this study investigated the effects of generative tasks that involve analogical problem posing (without solving) on learning and transfer. Pedagogical sequence (i.e., generation-first or instruction-first) and type of analogical reasoning task (i.e., generating one's own analogical problems or generating analogical mappings between given analogical problems) were the two factors manipulated. Preliminary analysis revealed no multivariate effects of the factors. Thus, we discuss the learning mechanisms enacted by analogical reasoning, reliability of the instruments, and the participants' prior condition as possible reasons and to inform future studies.

Introduction

While teachers often provide students direct instruction before having them generate their own ideas, problem solutions, or conceptions of topics, recent studies on *Productive Failure* (Kapur, 2008, 2010) have shown the benefits of allowing students to generate first, at the onset of learning, and afterward receive direct instruction (Kapur, 2014; Loibl & Rummel, 2014). These studies show that students who generate (G) first and receive instruction (I) later (which we refer to as the G-I sequence) experience an initial performance dip after the generation phase, yet have subsequent success in learning after instruction. This is the crux of the *Productive Failure* phenomenon. Schwartz and Martin (2004) explain that a generative type of task prepares students for their future learning in a subsequent phase of direct instruction, such as a lecture.

The existing *Productive Failure* work uses problem solving as the generative task, and Kapur (2012) has found that the number of (typically suboptimal) solutions that students generate is positively correlated to learning outcomes. However, there are other types of generative tasks and it is an open question as to whether the Productive Failure phenomenon would still apply. For example, prior work has shown that the generative activities of summarizing a text (Grabowski, 2004) and self-explaining (Chi, De Leeuw, Chiu, & LaVancher, 1994) produce better outcomes than less generative activities (highlighting and reading the text twice, respectively). However, these works did not involve a manipulation of pedagogical sequence.

An extended question is whether activities that offer more degrees of freedom in generation may lead to better learning and transfer. By this, we mean tasks that have multiple possible canonical answers. It makes sense that in general, learning activities that cognitively engage students in similar ways (e.g., generating ideas, rules, principles, solutions, etc.) would produce similar learning outcomes (Chi, 2009). However, studies by Loibl and Rummel (2014) and Kapur (2014) both suggest that different types of generative tasks may lead to different learning effects. For example, in Kapur's (2015) work on problem posing, which has more degrees of freedom as students may pose problems with different problem goal states, results showed that posing problems led to better transfer outcomes than solving problems, but worse conceptual knowledge outcomes. Thus, differing degrees of freedom may lead to learning and transfer trade-offs, and testing these effects is necessary.

In this study, we examine analogical problem posing (Gentner & Holyoak, 1997) as the generative task, because analogical reasoning—the underlying cognitive process—has a substantial possibility for varying the degrees of freedom. In analogical reasoning, leaners perceive relational similarities (Gick & Holyoak, 1980) between two problems, recognize them as an analogy, and transfer prior knowledge (i.e., problem solution) from one problem (i.e., source) to the other (i.e., target). Holyoak and Koh (1987) regard relational similarities as problem structures that are functionally relevant to solving the source and target problems. Analogical reasoning is found to help students focus on relational similarities rather than on idiosyncratic surface features between two situations (Gentner & Markman, 1997). For example, Gentner, Loewenstein and Thompson's (2003) study on the transfer of negotiation strategies shows that explicitly comparing source and target cases that share a common underlying principle can be illuminating even if the common principle is only partially understood in either case.

We used analogical reasoning in two different ways to correspond to different degrees of freedom in generation. One type of activity is called Student-Generated Analogous Problems (SGAP). SGAP requires students to generate their own analogical problems to a given problem, compare and contrast the similarities and differences, and make modifications of their analogies. The other type of activity is called Teacher Given Analogous Problems (TGAP). TGAP requires students to generate and encode relational similarities between the source and target problems (which are both given by the teacher). This process is called analogical mapping (Gentner & Markman, 1997) or analogical encoding (Gentner et al., 2003). The SGAP and TGAP tasks differ in degrees of freedom in generation; in the SGAP task, one problem may have many possible analogues whereas in the TGAP task, the similarities between two given problems are limited and fixed. Thus, to extend the understanding of the *Productive Failure* phenomenon, in this study we used SGAP and TGAP tasks to differentiate generativity of the tasks.

The main research questions addressed (a) whether the effect of pedagogical sequence holds in analogical problem posing and (b) whether the degrees of freedom in generation in analogical reasoning tasks affect the outcomes of pedagogical sequence.

Methods

A 2x2 experimental study was conducted using the factors of (a) pedagogical sequence (Generation first-Instruction later, G-I, vs Instruction first-Generate later, I-G) and (b) generative task (student-generated analogous problems, SGAP, vs teacher-given analogous problems, TGAP). Thus, there were four conditional groups in this study: SGAP-I (i.e., students generating analogous problems, followed by instruction), TGAP-I (i.e., students generating analogical mappings for teacher given analogous problems, followed by instruction), I-SGAP, and I-TGAP.

Sample

The study was conducted in an all-girls high-achieving secondary school in Singapore. A total of 117 seventh-grade girls from five classes took part in the study at the end of the school year. Participants were randomly assigned to conditions, which took place in two different instructional environments (1). Six students had missing data, representing 5.13% of the total participants. Therefore, our total sample size was 111: 25 in SGAP-I, 29 in TGAP-I, 26 in I-SGAP and 31 in I-TGAP.

Materials and measurements

The learning goal was to formulate a system of linear equations for word problems. Three measurements were collected on student prior ability. As formulating equations for word problems requires math knowledge and language comprehension ability, the participants' national exam results from sixth grade (referred to as PSLE) were collected as a measurement of their general academic ability (2). We collected the participants' analogical reasoning ability (i.e., AR ability) using domain-free pictorial items isomorphic to Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998). The instrument measured the participants' domain-free general ability. The Cronbach's α of the 16-item AR test was .65. The participants' prior knowledge specific to algebra (i.e., prior algebraic knowledge) was measured by a 10-item pretest concerning the conceptual knowledge of algebra and the ability to formulate linear equations for one-variable and two-variable word problems. Of the ten items, one was excluded from data analysis due to its low reliability. The Cronbach's α of the final 9-item instrument was .64.

The posttest comprised three measures. A six-item learning instrument measured the participants' ability to formulate a system of linear equations for word problems with two and three unknown variables. One item was excluded from data analysis due to its low reliability. The Cronbach's α of the final five-item instrument was .55. A posttest with four items on conceptual knowledge measured the participants' understanding of algebraic representations involving three variables. One item was excluded from data analysis due to its low reliability. The Cronbach's α of the final three-item instrument was .75. A six-item posttest on transfer measured the participants' ability to formulate equations for word problems involving four unknown variables, nonlinearity, and inequalities. The Cronbach's α of this instrument was .60.

Procedures

The study involved four sessions and each session lasted about one hour. All participants took the AR and algebra pretests in Session 1 and posttest in Session 4. The manipulations occurred in Sessions 2 and 3.

In the SGAP-I condition, the SGAP task was implemented in Session 2. Students were given a word problem with two unknown variables and asked to generate an analogous problem that involved three unknown variables. In the instruction phase (Session 3), the teacher presented the step-by-step procedures on how to formulate a system of linear equations for a word problem with two unknown variables. She then showed another worked example and compared the two problems to highlight their similarities and differences. The participants then practiced on four new word problems (all with two unknown variables). The answers were given at the end of the session. For the TGAP-I condition, the TGAP task was implemented in Session 2. Students were given two word problems with the same context; one involved two unknown variables and the other three unknown variables. Rather than generating analogous problems, they identified the similarities and differences between the problems. In Session 3, the same instruction was received as in the SGAP-I condition.

The I-SGAP and I-TGAP conditions were similar to the respective SGAP-I and I-TGAP conditions, except that the instruction was given in Session 2 and the generative task (SGAP or TGAP) in Session 3. To minimize teacher effects, the same teacher taught the instruction phase in all conditions. Similarly, the same researcher and teacher jointly facilitated all the SGAP and TGAP tasks. All participants were also given the same generic training on analogies, which was not specific to mathematics.

Data analysis and initial findings

The descriptive statistics for the PSLE and pre- and posttest scores are presented in Table 1. ‘

Table 1: Summary of PSLE, AR ability, pretest, and posttests

	Max Score	SGAP – I		TGAP – I		I – SGAP		I – TGAP	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
PSLE	300	262.1	4.93	258.3	7.74	263.5	4.65	260.2	5.27
AR ability	16	8.56	2.93	8.17	2.93	9.62	2.90	9.03	2.63
Prior algebraic knowledge	90	40.60	19.41	38.10	13.73	47.33	19. 16	39.85	15.43
Posttest–Learning	50	28.36	7.83	24.62	8.86	27.23	8.66	26.35	8.75
Posttest–Conceptual knowledge	30	9.60	8.16	5.24	7.64	9.08	8.88	9.29	8.42
Posttest–Transfer	60	20.36	8.47	17.02	8.54	24.12	9.88	22.06	8.63

SGAP: Student-Generated Analogous Problems; TGAP: Teacher-Generated Analogous Problems; I: Direct Instruction

We conducted a One-Way ANOVA to test the mean differences in prior knowledge across the four conditional groups. There were no significant difference in AR ability, $F(3, 107)=1.308$, $p=.276$, nor prior algebraic knowledge, $F(3, 107)=1.512$, $p=.216$. There was a difference in PSLE, $F(3, 107)=4.027$, $p=.009$.

We conducted a two-way MANCOVA with posttest scores on learning, conceptual knowledge and transfer as the three dependent variables, pedagogical sequence and generative task as the two factors, and PSLE, AR ability and prior algebraic knowledge as the covariates. The data analysis revealed that pedagogical sequence, generative task and their interaction did not have overall multivariate effect (Wilk's Λ from .95 to .984, p from .153 to .642, partial η^2 from .016 to .050). To inform future studies, the data were explored further. Some speculations from this exploration are discussed in the next section.

Conclusion and discussion

This study investigated the *Productive Failure* phenomenon by testing the effects of (a) pedagogical sequence and (b) degrees of freedom of generation in a task via analogical problem posing, thereby extending prior work, using a 2x2 experimental design. We used PSLE, AR ability and prior algebraic knowledge as covariates for the following reason. Prior research on *Productive Failure* and analogical transfer suggests that such measures of prior knowledge/ability affect learning and transfer. In particular general academic ability (in our study, PSLE score), domain-free general ability (e.g., AR ability), and domain-specific knowledge (e.g., prior algebraic knowledge) have shown effects on post performance, thus, we deemed it important to include these in our model.

A two-way MANCOVA test did not reveal any significant multivariate effect of the two factors or their interaction. From the data exploration, we speculate that the participants' prior knowledge/ability (e.g., prior algebraic knowledge, PSLE and AR ability) might have had strong enough effects on learning and transfer that they possibly overshadowed any effects of the conditional manipulation. Another interpretation is that the learning mechanisms manipulated by the two factors were not strong enough to be detected. For example, while pedagogical sequence, generative task and their interaction did not have overall multivariate effect (Wilk's Λ from .95 to .984, p from .153 to .642, partial η^2 from .016 to .050), prior algebraic knowledge had strong multivariate effect (Wilk's $\Lambda=.826$, $p=.000$, partial $\eta^2=.174$). If this speculation is true, it suggests that compared to the problem-posing task examined by Kapur (2015), an analogical problem posing task might be weaker in enacting learning and transfer mechanisms. The speculation will be examined when we analyze the artifacts generated by the participants in the SGAP and TGAP tasks.

The relative low reliability of the instruments, especially the posttest instruments on learning (Cronbach's α =.55) and transfer (Cronbach's α =.60), might have also made it difficult to detect any learning and transfer effects that were induced by the two factors. For example, in our data exploration, we conducted the pair-wise comparison between SGAP-I vs TGAP-I with prior algebraic knowledge as a covariate. The univariate analysis showed a main effect of generative task on conceptual knowledge (p=.051, partial η^2=.073) with SGAP-I condition outperforming TGAP-I condition. In this model, the variance could be explained by the factor (b) generative task, but the overall univariate model was not significant (p=.144, partial η^2=.073). Thus, we will improve the reliability of the instruments in our future studies.

Another reason for the non-significant multivariate effect could be the profile of our participants. The participants in this study were from an all-girls high-achieving school. It is possible that differences in outcomes could not be detected due to a strong effect of their high prior ability. We aim to conduct future studies using a more general population of students (i.e., represented a range of ability levels) to test this effect.

In summary, we attempted to extend the *Productive Failure* design by replacing the problem solving task with analogical problem posing tasks. The data did not show an overall multivariate effect of pedagogical sequence and generative task. This could due to multiple reasons, such as how strongly the learning mechanisms were enacted, how reliable the instruments were, and how high-achieving students' prior ability influenced their learning. In our further analysis, we will examine the types of analogical problems and analogical mappings generated by the participants in the four conditions to understand the mechanisms enacted by each conditional group.

Endnotes

(1) The SGAP-I and TGAP-I groups were hosted in a lecture hall and the I-SGAP and I-TGAP groups in a computer lab.
(2) The national exam is called the Primary School Leaving Examination (PSLE). The PSLE result refers to the T-score, the relative rank of a student's performance compared to all the other students in the whole cohort.

References

Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework of differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73-105.

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Gentner, D., & Holyoak, K. J. (1997). Reasoning and learning by analogy. *American Psychologist, 52*(1), 32-34.

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

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*(1), 45-56.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Grabowski, B. (2004). Generative learning contributions to the design of instruction and learning. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed., pp. 719-743). Mahwah, NJ: Lawrence Erlbaum Associates.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15*(4), 332-340.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science, 40*(4), 651-672.

Kapur, M. (2014). Productive Failure in learning math. *Cognitive Science, 38*, 1008-1022.

Kapur, M. (2015). The preparatory effects of problem solving versus problem posing on learning from instruction. *Learning and instruction, 39*, 23-31.

Loibl, K., & Rummel, N. (2014). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science, 42*, 305-326.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.

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.

Acknowledgments

This work is funded by the National Institute of Education's research grant (OER 02/15 HJS). We thank the participating students, the teacher who conducted the instruction section, and the researcher and the teacher who jointly facilitated the SGAP and TGAP sessions. We thank school partner (Ms Chew Meek Lin), collaborator (Madam Kum Fong FOO), and research assistant (Ms Lena HENG) for their respective contributions.