Combining Generation and Expository Instruction to Prepare Students to Transfer Big Ideas Across School Topics

Inga Glogger (Organizer), University of Freiburg, Germany, glogger@psychologie.uni-freiburg.de
Lennart Schalk, ETH Zurich, Switzerland, lennart.schalk@ifv.gess.ethz.ch
Claudia Mazziotti, Ruhr-Universität Bochum, Germany, claudia.mazziotti@rub.de,
Nicole R. Hallinen, Stanford Graduate School of Education, hallinen@stanford.edu
Armin Barth, Ralph Schumacher, ETH Zurich, Switzerland,
Email: armin.barth@ifv.gess.ethz.ch, ralph.schumacher@ifv.gess.ethz.ch
Katharina Gaus, Alexander Renkl, University of Freiburg, Germany,
Email: katharina.gaus@web.de, renkl@psychologie.uni-freiburg.de
Katharina Loibl, Nikol Rummel, Ruhr-Universität Bochum, Germany,
Email: katharina.loibl@rub.de, nikol.rummel@rub.de
Doris B. Chin, Kristen P. Blair, Daniel L. Schwartz, Stanford Graduate School of Education,
Email: dbchin@stanford.edu, kpilner@stanford.edu, danls@stanford.edu
Katherine McEldoon (Discussant), Learning Sciences Institute, Arizona State University, k.mceldoon@asu.edu

Abstract: Approaches combining generation and expository instruction have been shown to be beneficial for transfer. This symposium focuses on the approaches of inventing and productive failure. Both approaches delay expository instruction. Students are asked to work on a generation activity with a problem or with contrasting cases. This activity is aimed to foster transfer of the knowledge gained during subsequent expository instruction. The goal of the symposium is to contribute to a better understanding of how the generation activity is best designed to foster transfer. The contributions of this symposium varied the amount and type of generating by varying the materials (e.g., content of cases), the task (e.g., inventing vs. self-explaining), or the setting (collaborative vs. individual) of the generation activity. Mediating processes are also addressed. The symposium will clarify moderating conditions and mediating processes of how generation activities combined with expository instruction deploy their full potential.

Overall Focus and Major Issues of the Symposium

Generative learning activities as well as expository instruction have their unique advantages for learning (Lee & Anderson 2013). Attempts to combine the two have been shown to be beneficial for transfer (Kapur, 2012; Schwartz & Martin, 2004). Transfer of big ideas across school topics is a much wanted outcome of instruction, but hard to achieve (Barnett & Ceci, 2002). The big ideas studied in the different contributions to this symposium comprise linear functions (Schalk et al.), ratios in science and mathematics (Glogger et al., Hallinen et al.), and fraction expansion (Maziotti et al.) - all of which have great significance for various school subjects. In this symposium we focus on two forms of combining generation and expository instruction to foster transfer: First, *inventing* asks learners to generate a principle or method that allows evaluating several cases from the learning domain. 'Cases' in the present studies refer to examples of a domain such as the diagram of a linear function with a positive gradient, or buses with densely packed passengers as cases of a type of density. The cases usually vary critical features of the domain so that contrasting the cases helps to identify and understand the critical features. The inventing with contrasting cases activity is used as preparation for future learning from expository instruction. Second, in the approach of productive failure, the generative phase includes a complex problem. Learners usually fail to solve this problem in a canonical way. However, the combination with subsequent expository instruction is theorized to make the failure productive (Kapur, 2012). The expository instruction encompasses contrasting of suboptimal (students') and canonical solution methods.

The benefit in transfer of these activities could be based on the deep processing of good examples (or solution methods) in an early stage of skill acquisition (Lee & Anderson, 2013). The deep processing could result in a richly elaborated mental representation of the examples. At the same time, the understanding of a deep structure, an abstract principle that "interconnects" the examples (and can be abstracted by contrasting and comparing), can be the basis for transfer (Barnett & Ceci, 2002). Contrasting cases were shown to help recognizing deep similarities between cases and abstracting generalizable principles (e.g., Alfieri, Nokes, & Schunn, 2013). The deep processing and abstracting of principles prepares to learn from the subsequent expository instruction and to transfer "big ideas" across topics. The contributions to this symposium focus on ways how to best foster deep processing and abstracting deep structures. They analyze primarily experimental variation of the generation phase.

Major issues of this set of studies refer to the questions of what conditions moderate and what processes mediate the effectiveness of different combinations of generation and expository instruction.

Moderating conditions can be the type and amount of generation. The type and amount of generation can vary depending on the task, the materials, and the setting. Regarding the task, the level of generation is highest in an inventing or productive failure task (all contributions encompass such a condition). For example, an index to evaluate the density in buses has to be invented. A somewhat lower amount of generation is required when students self-explain already evaluated cases (Schalk et al.). They still have to think towards the abstract principle that forms the basis for the evaluation of the presented cases. The next lower level of generation is to self-explain a completely worked example of the inventing problem (index and resulting evaluation of cases are provided, Glogger et al.). The type and amount of generation can also vary with the specifics of the material, here the characteristics of the cases. Characteristics of the cases that were varied in the studies of this symposium include the concreteness of the cases (idealized, abstract vs. concrete cases in Schalk et al.) and contents of the cases (isolate the main effects of key variables or include main effect and interaction of variables in Hallinen et al.). Finally, the kind of generation can vary depending on the setting such as collaborative or individual work, addressed in the study of Mazziotti et al.

Mediating processes can explain how the generation phase affects processing of the subsequent expository instruction and, in the end, how transfer is fostered. Processes discussed in this symposium are motivational such as self-efficacy and cognitive such as cognitive load (Glogger et al.).

The potential significance of each contribution is summarized in the following: Schalk, Barth, and Schumacher could show that the kind of material and the kind of task interact when students learn about linear functions: Specifically, their results suggest that self-explanation prompts should be combined with concrete cases, while invention prompts should be combined with idealized cases. They did not find differences between the generative preparatory conditions and a tell-and-practice condition.

Glogger, Gaus, and Renkl found that a higher level of generation (inventing) led to deeper encoding of the deep structure of physics ratios such as density than a lower level of generation (worked examples). Deep encoding mediated transfer. A comparison with a previous study suggests that students react differently to the task of inventing or explaining worked examples, respectively, depending on their experience with generative group activities.

Mazziotti, Loibl, and Rummel could show that productive failure in mathematics (fraction expansion) is also beneficial for a new age group, namely, elementary school children. Additionally, they varied if students worked on the preparation problem in groups or individually. A simple comparison did not reveal differences in learning outcomes depending on the social setting. However, further analyses of video data will provide more differentiated information about effective processes during collaborative generative activities.

Hallinen, Chin, Blair, and Schwartz summarize several studies on the use of contrasting cases in combining generation and expository instruction, about task orientation, and about the content of materials. They could show that more complex material (including main effects and interaction of key variables instead of just the main effects of variables) enhances future problem solving. Consistent with the symposium as a whole, they conclude that the well-chosen combination of well-designed materials and an explicit focus on generation (such as an invention task) are two key aspects to transfer big ideas across school topics.

The symposium will give the audience a chance to generate and invent hypotheses about key ideas in the symposium by providing some materials from the studies. The interactivity of the symposium will further be facilitated by discussion questions posed by the presenters.

Understanding the Gradient of Linear Functions: Comparing Students' Transfer Performance Resulting from Different Preparatory Constructive Learning Activities and Tell-and-Practice Instruction

Lennart Schalk, Armin Barth, Ralph Schumacher

Introduction

A major challenge to science and math education is to enable students to use the knowledge learned in the classroom flexibly. An important aim is therefore to foster the construction of knowledge structures which enable transfer to new situations. In the present study, we compared different approaches of how to introduce a novel concept in math education which is of high importance for science in general, namely learning how to determine gradients of linear functions. One way to introduce a new topic is to tell (i.e., directly instruct) learners about the concept and have them practice its application in several tasks; another way is to withhold the expository instruction and start with a preparatory constructive learning activity. These activities have been shown to better support transfer in comparison to the tell-and-practice approach (e.g., Schwartz, Chase, Oppezzo, & Chin, 2011).

We implemented contrasting cases in learning materials to prepare an expository instruction on how to determine the gradient of linear functions. Contrasting cases help to recognize deep similarities between cases and to abstract generalizable schemata (e.g., Alfieri, Nokes, & Schunn, 2013) which in turn can enhance the effectiveness of an expository instruction (e.g., Schwartz et al., 2011). However, contrasting cases could be implemented in several different ways. First, cases can vary in concreteness. Cases could make a reference to concrete, realistic situations (e.g., using realistic concepts to label the axes of coordinate systems in which linear functions are depicted) or they could represent a concept idealized without direct reference to the real world (e.g. removing labels of the axes). Concrete cases have the advantage of activating real-world knowledge which might help to solve the task, but this activation of prior knowledge could also distract learners from deriving the underlying generic concept (e.g., Kaminski & Sloutsky, 2013). Therefore, the underlying concept might be easier to abstract from idealized cases, which however do not have the potential of activating real-world knowledge. Second, when several cases are presented simultaneously, participants should be prompted to actively compare the cases to increase transfer performance (e.g., Gick & Holyoak, 1983). Promising approaches are self-explanation prompts and prompts to invent a procedure or a formula to describe several different cases coherently. For the present research, we manipulated these two factors, concreteness of cases and kind of prompts, to construct four different preparatory learning materials and compared the transfer performance resulting from learning with these materials to a tell-and-practice condition.

Materials, Procedure, and Hypotheses

All learning materials consisted of printed booklets and were randomly distributed within standard math lessons. In the preparation conditions, students started with constructive learning activities before they received expository instruction. We always presented eight cases of linear functions (i.e., 3 graphs with a positive gradient, 3 graphs with a negative gradient, and 2 graphs with a gradient of 0). Cases were either concrete or idealized and combined with either self-explanation (EXPLAIN_CONCRETE, EXPLAIN_IDEALIZED) or invention prompts (INVENT_CONCRETE, INVENT_IDEALIZED). In the CONCRETE conditions, axes of the graphs were labeled with meaningful concepts; these labels were removed in the IDEALIZED conditions. In the EXPLAIN conditions, cases were presented with a gradient (i.e. a number) that indicated the steepness of the slopes. Participants were prompted to explain how these gradients had been derived. In the INVENT conditions, cases were presented without the gradient and participants were prompted to invent a coherent method to derive an index for the steepness of the slopes. After participants finished working on the preparatory materials (for max. 30 min), these materials were removed and participants read a short instructional text on the formal definition of gradients of linear function (max. 5 min). In the tell-and-practice condition (T&P), participants read the short instructional text first (max. 5 min). Afterwards, they practiced application of the concept by computing the gradients of eight cases of linear graphs (max. 30 min).

We assessed transfer performance immediately after learning and four weeks later. The test consisted of 20 questions embedded in contexts that differed strongly in superficial details from the learning materials (e.g., physics contexts, geometry problems). Participants had 45 minutes to answer the questions. We assumed that writing self-explanations is more productively combined with concrete cases given as the context might support writing explanations, while this reference might make it more difficult to invent a method because the labeling might serve as unnecessary seductive details. Accordingly, we expected a qualitative interaction (stable over time) for the preparation conditions: INVENT_IDEALIZED > INVENT_CONCRETE, and EXPLAIN_IDEALIZED < EXPLAIN_CONCRETE. Furthermore, we expected that at least the two preparation conditions INVENT_IDEALIZED and EXPLAIN_CONCRETE outperform T&P.

Participants

We tested 148 nine-grade students (73 female; mean age 15;6) from a Swiss town in the first week after the summer break. Linear functions are introduced in grade nine, thus, no participant had received any instruction on linear functions prior to the instructions delivered in the present study. Students were newly mixed in the ninth grade. We expected differences in prior math knowledge between students coming from different classrooms and therefore, we used the final math grades of grade eight as a covariate in all analyses.

Results

Transfer questions were scored with 1 point if solved correctly, with 0.5 points if only parts of the solution were correct and with 0 points if the solution was incorrect or missing. The statistical analyses followed a two-step process. First, we computed 2 (time of transfer test: immediate vs. delayed) x 2 (concreteness: concrete vs. idealized) x 2 (prompt: invent vs. self-explain) ANCOVA (covariate: math grade) for the preparation conditions to test our hypothesis of a qualitative interaction. Second, we compared performance in the preparation conditions to T&P in a 2 (time) x 5 (condition) ANCOVA. We only report statistically significant effects with $p < 0.05$. The covariate strongly predicted transfer in both ANCOVAs, but did not interact with conditions.

For the preparation conditions, the ANCOVA indicated a slight decrease in transfer performance over time ($F(1,115) = 4.2$, $\eta_{\text{partial}}^2 = .04$). More importantly, a statistically significant concreteness x prompt interaction ($F(1,115) = 4.0$, $\eta_{\text{partial}}^2 = .03$) supported our hypothesis of a qualitative interaction:

INVENT_IDEALIZED outperformed INVENT_CONCRETE (marginal estimates means: 12.67 vs. 11.40), while SELF_CONCRETE outperformed SELF_IDEALIZED (marginal estimates means: 12.35 vs. 11.05). Comparing transfer performance to T&P revealed no statistically significant differences between conditions.

Discussion
The present study in which participants learned to determine the gradient of a linear function shows that T&P is not inferior to learning materials in which expository instruction is preceded by a preparatory constructive learning activity. One explanation could be that the concept was quite easy, which might make a preparation unnecessary. For the preparation conditions, our results indicate that different kinds of prompts should be combined with different kinds of cases. Self-explanation prompts should be combined with concrete cases. A possible explanation is that this activity requires participants to describe materials and concepts in their own words; this requires the use of prior knowledge which would not be as easily available when cases are represented idealized. In contrast, invention prompts should be combined with idealized cases. We assume that this advantage emerges as the prompt is highly demanding when several cases have to be considered and idealized cases make it easier to grasp the underlying concepts. However, these interpretations would benefit from fine grained analyses of the quality of self-explanations and invented solutions. These analyses are conducted at the moment and will be presented at the conference. At the moment, the results indicate that different approaches can lead to successful transfer performance and that the T&P approach, which is popular among educators, need not to be considered a deficient instructional technique generally.

Two Times Inventing Beats Worked Examples as Preparation for Learning from Expository Instruction
Inga Glogger, Katharina Gaus, Alexander Renkl

Introduction and Theoretical Background
Inventing includes a generation task based on contrasting cases. More specifically, a formula or common principle is to be invented that allows for the evaluation of the cases (e.g., crowdedness of passengers in several buses). Such a generative inventing problem aims to prepare learners for subsequent expository instruction (e.g., about density; Schwartz, Chase, Oppezzo & Chin, 2011). This preparation can consist of cognitive effects, such as making aware of knowledge gaps, and motivational effects such as lifting curiosity and interest. Enhanced motivation can foster transfer. The generative processes can deepen encoding of the cases. Encoding cases (examples) deeply in an early stage of skill acquisition can foster transfer as well. On the other side, inventing activities may be problematic because learners are not directed towards the most important aspects and might (thus) not generate canonical solutions. Additionally, previous findings on inventing can often be explained by longer time-on task or weak control groups (Sweller, Kirschner & Clark, 2007). Having learners study a worked version of the inventing problem with contrasting cases (worked example) as control condition avoids these problems. In a recent study, we have actually found that enhanced self-efficacy (motivational), and less extraneous load (cognitive) in the worked example group, as compared to a inventing groups, fostered deep-structure encoding and the transfer of big ideas (Glogger, Fleischer, Grüny, & Renkl, 2013). Students, however, might have had too little experience and practice with generative tasks such as inventing (Lee & Anderson 2013). Thus, in the present experiment, we applied two inventing phases instead of one and asked if the previously found advantage of worked examples can be replicated. More specifically, we tested the following hypotheses and asked the following research questions:

1. Inventing leads to higher interest, curiosity, and perceived knowledge gaps than the worked solution. 2. The worked solution induces less extraneous load and leads to higher self-efficacy than inventing. 3. Do the two groups differ in encoding of the deep structure of preparation problems? 3. Do the two groups differ in learning outcomes (transfer)? 4. Are there process variables that mediate the effects on learning outcomes?

Method
We randomly assigned 108 eighth-grade students (sex: 89 female; $M_{\text{age}} = 13.62$, $SD = 0.65$) to two conditions: inventing ($n = 52$) and worked solution ($n = 56$). The experiment had two preparation phases, both preparing to learn the concept of density, and moreover, of ratio indices as "big idea" in physics (the materials were adapted from Schwartz et al., 2011). The phases varied in the level of generation (inventing the solution or worked solution). In a first preparation phase, we gave three cases of different numbers of clowns that were squeezed in differently sized buses along with the task to find an index (with certain properties) for the crowdedness of the buses (the second preparation task was about gold quality). Students can discover the ratio structure of the index by comparing cases. Both experimental groups worked in pairs for 20 minutes. The inventing group invented the index. The worked example group worked on the same problem, but it was worked out: a fictional student already solved the task. Some of her thoughts were written on the worksheet. Participants were asked to explain

their solution to each other. This way, we kept communicating and debating explanations comparable; the groups differed in whether the solution was generated or worked.

We assessed extraneous load, self-efficacy, interest, epistemic curiosity, and perceived knowledge gaps by questionnaires immediately after the preparation task (see Glogger et al., 2013). Two days later, a recall test was given to measure the encoding of the deep structure during the preparation phase 1. The subsequent preparation phase 2 was about an index for gold quality, also preparing for understanding the concept of density. Students then listened to a lecture. Finally, we determined learning outcomes by calculation tasks and transfer tasks (e.g., invent density index, new cover story, or invent spring constant index, i.e., another ratio index in physics).

Findings

Groups did not differ in interest, curiosity, self-efficacy, and knowledge gaps (all $p > .23$; see Table 1 for means and standard deviations). Working on a worked-out version led to less extraneous load, $F(1,97) = 3.47$, $p = .033$ (one-tailed). Further analyses showed that this effect is mainly due to students who perceive their performance in sciences as low (significant interaction $b = -0.181$, $t(95) = -2.22$, $p = .029$). Encoding of the deep structure of the preparation problem, measured by the recall test, was higher in the inventing group, $F(1,97) = 6.22$, $p = .008$, $d = 0.54$. The inventing group outperformed the worked solution group in transfer tasks, $F(1,97) = 11.20$, $p = .001$, $d = 0.67$. We tested mediation effects with a set of related multiple regression equations, following a products-of-coefficients strategy (MacKinnon, 2008). Encoding of deep structure mediated part of the group effect on far transfer; path A: $b = .50$, $SE = .186$, $\beta = .26$, $p = .008$, $R^2 = .28$; path B: $b = 1.14$, $SE = .427$, $\beta = .26$, $p = .009$, $R^2 = .40$; Sobel-test: $z = 2.30$, $p = .021$, $d = 0.14$, small effect. Deeper exploratory analyses revealed that more students found the canonical solution to the second inventing problem (56 %) than to the first (34 %), whereas self-explanations to the second worked example tended to get worse. We did not find any gender effects.

Table 1: Means (Standard Deviations) of process and outcome variables in both groups and ranges of variables.

	Inventing *M* (*SD*) (*N* = 49)	Worked-Example *M* (*SD*) (*N* = 50)	Range
Perceived Knowledge Gaps	3.72 (1.08)	3.91 (0.85)	1-6
Extraneous Load	**3.93** (0.88)	**3.62** (0.76)	1-6
Interest	3.33 (0.94)	3.41 (0.85)	1-6
Epistemic Curiosity	2.91 (0.89)	3.11 (0.73)	1-6
Self-efficacy	3.60 (2.20)	3.87 (2.32)	0-10
Recall of Deep Structure	**1.06** (1.07)	**0.56** (0.76)	0-3
Transfer	**8.10** (4.23)	**5.40** (3.79)	1-16

Note. Significantly different means are printed in bold.

Discussion

In this study with two preparation phases, a higher level of generation (inventing) led to deeper encoding of the deep structure of the preparation problems. Deep encoding partly mediated transfer performance. These findings are in contrast to own previous studies (with one preparation phase), but in line with Schwartz' studies (with several preparation phases). One explanation of the effects could thus be that several preparation phases are required for inventing to be successful in comparison with a worked example. Inventing solutions got actually better in the second phase. Other studies have been shown that generation tasks need practice to be successful (Lee & Anderson, 2013).

However, we have found differences in encoding the deep structure already after the first preparation phase. A reason could be that students' practice with generative activities was already higher than in a previous sample (Glogger et al., 2013). If students need practice in generative tasks, students with higher levels of experience with generative activities, as compared to little experience, should react and benefit differently from them. We have informal hints that the present sample differed in this experience from the previous sample (teachers' reports). We compared the patterns of results of the previous study with less experienced students with the present one with more experienced students. The process variables were measured by the same items in both studies. This comparison suggests that the preparation activities elicited different processes across studies. The worked example led to significantly higher self-efficacy and less extraneous load in the previous study. Both latter variables were significant mediators of the group effect on transfer. Motivation was higher in the present sample. These different reactions towards the preparation activities could be an effect of different levels of experience with generative group activities in the two samples. However, further research is needed to systematically test this assumption.

Worked examples were shown to be most effective with a fading procedure (Renkl, 2013). In the present study, the second fully worked-out example might have reduced interest, curiosity, and deeper processing because there was no fading. Subsequent studies could use a fading procedure and further look at beneficial motivational and cognitive processes during a preparation activity. In sum, this experiment suggests that students' transfer of big ideas in physics is best fostered by several phases of a generative task such as inventing instead of several worked examples.

Does Collaboration Affect Learning in a Productive Failure Setting?

Claudia Mazziotti, Katharina Loibl, Nikol Rummel

Introduction and Theoretical Background

Researchers as well as practitioners struggle with the so called assistance dilemma, which targets the question of the timing and degree of assistance that students should receive when learning new concepts and procedures. Some researchers argue that a high degree of assistance in form of instruction should be provided at the beginning of the learning process to avoid a high level of cognitive load (Kirschner, Sweller, & Clark, 2006). Others claim that at the beginning of the learning process students should be enabled to discover new concepts and procedures on their own (i.e., low degree of assistance) to support the acquisition of conceptual knowledge (e.g., Schwartz & Martin, 2004; Kapur, 2012). One representative of this group is Manu Kapur with his two-phase instructional design of Productive Failure (PF) (e.g., Kapur, 2012): In the first phase, small groups of students try to solve a new problem on their own. During this problem-solving phase they usually generate incomplete or erroneous solution ideas. In the second phase, students receive instruction on the canonical solution in a class session (instruction phase). It has been discussed that PF designs trigger the following learning mechanisms (e.g., Kapur, 2012): First, students' problem-solving attempts prompt the activation of prior knowledge. Second, during instruction students focus their attention on the components of the target concepts that they did not yet discover during the problem-solving phase. Additionally, Loibl and Rummel (2013) argue that the form of instruction is crucial for learning: During the instruction phase, the teacher can build upon students' generated solution ideas by comparing these ideas to each other and contrasting them with the canonical solution. Loibl and Rummel (2013) showed that this form of instruction was more effective than instruction focusing on the canonical problem-solving procedure and the underlying concept.

While the effects of the timing and the form of instructions on the acquisition of conceptual knowledge are studied well, the role of collaboration during the first phase of PF remains unclear (Collins, 2012). Against the background of research on collaborative learning, we assume positive effects of trying to solve a problem collaboratively, because in a collaborative learning setting students are encouraged to verbalize and explain their problem-solving ideas to each other (Slavin, 1996). In this way, elaborative processes and sense-making activities are initiated, which in turn support the acquisition of conceptual knowledge (Cohen, 1994). Indeed, Mullins, Rummel, and Spada (2011) showed that trying to solve a problem collaboratively is more efficient for the acquisition of conceptual knowledge than individual problem solving. However, research on collaborative learning has also emphasized that students need collaboration support (e.g., a role script, a group goal) to ensure fruitful interaction (Slavin, 1996; Cohen, 1994). Against this background, we hypothesize that learning collaboratively during the first phase of PF facilitates the acquisition of conceptual knowledge in comparison to individual learning (hypothesis 1). Given that several studies with secondary school students or university students could show that PF designs foster conceptual knowledge and transfer in comparison to expository instruction conditions (e.g., Schwartz & Martin, 2004; Kapur, 2012), we further investigated whether the beneficial effect of PF on the acquisition of conceptual knowledge can be replicated with elementary school students (hypothesis 2).

Method

In order to test our hypotheses, we conducted a quasi-experimental study with 4th graders (N = 55; age 10) in Germany. Overall, we installed three conditions: We implemented two PF conditions as they were described above differing only in the social form during the problem-solving phase (i.e., PFCo = problem solving in dyads; PFIn = individual problem solving). As a control condition, we set up a third condition (EICo), in which students received instruction first. As in the PF conditions, the experimenter compared and contrasted typical student solutions, which were collected prior to the experiment during pilot studies. Afterwards students solved an isomorphic problem collaboratively. In both collaborative conditions (PFCo & EICo) we supported the collaboration via a role script ("thinker" and "asker") and a group goal (the pair who collaborated best won a price). During the collaboration students were videotaped. The to-be-solved problem was a story problem requiring expanding two fractions to compare them. Prior to the first learning phase, all students answered a pretest measuring students' mathematical prerequisites. After the second phase, students answered a posttest measuring conceptual knowledge of the target concept. The study took place on two days (45 minutes per day) and the three classes were assigned to the three conditions as a whole. In all classes, students have not learnt fraction expansion and comparison prior to the study.

Results

Only the 52 students who were present during both learning phases were included in the analyses. To assess differences between the experimental conditions, we calculated an ANCOVA with the factor condition and the covariate pretest score (i.e., mathematical prerequisites). Mean scores and standard deviation of the three conditions are displayed in Table 1.We did not find significant differences between the three conditions ($F[3,48] = 2.3$, $p = .11$). We further defined two a priori contrasts in line with our hypotheses. The first a priori contrast compared PFCo and PFIn to test the effect of collaboration (hypothesis 1). For this contrast, we did not find significant differences between conditions ($F[1,48] = 0.4$, $p = .84$). The second a priori contrast compared both PF conditions to EICo to test if the PF effect can be replicated for young children (hypothesis 2). This comparison revealed a significant difference between EICo and the two PF conditions ($F[1,48] = 4.6$, $p = .03$), favoring the PF conditions.

Table 1: Mean scores and standard deviations of the posttest.

	N	*Mean*	*SD*
PFCo	16	7.88	5.43
PFIn	17	7.71	3.29
EICo	19	5.26	3.91

Discussion and Outlook

In summary, we investigated two hypotheses: Our main hypothesis, that collaborative problem solving within the first phase of PF is more beneficial for learning than individual problem solving, could not be confirmed by our findings. This result is surprising considering the above mentioned beneficial effects of collaborative learning on the acquisition of conceptual knowledge. Although we implemented a group goal and a role script to support the collaboration, it seems that students did not take advantage of the collaborative setting. The analysis of the video data will provide insights on the quality of students' collaboration. Given the high SD in the PFCo condition we assume that the quality of collaboration differed strongly across dyads and moderated the effect of collaboration on conceptual knowledge. Hence, for future work we aim to intensify our efforts to implement a collaborative support structure (e.g., by implementing a training for collaborative learning). Our additional hypothesis, that also for young students PF is more beneficial for the acquisition of conceptual knowledge than expository instruction, could be confirmed. In other words, this finding replicated other PF studies (e.g., Kapur, 2012; Loibl & Rummel, 2013) with a younger age group.

Due to the small sample size we should treat our results as preliminary. We plan to conduct a follow-up study with a 2x2 design with the factors timing of instructions and social form. With this study we aim at determining whether the possible effect of collaborative learning interacts with the timing of instructions.

Using Contrasting Cases for Generation and Instruction

Nicole R. Hallinen, Doris B. Chin, Kristen P. Blair, Daniel L. Schwartz

Contrasting Cases

An important part of science and math education is the learning of functional relations among variables. These ideas map to Mathematical Practices in the Common Core State Standards (e.g. *Reason abstractly and quantitatively; Look for and make use of structure*), as well as the Ratio & Proportion Content area throughout middle school grades (National Governors Association Center for Best Practices, 2010). In science assessments, it has been shown that 4th grade students who know about functional relations between plant growth and amount of light are better able to explain their answers, indicating that this knowledge affects a students' ability to express his or her understanding (NCES, 2012).

Contrasting cases are instructional tools that, in our instantiation, present a series of examples of a phenomenon; they are particularly well-suited to the task of highlighting variation to show functional relations. During these short, inductive activities, students learn to appreciate the range of variation in a problem space and also notice the invariant relationship that underlies each example. Contrasting cases have been developed for a number of STEM topics, such as statistical variation (Schwartz & Martin, 2004), electromagnetic fields (Chase, Shemwell, & Schwartz, 2010), and density and other ratio-based topics (Schwartz, Chase, Oppezzo, & Chin, 2011), and for use with audiences ranging from 4th grade to graduate students.

Our research focuses on two aspects of using contrasting cases in math and science instruction: task orientation and design of materials. In both topics, we measure the effectiveness of using contrasting to learn about scientific relationships and eventually transfer these ideas to new topics.

Task Orientation

To use contrasting cases in instruction, educators must instill the proper task orientation in students. We find that encouraging students to adopt an "inventing" task orientation is productive for discovering important features and patterns among the examples shown.

In a study conducted in an 8th grade classroom, contrasting cases were used to teach students about the underlying concept of ratio, an important idea that relates to many ideas in middle school science and mathematics topics, such as density, and speed. Some students were taught about ratios first and told to use this knowledge to compute ratios about the contrasting cases. Another condition was asked to invent an index to describe the cases. In the inventing condition, students invented the ratio structure of density and speed. The students who invented the ratio idea transferred the concept to the spring constant, a new domain, on the posttest (Schwartz et al., 2011). The students who were explicitly taught about computing density and speed did not see the underlying ratio structure across examples, which led to lower transfer. This finding is in line with other research showing that students fail to notice commonalities across problems when instances are presented sequentially (Star & Rittle-Johnson, 2009).

A different study with 40 6th grade students used a set of contrasting cases to teach the relationships among range, hang time, and speed in projectile motion. In this between-subjects study, one condition was asked Compare and Contrast (CC) the cases. The other condition was instructed to use the same set of cases to Invent (INV) a method to predict the range of any shot. The CC students focused on isolating individual components of projectile motion for comparison and were less likely to create an explanation that linked all three factors (range, time, and speed) than the INV condition. As a result, INV students outperformed their CC peers on a posttest about predicting the range of a projectile. Compare and Contrast instructions were simply not sufficient for learning structural relationships from contrasting cases; activities must ask students to generalize. (Chi et al., 2012)

These examples highlight the benefit of encouraging students to search for a general solution. In each study, the same contrasting cases materials were provided to each condition, indicating that task orientation provides additional support for learning beyond the cases themselves.

Designing Contrasting Cases Materials

Even with the right task orientation, students need well-designed materials to support learning. The examples included in sets of contrasting cases for multivariate topics must show varying levels of each factor as well as the variety of possible combinations of variables that contribute to the solution. Fifty community college students participated in a recent study to investigate the effects of using different materials to learn early physics concepts.

For this study, we designed two sets of contrasting cases showing two-dimensional inelastic collisions. The result of each collision (move left, right, or stop) is determined by comparing the momentum (mass * speed) of the objects. We manipulated the examples as shown in Figure 1A. Individuals in the *Main Effects* (ME) condition received cases that isolate the main effects relationships between the two variables – cases where mass or speed is held constant across the two objects, or where mass and speed both indicate the same result. In the *Main Effects + Interaction* (ME+I) condition, some cases showing interaction relationships between the two key variables were included. In these cases, attending only to mass or only to speed would be insufficient to determine the result.

In a transfer task, participants in both conditions received contrasting cases for the balance scale, an analogous science domain in which torque must be calculated by multiplying mass and distance from the fulcrum to determine whether the scale balances or tips. These cases included main effects and interaction cases, showing enough variation that simple qualitative rules about mass and distance would not be sufficient to explain the results. These cases are shown in Figure 1B, with the interaction cases outlined for the reader. The worksheet that students saw did not include this outlining.

We found that using the ME collision materials led to a qualitative understanding of the relationship between mass and speed. Participants who receive the ME+I materials were more likely to find the multiplicative relationship between the variables, and, in turn, performed better on a set of 18 prediction questions ($t(32) = -3.14$, $p < 0.01$).

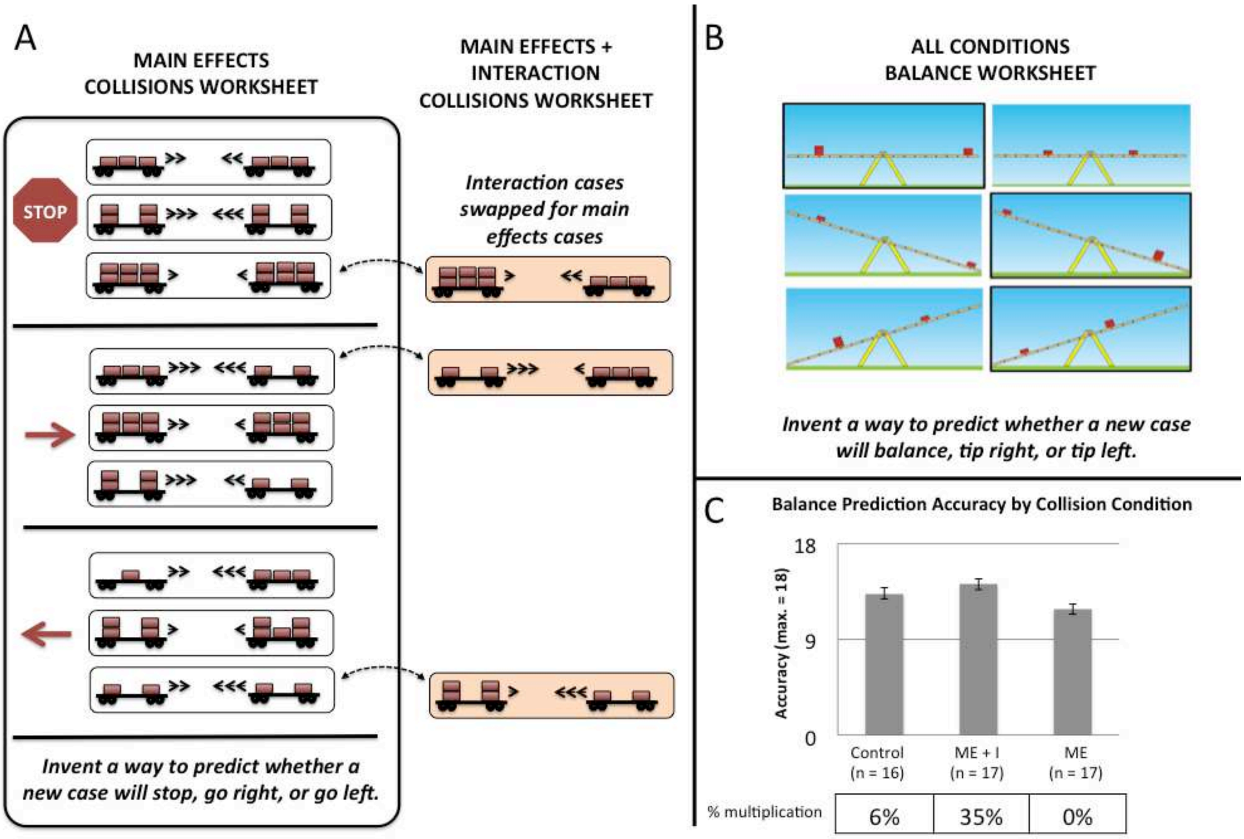

Figure 1. Contrasting cases materials for A) momentum and B) torque. C) Transfer task results.

As shown in Figure 1C, participants who used the momentum ME materials learned less from the transfer balance scale cases (0% found a multiplicative solution). They also performed worse on balance scale prediction questions than their ME+I counterparts. In fact, they performed even lower than control students who did not do the momentum activity at all. Using ME materials resulted in negative transfer to a new content area.

This result demonstrates that the effects of using incomplete examples extend beyond one topic and can impact learning on future problem solving tasks. Therefore, educators should carefully choose contrasting cases that systematically demonstrate the structural relationships between variables in these multivariate conceptual domains.

Conclusion

Contrasting cases have the potential to facilitate a deep learning and eventual transfer of quantitative structures, such as ratio, that recur throughout science and math curricula various science and math topics. As demonstrated, instructional designers must thoughtfully select cases to show the range of possible variability. Additionally, the studies reviewed here emphasize the utility of adopting an "inventing" task orientation for learning from contrasting cases. We suggest that the combination of well-chosen materials and an explicit focus on invention and generalization are two key aspects of using contrasting cases to promote learning across STEM domains.

Discussant

Katherine McEldoon, Learning Sciences Institute, Arizona State University
(Chairperson: Inga Glogger, University of Freiburg)

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, 48(2), 87-113.

Chase, C. C., Shemwell, J. T., & Schwartz, D. L. (2010). *Explaining across contrasting cases for deep understanding in science: An example using interactive simulations.* Proceedings of the International Conference of the Learning Sciences.

Chi, M., Chin, D., Hallinen, N., & Schwartz, D. L. (2012, August). *A comparison of two instructional models using contrasting cases*. Physics Education Research Conference, Philadelphia, PA.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*(1), 1-35.

Collins, A. (2012). What is the most effective way to teach problem solving? A commentary on productive failure as a method of teaching. *Instructional Science*, *40*, 731-735.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*,1-38.

Glogger, I., Fleischer, C., Grüny, L., & Renkl, A. (2013, August). *Inventing compared to worked-out problems: which processes foster transfer in physics?* Paper presented at the 14th Biennial Conference of the European Association for Learning and Instruction (EARLI), Munich, Germany.

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology*, *105*(2), 351-363.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*, 651–672. doi:10.1007/s11251-012-9209-6.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75-86.

Lee, H. S., & Anderson, J. R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology, 64*, 445–469. doi:10.1146/annurev-psych-113011-143833

Loibl, K., & Rummel, N. (2013, June). *Delaying instruction alone doesn't work: Comparing and contrasting student solutions is necessary for learning from problem-solving prior to instruction*. Paper presented at the 10th international conference on computer-supported collaborative learning.

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

Mullins, D., Rummel, N., & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *International Journal of Computer Supported Collaborative Learning*, *6*(3), 421-443.

National Center for Education Statistics (2012). *The Nation's Report Card: Science in Action: Hands-On and Interactive Computer Tasks From the 2009 Science Assessment (NCES 2012-468)*. Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). *Common Core State Standards (Mathematics)*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers.

Renkl, A. (2013). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, *37*, 1–37. doi:10.1111/cogs.12086

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. doi:10.1037/a0025140.

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.

Slavin, R. E. (1996). Research on Cooperative Learning and Achievement: What we know, what we need to know. *Contemporary Educational Psychology*, *21*(1), 43-69.

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102*(4), 408-426.

Sweller, J., Kirschner, P., & Clark, R. (2007). Why minimally guided teaching techniques do not work: a reply to commentaries. *Educational Psychologist*, *42*, 115-121. doi:10.1080/00461520701263426