Learning and Instruction 41 (2016) 41–51

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Consistent advantages of contrasted comparisons: Algebra learning under direct instruction

Esther Ziegler*, Elsbeth Stern

Center for Research on Learning and Instruction, ETH Zurich, Switzerland

A R T I C L E I N F O

Article history:
Received 23 August 2014
Received in revised form
31 August 2015
Accepted 10 September 2015
Available online xxx

Keywords:
Comparison
Contrasting
Algebra learning
Concept learning
Direct instruction

A B S T R A C T

Contrasted comparisons are an effective means of concept learning. Many studies have investigated the effects of comparisons in self-learning settings. However, because direct instruction remains a widespread instructional method, we adapted self-learning materials from a previous experiment that demonstrated the benefits of a contrasted introduction of algebraic addition and multiplication for direct instruction at the blackboard. Ninety-eight sixth-graders were randomly assigned to two groups: in the contrast group, addition and multiplication were simultaneously introduced and compared at the blackboard, whereas in the sequential group, addition was taught for two days followed by two days of multiplication training. The contrast learners clearly outperformed the sequential learners in differentiating superficially similar algebraic concepts on three follow-up measures. Furthermore, similar to our previous results, the benefits of contrasted comparison only appeared with delay. Our results extend the applicability of contrasted comparison to the direct instruction method.

1. Introduction

Comparison is an effective principle in concept learning across a variety of applications and contexts in both laboratory and classroom settings (for an overview, see Alfieri, Nokes, & Schunn, 2013; Gentner, 2010; Rittle-Johnson & Star, 2011). Alfieri et al. (2013, p. 108) summarized comparison research and demonstrated that "most studies in the current review utilized case comparisons to illustrate only a single procedure or concept". This approach of using two or more cases to learn a single concept, procedure or method, as opposed to studying the cases separately, can serve as a reference for complex concept learning, as demonstrated in the following studies: the use of two different cover stories for the same negotiation strategy (Loewenstein, Thompson, & Gentner, 2003), the study of two different scenarios that depict the same concept of heat flow (Kurtz, Miao, & Gentner, 2001), processing two phenomena related to the concept of changing the state of matter (Mason, 2004), and learning with three cases that illustrate the concept of electromagnetic flux (Chase, Shemwell, & Schwartz, 2010). When two cases are juxtaposed, the comparison helps emphasize the essential features of the underlying concept beyond the distracting surface features. Thus, a comparison of examples helps learners abstract the common principle, which can subsequently be used to solve novel problems. The results of the studies indicated several important cognitive skills that can be improved by comparison, including categorization, concept development, concept application, and transfer.

Rittle-Johnson and Star (e.g., Rittle-Johnson & Star, 2007; Star & Rittle-Johnson, 2009) further elaborated and presented two different solution methods for learning a complex and challenging mathematical procedure, methods that helped students flexibly choose between appropriate solution strategies. The results demonstrated that the students became more flexible and improved their understanding of the procedure. Comparisons of similar solution methods emphasize differences in the methods, which helps to distinguish between them. This fact has been thoroughly described in analogical reasoning research (Gentner, 2010; Loewenstein & Gentner, 2001; Markman & Gentner, 1993).

In general, the acquisition of demanding conceptual knowledge requires deliberate instructional support. Students often have considerable difficulties in distinguishing between concepts because they are usually introduced consecutively. To help the learner distinguish between different concepts and strategies, we successfully introduced different concepts simultaneously by contrasting them (Ziegler & Stern, 2014). In this previous experiment,

* Corresponding author. Research on Learning and Instruction, Institute for Behavioral Sciences, ETH Zurich, Clausiusstrasse 59, 8092, Zurich, Switzerland.
E-mail address: esther.ziegler@ifv.gess.ethz.ch (E. Ziegler).

http://dx.doi.org/10.1016/j.learninstruc.2015.09.006
0959-4752/

we chose the similar but easily confused principles of addition and multiplication in algebra and developed an extensive four-day, self-study program. In a group by time design, sixth graders were randomly assigned to either the contrast or the sequential condition. In the contrast condition, addition and multiplication problems were mixed and juxtaposed from the beginning, whereas in the sequential condition, addition problems were presented first and multiplication problems were presented second. We demonstrated the clear benefits of a simultaneous and contrasted introduction of algebraic addition and multiplication in three follow-up sessions both in tests on algebraic transformations and in measures on an explicit understanding of algebraic principles. It was shown that the students in the contrast condition were better able not only to distinguish between algebraic addition and multiplication but also to correctly apply them in the long run, although they performed worse on immediate learning measures. This short-term disadvantage and long-term benefit of contrasting sends an impressive message to those in learning settings. Two theoretical explanations are worth considering (see detailed discussion in, Ziegler & Stern, 2014). First, this pattern of delayed benefits, occasionally called desirable difficulties, matches the higher effort that contrast learners must invest, which corresponds to the retention effects observed for intermixed concept learning (E. Bjork & Bjork, 2011; R. Bjork, 1994; Kapur & Bielaczyc, 2012; Rohrer & Pashler, 2010). Our results showed contrasting as a method that makes material more challenging but at the same time more efficient to learn. Second, although high extraneous cognitive load is not recommended, complexity of material may result in high but inherent intrinsic cognitive load (Sweller, 2003; Sweller & Chandler, 1994; Sweller, Van Merrienboer, & Paas, 1998). Our study's addition and multiplication principles were not complex per se, but distinguishing between them was a challenge. Their juxtaposition increased the load but helped students to focus on key differences and to interconnect the two similar, easily confusable principles. Therefore, the higher intrinsic cognitive load caused by contrasted problem presentation presumably prevents learners from engaging in superficial information processing and rash automation.

Other studies have examined the comparison of principles from two different categories (Cummins, 1992; Day & Goldstone, 2010; Kang & Pashler, 2012; VanderStoep & Seifert, 1993). However, the materials used in these studies were categorization and selection tasks, which contrasted with our more complex mathematical problems. Our goal was to extend the application of comparisons beyond learning single complex concepts or procedures to simultaneously learning two different, but similar principles. In this experiment, we seek to determine whether the short-term failure and long-term superiority of the contrasted comparison can be replicated under modified instructional settings.

1.1. The instructional setting of the comparison studies

Most studies of case comparisons have used self-learning settings, which have frequently included analyzing and describing scenarios (e.g., Kurtz et al., 2001; Mason, 2004), reading and rating cases (e.g., Gentner, Loewenstein, Thompson, & Forbus, 2009; Loewenstein et al., 2003), reading stories and writing summaries (e.g., Catrambone & Holyoak, 1989; Schwartz & Bransford, 1998), and studying sets of worked examples (e.g., Gerjets, Scheiter, & Schuh, 2008; Rittle-Johnson & Star, 2007). In comparison studies with preschoolers, self-learning settings were not possible because of the absence of sufficient reading and writing abilities. Instead, the children were individually tested by an experimenter, e.g., in categorization tasks (Graham, Namy, Gentner, & Meagher, 2010), spatial mapping tasks (Christie & Gentner, 2010), or hide-and-find games (Loewenstein & Gentner, 2001).

In their review, Alfieri et al. (2013) described only two studies that investigated comparisons in the direct instruction format. One study analyzed how frequently teachers used analogous comparisons in instruction and demonstrated that comparisons used in direct instruction must be accompanied by cues provided by the teacher to be visible and therefore helpful to students (Richland, Zur, & Holyoak, 2007). A second study investigated how students learn to distinguish between three language learning theories, demonstrating that a video-based anchored instruction group outperformed a traditional, lecture-based instruction group (Michael, Klee, Bransford, & Warren, 1993). However, neither study used explicit case comparisons similar to those of either our study or many other self-learning studies.

To summarize, the benefits of comparisons have been confirmed for different age groups (children, university students, and adults) for various group sizes and under various presentation forms (from paper-pencil to computer-based). Most studies have been conducted in controlled laboratory settings over brief periods, although Rittle-Johnson and Star (for an overview, see Rittle-Johnson & Star, 2011) designed various materials to implement comparisons in classroom mathematics. Overall, the comparison results were relatively robust. Nevertheless, for comparisons to be a natural element of daily classroom practice for teachers, further research is needed to identify the conditions under which they can be implemented as an effective means of instruction. Whether the superiority of contrasted comparison that has been so convincingly demonstrated in various self-learning settings also applies to direct instruction remains an open question.

1.2. The role of guidance and prompting in learning from comparisons

Systematic guidance and coherent structure have repeatedly been shown to be crucial for any kind of self-learning (Hardy, Jonen, Moller, & Stern, 2006; Mayer, 2004). Similarly, Alfieri et al. (2013) noted that “guidance” and “prompting” played important roles in all comparison settings. By contrast, the benefits of comparisons will be minimal when the learner merely reads or solves contrast problems (Loewenstein, Thompson, & Gentner, 1999; Ziegler & Stern, submitted). Precise and explicit instructions are essential for effective comparisons, e.g., describing similarities and differences in stories (Catrambone & Holyoak, 1989; Kurtz et al., 2001; Loewenstein et al., 1999; Mason, 2004), analyzing contrasting cases and finding relevant patterns (Schwartz & Bransford, 1998), or answering specific comparison questions (Gadgil, Nokes, & Chi, 2012; Ziegler & Stern, submitted).

Recent research has demonstrated the benefits of mixed learning of different rules, tasks or formulas instead of learning them one after the other in succession (Kang & Pashler, 2012). Interestingly, there was no difference between settings in which material was provided mixed simultaneous, e.g., abc, abc, abc, and settings in which material was presented mixed but in succession, e.g., abcabcabc. Mixed simultaneous processing provides a better opportunity to compare materials than successive processing, an opportunity, however, that was apparently not spontaneously used by learners. This result suggests that simultaneous processing is not sufficient and that for comparison to be an effective means of learning, learners must be directed to actively engage in systematic comparisons. Precisely such prompting can make a difference, as we have shown in an experiment on contrasted algebra learning with and without prompting comparisons (Ziegler & Stern, submitted). The group that received explicitly prompted comparisons outperformed the group presented with a simple opportunity to compare. In line with Sweller and Chandler (1994) and Sweller et al. (1998), we argued that prompting might force students to connect two similar principles and become aware of their distinguishing features and rules, thereby gaining a better overall comprehension.

These crucial elements of guidance and explicit prompting raise the question of whether direct instruction is an efficient method of supporting learning by comparison. Direct instruction is a widespread and common method of instruction practiced in regular schools all over the world. Although direct instruction is sometimes discredited by teachers, policy makers, and educational administrators, empirical evidence has shown that carefully prepared and thoughtfully applied direct instruction can be an efficient means of communicating information (Hattie, 2009; Kirschner, Sweller, & Clark, 2006). Direct instruction is characterized by explicit explanations and demonstrations of concepts and principles offered by an expert – usually the teacher. In direct instruction, the teacher directs the learners' attention to crucial aspects of the material.

Both comparison and direct instruction have the independent potential to direct learners' attention to important aspects and thereby to highlight the inherent key features and difficulties of the concepts to be learned. Regarding the particular potential of direct instruction, the question arises of whether teacher guidance in contrast learning that explicitly emphasizes differences between juxtaposed concepts reduces cognitive load to a similar level as that experienced in sequential learning. Through this reduction, the desirable difficulty and interconnecting effects of contrasted comparison might be neutralized. For the current experiment, this could mean that if the teacher strenuously emphasizes that addition and multiplication algebra problems must be treated differently, it may not matter whether direct instruction is based on contrasted or sequential material. Put differently, contrasted comparison may merely compensate for the deficiencies of self-learning, thereby rendering itself superfluous under direct instruction.

Despite these objections, there are nonetheless good reasons to assume that contrasting will also unfold its particular potential under direct instruction. When addition and multiplication are directly instructed in sequential order, key features and difficulties of both principles can be highlighted. However, the sequential practice of addition and multiplication problems may prevent learners from developing deeper insights into the differences between the two operations.

We therefore consider it necessary to examine the effects of our successful contrasted introduction of two algebraic principles in a direct instruction format. When new instructional materials, techniques or tools are developed, they are often investigated using self-learning material or computer-based methods. Because findings may easily be too broadly generalized, it is important to examine promising methods in alternative formats.

1.3. The current project

To examine the contrasting of cases under direct instruction, we adapted our algebra self-learning setting (Ziegler & Stern, 2014) to a direct instruction setting. As with the carefully designed self-learning program, we aimed to present well-thought-out and well-established direct instructions with guided blackboard explanations and structured training phases (e.g., Liem & Martin, 2013; Rosenshine & Stevens, 1986).

We address two questions: First, will contrasting of cases under direct instruction lead to better performance in solving algebra problems than sequential learning? For the reasons stated above, contrasting effects cannot be taken for granted under direct instruction. Nonetheless, they are highly probable, given the strength of contrasting effects found in previous studies, including our own. Being confronted with contrasted cases during both direct instruction and subsequent practice can be expected to highlight the inherent complexity of the concepts by directing the students' attention to the relevant aspects. Referring to our study, juxtaposing addition and multiplication problems helps students to interconnect both algebraic principles. If both types of problems are processed sequentially, cognitive load might be lower at the expense of deeper comprehension. Thus, we predict an advantage of contrasting on all the follow-up measures. The following two hypotheses were examined: 1) Contrasted comparison is superior one day after the intervention both on algebraic transformation knowledge and on measures of explicit understanding of algebraic principles. 2) The effects persist over the long term, at both one- and ten-week follow-ups.

Second, will the disadvantage of the contrast group with respect to immediate learning measures also occur under direct instruction? In the prior self-learning study, contrast students made more errors in the short term during intervention, but showed better long-term gains. This pattern of delayed benefits depicts the desirable effort that students had to invest during the acquisition phase. We argue that a teacher's explanations reduce the extraneous cognitive load and therefore will support contrast students in concentrating on the differences between addition and multiplication from the beginning of being presented with learning tests. It is therefore plausible that under direct instruction, learners from the contrast group can save themselves a detour via productive failures. Thus, it is worthwhile to examine whether students' learning trajectories are equivalent under both conditions. Accordingly, the third hypothesis was examined: 3) Both groups are equivalent in terms of immediate learning measures.

2. Methods

2.1. Participants

Sixth graders with no prior algebra knowledge were chosen as participants. In the Swiss mathematics curriculum, algebra is not introduced until secondary school, which begins in grade 7. Although sixth graders are unlikely to have received any formal instruction in algebra, based on their arithmetic competencies they can be expected to be ready to learn elementary algebraic principles. The participating classes were recruited from urban and suburban public schools in the canton of Zurich. For the analyses, we excluded students (a) with insufficient German language comprehension, (b) with special needs, or (c) who were unable to fulfill the minimum standards of school performance. All of the students were volunteers, and their parents provided written consent. Every class was rewarded with 200 Swiss francs (approximately 160 Euros), and each student received a small gift.

A total of 98 students (47 females) from five classes participated. Seven students were excluded: four students because they did not fulfill our criteria and three because they missed training days and did not finish the program. Students in each class were matched to the two groups based on their grades in mathematics and German. The contrast group was composed of 46 students ($M = 12.2$ years, $SD = .4$), and the sequential group was composed of 45 students ($M = 12.3$ years, $SD = .5$).

2.2. Design and procedure

In a 2 (group: contrast, sequential) x 3 (time: one day, one week, ten weeks) mixed-factorial design, we investigated effects on algebraic transformation and explicit transformation explanations. Each student participated in four training sessions and three follow-up sessions, as listed in Table 1. Both groups participated in

Table 1
Overview of the activities of each session.

	Session	Duration	Activities
Training sessions	1st day (Mon.)	2 lessons[a]	Pretest – prior algebra knowledge (5 min) Introduction – short slide presentation: (5 min) Session 1: units[b] and learning tests[d] 1 – 3
	2nd day (Tue.)	2 lessons[a]	Repetition test 1 (5 min) Session 2: units[c] and learning tests 4 + 5
	3rd day (Wed.)	2 lessons[a]	Repetition test 2 (5 min) Session 3: contrast group: units[c] and learning tests 6 + 7, sequential group: units[b] and learning tests 6 – 8
	4th day (Thu.)	2 lessons[a]	Repetition test 3 (5 min) Session 4: contrast group: units[c] learning tests 8 + 9, sequential group: units[c] and learning tests 9 + 10
Follow-up sessions	1 day later (Fri.)	2 lessons	Follow-up session 1: "transformation test" and "explicit transformation explanations" (60 min) Survey part: personal data, logical reasoning test, and arithmetic test (45 min)
	1 week later	1 lesson	Follow-up session 2: "transformation test" and "explicit transformation explanations" (60 min)
	10 weeks later	1 lesson	Follow-up session 3: "transformation test" and "explicit transformation explanations" (60 min)

Note.
[a] In total 90 min.
[b] For each unit: blackboard instruction (appr. 10′), and work sheet (appr. 10′).
[c] For each unit: blackboard instruction (appr. 15′), and work sheet (appr. 15′).
[d] For each learning test (appr. 3–5′).

90-min training sessions on four consecutive days for a total of 9 units. Each unit consisted of three components: receiving blackboard instruction, completing a worksheet, and solving a learning test. The blackboard instruction was provided by the research leader (the first author), who is a licensed primary school teacher. The teacher guided the derivations of the principles by emphasizing the important points, and the students were asked to participate. After the blackboard introduction, the students copied the worked problems from the blackboard on their worksheets and then solved similar problems to apply and reinforce the learned rules. As an indicator of immediate learning gains, the students completed a learning test both after each unit and at the beginning of each day.

In the contrast group, addition and multiplication were simultaneously introduced, i.e., in parallel on the left and right blackboards, by comparing them. In the sequential group, the students practiced addition problems for two days, followed by two days of multiplication training. The two programs contained identical algebra problems, but they differed in the order of the presentation of the problems and tasks (Fig. 1 and Table 2).

Training occurred in groups of 10–15 students at the school. The students were sufficiently spaced so they could not view other students' worksheets when solving the trial tasks and tests. The students were instructed to work on their own and ask the instructors directly if they had questions or problems with the material. The research leader and a research assistant were present at all times to guide the training and testing.

2.3. Materials

We used algebra learning material from a previous study (Ziegler & Stern, 2014). For this study, we converted the self-study introduction section of each unit into guided blackboard instruction, i.e., the self-learning material was converted to direct instruction material. Otherwise, identical material was used. This included the pretest algebra, trial tasks, immediate learning tests, and follow-up tests.

2.3.1. Pretest on prior algebra knowledge: presented in training session 1

Although Swiss sixth graders have not yet received any formal algebra instruction, it is possible that some of them could spontaneously solve algebraic transformation problems by referring to their arithmetic knowledge. The test presented at the very beginning of Session 1 was composed of the following eight algebra problems requiring transformations: "$a + a + a + a =$", "$5 + a + a + 5 + a =$", "$c \cdot c \cdot c =$", "$2 \cdot 2 \cdot z \cdot 2 \cdot z =$", "$7b + 7b =$", "$7b \cdot 7b =$", "$ab \cdot 4\,ab =$", and "$xy + xy + xy + xy =$".

2.3.2. Blackboard instruction: presented in training sessions 1–4

2.3.2.1. Blackboard instruction materials. The algebra material introduced on the blackboard consisted of two series of algebra problems, illustrating nine addition and nine multiplication learning steps (for an excerpt, see Fig. 1). Every learning step exemplified one or two rules of increasing difficulty. The addition and multiplication problems used identical numbers and letters; thus, the problems were differentiated only in the use of addition and multiplication signs, which naturally affected the solution steps and results. This characteristic allowed the algebra problems to be introduced simultaneously, side-by-side in a contrasted format (Fig. 2A). For details regarding the development of the algebra materials, see Ziegler and Stern (2014). In the sequential group, all addition steps were initially introduced, followed by the multiplication steps. For the unit assignments, see Table 2. The contrast group was assigned nine units, whereas the sequential group was assigned ten units. In the sequential version, two blocks were provided per unit to maintain similar amounts of processing. However, the fifth and tenth units contained only one learning step to avoid mixing addition with multiplication.

For each blackboard unit, the students were presented with two blocks of two or three algebra problems: either with one block of addition and one block of multiplication (contrast group) or with two consecutive blocks of the same operation (sequential group). An example of each group is presented in Fig. 2.

2.3.2.2. Introduction of the material by the teacher. The teacher guided the instruction while standing at the blackboard and providing step-wise solutions to the problems, and the students were asked to participate. The instruction was controlled for comparability in both groups and all classes using a script (Appendix A). The teachers' comments aimed to guide the students to more carefully examine the examples and emphasize the elements to which they should attend. In the contrast version, addition and multiplication were introduced in parallel in a contrasted manner. Thus, the teacher's questions, explanations, and hints had the following objectives: to prompt students to compare addition

A3	$xy + xy + xy = 3 \cdot xy = 3xy$ $2b + 2b + 2b + 2b + 2b = 5 \cdot 2b = 10b$ $3cx + 3cx = 2 \cdot 3cx = 6cx$	M3	$xy \cdot xy \cdot xy = x \cdot y \cdot x \cdot y \cdot x \cdot y = x \cdot x \cdot x \cdot y \cdot y \cdot y = x^3 \cdot y^3 = x^3y^3$ $2b \cdot 2b \cdot 2b \cdot 2b \cdot 2b = 2 \cdot b \cdot 2 \cdot b \cdot 2 \cdot b \cdot 2 \cdot b \cdot 2 \cdot b = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot b \cdot b \cdot b \cdot b \cdot b = 32 \cdot b^5 = 32b^5$ $3cx \cdot 3cx = 3 \cdot c \cdot x \cdot 3 \cdot c \cdot x = 3 \cdot 3 \cdot c \cdot c \cdot x \cdot x = 9 \cdot c^2 \cdot x^2 = 9c^2x^2$
A4	$c^2 + c^2 + c^2 + c^2 = 4 \cdot c^2 = 4c^2$ $a^4 + a^4 = 2 \cdot a^4 = 2a^4$ $x^3 + x^3 + x^3 = 3 \cdot x^3 = 3x^3$	M4	$c^2 \cdot c^2 \cdot c^2 \cdot c^2 = c \cdot c \cdot c \cdot c \cdot c \cdot c \cdot c \cdot c = c^8$ $a^4 \cdot a^4 = a \cdot a \cdot a \cdot a \cdot a \cdot a \cdot a \cdot a = a^8$ $x^3 \cdot x^3 \cdot x^3 = x \cdot x \cdot x \cdot x \cdot x \cdot x \cdot x \cdot x \cdot x = x^9$
A5	$2x + 5x + 2x = 9x$ $3bc + bc + 6bc = 3bc + 1bc + 6bc = 10bc$ $y^3 + 4y^3 = 1y^3 + 4y^3 = 5y^3$	M5	$2x \cdot 5x \cdot 2x = 2 \cdot x \cdot 5 \cdot x \cdot 2 \cdot x = 2 \cdot 2 \cdot 5 \cdot x \cdot x \cdot x = 20 \cdot x^3 = 20x^3$ $3bc \cdot bc \cdot 6bc = 3 \cdot b \cdot c \cdot b \cdot c \cdot 6 \cdot b \cdot c = 3 \cdot 6 \cdot b \cdot b \cdot b \cdot c \cdot c \cdot c = 18 \cdot b^3 \cdot c^3 = 18b^3c^3$ $y^3 \cdot 4y^3 = y \cdot y \cdot y \cdot 4 \cdot y \cdot y \cdot y = 4 \cdot y \cdot y \cdot y \cdot y \cdot y \cdot y = 4 \cdot y^6 = 4y^6$
A6	$m + m + a + m + a + m = a + a + m + m + m + m = 2 \cdot a + 4 \cdot m = 2a + 4m$ $4 + x + z + x + 4 + x = x + x + x + z + 4 + 4 = 3 \cdot x + 1 \cdot z + 8 = 3x + z + 8$	M6	$m \cdot m \cdot a \cdot m \cdot a \cdot m = a \cdot a \cdot m \cdot m \cdot m \cdot m = a^2 \cdot m^4 = a^2m^4$ $4 \cdot x \cdot z \cdot x \cdot 4 \cdot x = 4 \cdot 4 \cdot x \cdot x \cdot x \cdot z = 16 \cdot x^3 \cdot z^1 = 16x^3z$

Fig. 1. An excerpt of the algebra problems: addition learning steps (A3–A6) and multiplication learning steps (M3–M6).

Table 2
Worksheets: presentation order of learning steps.

Work sheet	Group	
	Contrast	Sequential
1	A1 + M1	A1 + A2
2	A2 + M2	A3 + A4
3	A3 + M3	A5
4	A4 + M4	A6 + A7
5	A5 + M5	A8 + A9
6	A6 + M6	M1 + M2
7	A7 + M7	M3 + M4
8	A8 + M8	M5
9	A9 + M9	M6 + M7
10	–	M8 + M9

Note. A1 − A9 = addition learning steps, and M1 − M9 = multiplication learning steps.

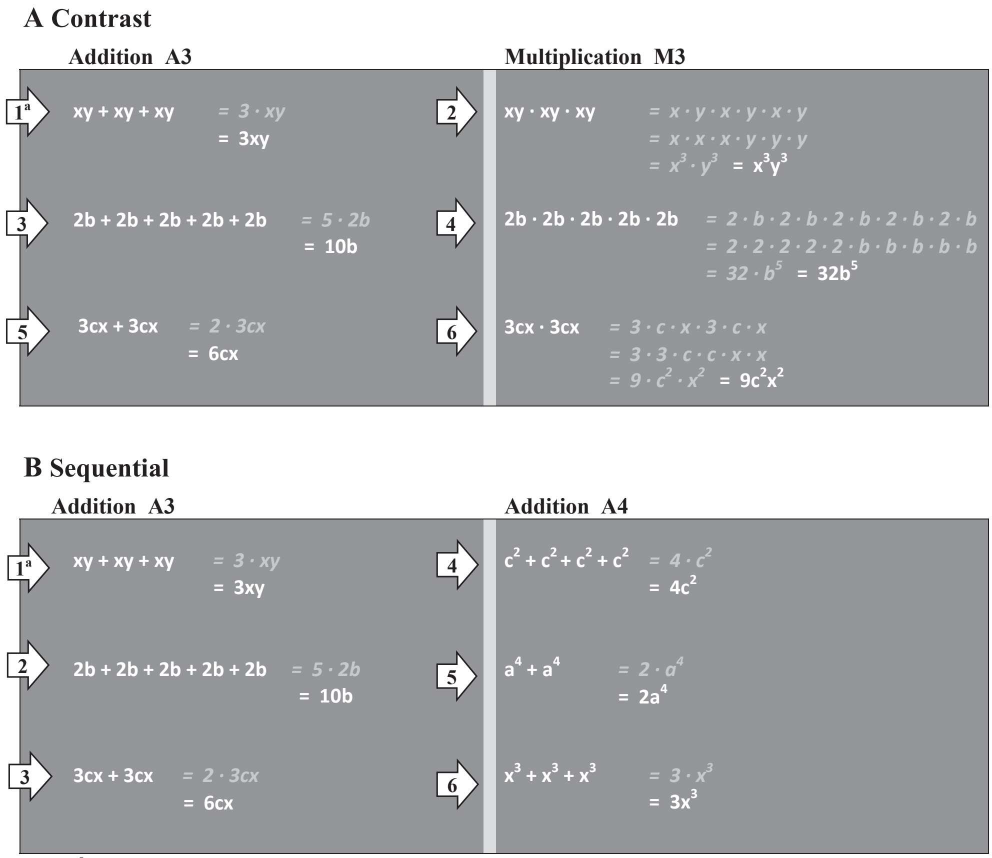
and multiplication, to emphasize the distinction between the two principles, and to clarify how to avoid confusing addition and multiplication. In the sequential group, the material was not appropriate for comparison, but the aim was to guarantee similarly deep processing; thus, the teacher's questions, explanations, and hints had the following objectives: to describe the solutions in detail, to explain how to solve such problems, and to indicate which elements upon which to focus. The length of instruction overall learning steps was identical for both groups because identical problems were introduced, only in a different order.

2.3.2.3. Implementation check. The research leader acted as instructor and was not blind to the research hypotheses. To ensure that comparable instructions were given to the contrast and sequential groups, the lessons were videotaped. The instructor followed detailed scripts, which were exactly observed. Therefore, the videos were not analyzed for content but instead were rated for cooperation of the students and engagement of the instructor. Six different single problems were chosen to be rated by twelve independent observers. The problems were cut from videos of three randomly chosen classes, and the video files were then transformed to audio files. In every case, two files with the same problem, one for a contrast group and the other for a sequential group, were juxtaposed. The raters listened to the two sets of instructions and assessed their degree of correspondence on a 3-point Likert scale (more, equally, or less engaged and more, equally, or less convincing). The audio files were rated by observers who were not informed about the experimental variation with regard to the instructions. There were two measures. *Student cooperation* had two sub-measures: 1) participation, as measured by the number of student statements or student answers during blackboard instruction; and 2) disturbances, as measured by the number of rebukes by the teacher, e.g., “put down the pen”, “wait to write”, or “stop chatting”. *Instructor engagement* also had two sub-measures: 1) atmosphere, i.e., “in which version does the voice of the instructor sound more engaged”; and 2) persuasiveness, i.e., “which set of instructions sounds more convincing”.

2.3.3. Worksheets: presented in training sessions 1–4

For each intervention unit, there was a worksheet that consisted of two sections: the copy section and the trial task section. The structure and content of the worksheets were identical for both conditions except for the different presentation orders of the problems, which resulted from their different introduction orders in the contrast and sequential versions.

2.3.3.1. The copy section. Free space was provided to copy the algebra problems from the blackboard to the worksheet. There were no other instructions or questions on the worksheet. For the contrast group, the space was divided into two columns for the two sides of the blackboard; for the sequential group, there was only one column with two rows.



Note. [a] The arrows indicate the introduction order of problems.

Fig. 2. Blackboard picture. The algebra problems of unit no. 3 for the contrast group (A). The algebra problems of unit no. 2 for the sequential group (B). White bold = the problems and results that were written in white chalk; grey italics = the intermediate steps to the solution that were written in red chalk.

2.3.3.2. The trial task section. In the first section, the students were given two trial problems, which were similar to the presented blackboard examples, for each block of algebra problems. The students were asked to record the intermediate steps of their solutions. The trials were assessed by the research leader and if necessary, corrected by the students before continuing to the next section. In the second section, the students were presented with more trial problems to apply and deepen the learned algebra principles. Both groups were given 4 to 6 trial problems per sheet, with solution instructions to record the intermediate steps in addition to the solution. Before the research leader corrected the trials, the students were told to review the problems on their own using the examples on the front of the page. For the contrast group, there were 46 problems; for the sequential group, there were 40 problems. The six missing problems were part of the 20 mixed algebra problems that the sequential group solved at the end of the fourth training day (see next paragraph).

2.3.3.3. Mixed algebra problems. The sequential learners were provided additional training on 20 mixed addition and multiplication problems at the end of the fourth training day on two additional sheets with the following instruction: “You learned addition and multiplication in algebra; now you will receive mixed problems to solve”. The problems were then checked and returned to the students for correction. This mixed training for the sequential group was designed both to ensure that the sequential learners would not be at a disadvantage and to mimic the teaching provided in an ordinary school setting. After the sequential introduction of the two operations, mathematics teachers usually assign mixed problems to ensure that both concepts are correctly understood. In the contrast group, these 20 problems were integrated into the normal learning materials of the intervention units (6 in the trial task problems, 10 in the immediate learning tests, and 4 in the repetition tests), because they always processed mixed addition and multiplication problems. The mixed problems of the sequential group were taken from the materials of the contrast program for two reasons: first, to avoid giving the sequential group additional training material that the contrast group did not process; and second, to keep time and the number of processed problems constant between the two groups. Although identical learning and practice material was used for both groups over the entire intervention, the presentation order of the material was unique to each group (for the only exception, see the repetition test information in Section 2.3.4.2).

When the students completed their worksheets, they were given optional non-mathematical tasks to prevent them from disturbing other students, e.g., determining the differences between two pictures or solving riddles or Sudoku puzzles.

2.3.4. Immediate learning tests: presented in training sessions 1–4

To compare the immediate learning gains of both groups, two types of algebra tests were given to all participants in sessions 1–4: learning tests and repetition tests.

2.3.4.1. Learning tests. Immediately after every worksheet, students were given a *learning test* in which they solved 3–8 algebra items per test sheet without being allowed to use the instruction material. These problems were similar to those presented on the work sheet that the students had just finished. When students did not remember the examples, they were instructed to guess the solution and told that the next learning step would repeat the solution steps for such problems. This step allowed us to assess what the students had learned during the session. Because there were 9 tests for the contrast group and 10 tests for the sequential group (see 2.3.2.1), the single tests are not directly comparable. Therefore, the scores of all the single learning tests of a student were totaled to measure performance on "learning tests". For the contrast groups, there were 58 items; for the sequential group, there were 48 items. The 10 missing items were items from the 20 mixed problems of the sequential group and were not included in the total score. They were parallelized items of similar items, such that every difficulty was represented in the sequential learning tests, e.g., "$e \cdot e \cdot e =$" was only in the contrast group, but there was a similar item "$p \cdot p \cdot p =$" in both groups, or "$m + c + c + m + c =$" compared to "$n + b + n + x + b + n =$".

2.3.4.2. Repetition tests. At the beginning of days 2–4, students were given a *repetition test* in which they solved 6–8 algebra items similar to those upon which they had worked the day before, to assess what they remembered. The solution rates of the three tests were totaled to measure solution rates on "repetition tests". For the contrast groups, there were 24 items; for the sequential group, there were 20 items. Because the presented blackboard examples varied in the introduction order over the two conditions and because there were only tests on days 2–4, two of the items were different in the contrast and sequential group. In the total score, we only included the items that were identical.

2.3.5. Control measures: presented in follow-up session 1

To control for individual characteristics that might affect algebra learning, four measures were assessed.

2.3.5.1. Logical reasoning test. Logical reasoning ability was assessed with a figural and a numerical subtest of a German intelligence test (subtests 3 and 4 of the LPS by Horn (1983)). The test is based on Thurstone's primary mental abilities with a maximum score of 40 for each subtest.

2.3.5.2. Arithmetic test. Arithmetic knowledge was assessed with two speed tests consisting of two sheets, each with 28 arithmetic items. Students were required to solve each sheet within 90 s and were only allowed to write the results. The first sheet contained two-digit additions, and the second sheet contained two-digit multiplications. The arithmetic knowledge score was determined by the solution rates of correct answers.

2.3.5.3. Mathematics and German grades. The students' mathematical achievement was measured in the form of their school grades in mathematics and German. The grades were reported by students' teachers. In Switzerland, grades range from 6 (best) to 1 (worst).

2.3.6. Transformation test: presented in follow-up sessions 1–3

The *transformation test* was applied to assess transformation knowledge. The test contained 58 items that assessed the students' ability to correctly apply transformations practiced in sessions 1–4 (e.g., "$a^2 \cdot a \cdot ay \cdot 4a =$", "$5ab + b + 3b + 2ab + 2b =$", "$y \cdot y^3 \cdot y^2 \cdot y =$", "$2 + 5x + 4 + 2x + 3 =$"). Examples from every work sheet were represented, with 2–4 problems representing each learning step. The items were matched to the problems of the worksheets with other variables and numbers, i.e., every item was an algebraic term that needed to be transformed into the shortest version by applying the correct rules. The items were ordered in increasing difficulty but always overlapped by 2 or 3 learning steps, within which the items were randomly mixed. The same algebra test was applied in all three follow-up sessions because students were not supposed to remember the items.

2.3.6.1. Scoring. Transformation knowledge was determined by the number of correct answers to the 58 algebra items, i.e., a maximum of 58 points were possible. In addition to the total score of the correctly solved problems, we analyzed for careless errors as a method of checking for the equality of the groups. Careless errors included mistakes such as miscounting the number of letters, e.g., "$n \cdot n \cdot n \cdot n \cdot n = n^4$", or an arithmetic error, e.g., "$b \cdot a \cdot 4 \cdot a \cdot 4 \cdot a = 18a^3b$".

2.3.7. Explicit transformation explanations: presented in follow-up sessions 1–3

The test of "*explicit transformation explanations*" was applied to assess explicit transformation knowledge and misconceptions. This test assessed students' ability to elucidate how to apply algebraic additions and multiplications. Students were asked to write down two separate descriptions that explicitly explained how to solve each type of problem. For each explanation, they were prompted with four hints, which were designed to help activate their knowledge: (a) "Describe in detailed steps how problems with letters are solved", (b) "Mention what one has to pay attention to", (c) "You can explain it by means of examples", and (d) "Imagine that you would like to explain the rules to classmates".

2.3.7.1. Scoring. The *explicit transformation explanations* of algebraic addition and multiplication were scored for accuracy and completeness of the answer. To do so, we developed a coding scheme, which is depicted in Appendix B. Two measures were assessed. First, *explicit transformation knowledge* was judged by the amount of correctly reported algebra concept features. For both addition and multiplication, the features were divided into defining features and secondary features, and the scores for each were added to the overall score of explicit transformation knowledge. Second, *misconceptions* were judged by the number of errors in the explicit transformation explanations, including errors in the examples used in the explanations. The errors of both the addition and multiplication explanations were added for a total misconceptions score.

Two trained raters independently coded the transformation explanations. The raters discussed and unified diverging judgments. Inter-rater reliability for coding the answers was 93.8% for the first point of measurement and 94.3% for the second point of measurement (exact agreement). Because of the high reliability, there was only one of the two raters who coded the third point of measurement.

3. Results

The results are presented in four sections: the students' preconditions, the implementation check, the group effects on the follow-up measures, and the group effects on the immediate learning measures.

3.1. The students' preconditions: algebra pretest and control variables

3.1.1. Prior algebra knowledge

As expected, given the age of students chosen for the study, there was no difference between groups on prior algebra knowledge, $F(1, 89) = .02$, $p = .878$, $\eta^2 = .00$. A floor effect was found, indicating that the students appeared to have almost no prior algebra knowledge before the training (for the contrast group, $M = 1.17$, $SD = 1.3$; for the sequential group, $M = 1.13$, $SD = 1.2$). The few correctly solved tasks showed that the students were solving almost only the algebraic addition problems with equal letters, similar to length measures that they had learned in school ("a + a + a + a = "; "7b + 7b =", and "xy + xy + xy + xy ="; however, they appeared to have no intuition about solving algebra problems that went beyond these simple additions.

3.1.2. Control variables

The control variables included logical reasoning, arithmetical knowledge, and grades in mathematics and German. A multivariate analysis of variance (MANOVA) revealed no significant difference between groups, $F(4, 86) = .44$, $p = .908$, $\eta^2 = .01$. None of the separate univariate ANOVA tests showed a significant effect (Table 3), thus indicating that the contrast and sequential groups did not differ either in cognitive preconditions or in school achievement.

3.2. Implementation check

3.2.1. Student cooperation

A MANOVA revealed no difference between groups, $F(2, 7) = .36$, $p = .707$, $\eta^2 = .09$. None of the separate univariate ANOVA tests showed an effect for participation (contrast group: $M = 155.8$ statements, $SD = 26.0$; sequential group: $M = 140.2$ statements, $SD = 29.7$), $F(1, 8) = .78$, $p = .403$, $\eta^2 = .09$, or for disturbances (contrast group: $M = 12.8$ interruptions, $SD = 4.8$; sequential group: $M = 13.6$ interruptions, $SD = 12.4$), $F(1, 8) = .02$, $p = .897$, $\eta^2 = .00$, indicating that the contrast group and sequential group did not differ.

3.2.2. Instructor engagement

There were 72 ratings for each measure. The distribution of the ratings showed a balanced picture. For atmosphere, the ratings were as follows: 24 found the versions equally engaging, 23 favored the contrast version, and 25 favored the sequential version. For persuasiveness, the ratings were as follows: 32 found the versions equally convincing, 23 favored the contrast version, and 17 favored the sequential version.

3.3. Effect of group on performance in the follow-up sessions

Students in the contrast group performed better than students in the sequential group on the three follow-up measures (Hypothesis 1): transformation knowledge, explicit transformation knowledge, and misconceptions. The effects were maintained over time (Hypothesis 2). Fig. 3 illustrates the means and standard errors across the three follow-up sessions for both groups. For each of the follow-up measures, separate mixed-factorial ANOVAs were conducted, with group as the between-subjects factor (contrast versus sequential) and time as the within-subjects factor (T1: one day, T2: one week, or T3: ten weeks). When there was a main effect of group, separate post-hoc comparisons were conducted for the three measurement points. The students in the contrast group were expected to perform better on all follow-up measures. Following the identification of a main effect of time, post-hoc tests were conducted to determine how stable the effects were over time.

Table 3
Students' performance by group on the control variables.

	Contrast $N = 46$	Sequential $N = 45$	
	M (SD)	*M (SD)*	*p*
Arithmetic knowledge (in %)	30.67 (12.8)	30.00 (10.4)	.785
Logical reasoning (max. 40)	27.83 (3.37)	27.54 (3.43)	.694
Mathematics grade[a]	4.71 (.70)	4.72 (.68)	.914
German grade[a]	4.75 (.52)	4.69 (.54)	.584

Note.
[a] Range from 6 (best) to 1 (worst).

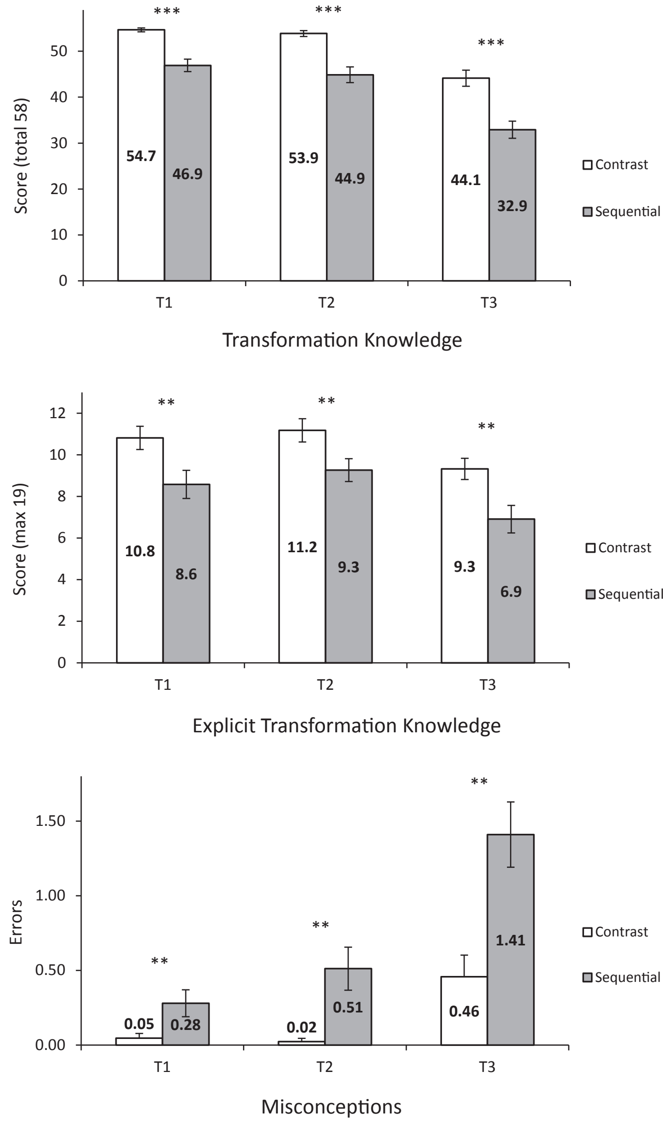


Fig. 3. Performance on the follow-up measures. T1 = 1 day later, T2 = 1 week later, and T3 = 10 weeks later. *$p < .05$, **$p < .01$, ***$p < .001$. The error bars represent the standard error.

3.3.1. Transformation knowledge

As expected, there was a main effect of group, $F(1, 86) = 35.98$, $p < .001$, $\eta^2 = .30$, which favored the contrast group. Post-hoc tests indicated significant group differences at T1, $t(86) = 5.48$, $p < .001$, $d = 1.21$, T2, $t(89) = 4.84$, $p < .001$, $d = 1.04$, and T3, $t(89) = 4.43$, $p < .001$, $d = .94$. These results indicate a strong advantage of contrast learning. There was also a main effect of time, $F(1.33, 114.50) = 92.73$, $p < .001$, $\eta^2 = .52$, but no interaction effect. The simple contrasts indicate a significant change from T1 to T2, $F(1, 86) = 5.32$, $p = .024$, $\eta^2 = .06$, and from T1 to T3, $F(1, 86) = 111.83$, $p < .001$, $\eta^2 = .57$, which demonstrate that scores declined after one week and then dropped even further. This decrease in performance after several weeks, during which the material was not repeated, is both normal and expected in learning environments. As expected, for careless errors, there was neither a group effect, $F(1, 86) = 1.17$, $p = .283$, $\eta^2 = .01$, nor an interaction effect, $F(2, 172) = 1.09$, $p = .337$, $\eta^2 = .01$, which indicates that careless errors were independent of group and intervention. However, there was a time effect, $F(2, 172) = 3.35$, $p = .037$, $\eta^2 = .04$, with more careless errors at the second measurement point and a low rate of careless errors at the third measurement point for both groups.

3.3.2. Explicit transformation knowledge

As expected, there was a significant main effect of group, $F(1, 84) = 10.29$, $p = .001$, $\eta^2 = .11$. Significant post-hoc tests at all points of measurement indicated that the students from the contrast group expressed more explicit knowledge in their written explanations compared to the students from the sequential group at T1, $t(85) = 2.55$, $p = .006$, $d = .55$; T2, $t(88) = 2.44$, $p = .009$, $d = .52$; and T3, $t(88) = 2.89$, $p = .003$, $d = .62$. There was also a main effect of time, $F(2, 168) = 11.94$, $p < .001$, $\eta^2 = .12$. The simple contrasts indicate a significant change from T1 to T3, $F(1, 84) = 11.40$, $p = .001$, $\eta^2 = .12$, but not from T1 to T2, $p = .179$, $\eta^2 = .02$, so that the effects remained stable for one week and then declined. There was no significant interaction effect, $p = .902$, $\eta^2 = .001$.

3.3.3. Misconceptions

As expected, there was a main effect of group on misconceptions, $F(1, 84) = 28.52$, $p < .001$, $\eta^2 = .25$. The post-hoc tests showed group differences that favored the contrast group at T1, $t(85) = -2.45$, $p = .009$, $d = .53$, at T2, $t(88) = -3.35$, $p = .001$, $d = .71$, and at T3, $t(88) = -3.63$, $p = .001$, $d = .78$. There was a main effect of time, $F(1.51, 126.79) = 20.38$, $p < .001$, $\eta^2 = .20$; the simple contrasts revealed a significant increase in errors from T1 to T3, $F(1,84) = 27.89$, $p < .001$, $\eta^2 = .25$, but not from T1 to T2, $p = .134$, $\eta^2 = .03$, and there was a significant interaction effect, $F(1.51, 126.79) = 5.00$, $p = .015$, $\eta^2 = .06$, so that the increase in the number of misconceptions during this time period was more pronounced in the sequential group.

3.4. Effect of group on immediate learning

The previous section indicates a clear superiority of the contrast group on all three follow-up measures. However, contrary to our expectations for immediate learning, the sequential group clearly outperformed the contrast group (Hypothesis 3). Fig. 4 illustrates the means and standard errors for the two measures.

3.4.1. Immediate learning gains

A MANOVA shows a significant effect of group, $F(2, 87) = 10.45$, $p < .001$, $\eta^2 = .19$. The students in the sequential group outperformed the students in the contrast group on both immediate learning measures. The separate univariate ANOVA tests on the outcome variables indicate significant treatment effects on the repetition tests, $F(1, 88) = 9.14$, $p = .003$, $\eta^2 = .09$, and learning tests, $F(1, 88) = 20.36$, $p < .001$, $\eta^2 = .19$. Therefore, relative to short-term learning, the results demonstrate a clear advantage for the sequential group.

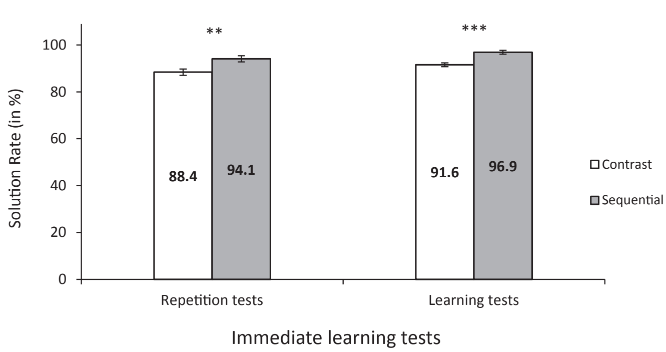


Fig. 4. Performance on the immediate learning tests. $^{**}p < .01$, $^{***}p < .001$. The error bars represent the standard error.

Inspecting the solution rates of the single learning tests after each session revealed no consistent trends in solution rates over time (see Table 4). Rather, they appear to have varied according to the difficulty of the learning material. No statistical analysis was performed because the single learning steps were not comparable due to the different introductory order of problems in the two groups (see Section 2.3.2.1).

4. Discussion

Previous research has demonstrated the benefits of contrasting in self-learning settings (Alfieri et al., 2013). Our results extended the applicability of contrasting to the direct instruction method, which remains a common and widespread method of instruction in school. We adapted the algebra self-learning materials from a previous experiment on direct instruction at the blackboard. The instructor acted as a classroom teacher, aided only by one assistant who corrected the materials continuously to enable intervention on four successive days. Thus, we replicated previous results regarding the effects of repeated contrasted comparisons in an ecologically valid classroom setting. The concern that the teacher's explicit highlighting of differences between concepts might neutralize the effect of contrasting was removed; that is, contrasting under teacher-directed instruction appears not to excessively reduce cognitive load, which might have eliminated the advantage of contrasting.

Overall, we clearly replicated the finding that contrasting is an effective means of algebraic concept learning (Ziegler & Stern, 2014). Our first aim was to examine the contrasting effects on algebraic problem solving. As expected, the contrast group outperformed the sequential group at the one-day follow up on all the measures (Hypothesis 1). The effects remained strong at the one- and ten-week follow-ups for transformation knowledge and became even stronger for explicit transformation knowledge and misconceptions, thus indicating that contrast learning improves long-term memorization (Hypothesis 2).

Our second aim was to determine whether the learning trajectories of the contrast and the sequential approaches were equivalent. Contrary to our expectation, however, there was a pattern of

Table 4
Mean solution rates and standard deviations (in parentheses) of students' performance by group on the single learning tests and repetition tests.

		Contrast group	Sequential group
		M (SD)	*M (SD)*
Learning tests	W1	89.1 (23.1)	98.9 (4.2)
	W2	96.4 (7.0)	96.8 (6.7)
	W3	89.8 (16.1)	98.5 (6.9)
	W4	92.8 (14.3)	94.2 (11.8)
	W5	91.7 (13.0)	95.0 (10.1)
	W6	95.7 (10.9)	98.5 (4.8)
	W7	87.3 (15.4)	96.7 (8.4)
	W8	89.1 (31.5)	97.0 (9.6)
	W9	93.1 (10.3)	96.4 (8.8)
	W10	–[a]	96.3 (10.6)
Repetition tests	d2	90.8 (11.1)	97.0 (6.4)
	d3	89.8 (15.1)	99.1 (4.2)
	d4	83.1 (23.7)	87.9 (18.5)

Note.
[a] The contrast group processed only 9 worksheets (see Section 2.3.2.1). W1–W10 = worksheets 1–10, d2 – d4 = days 2–4.

long-term advantages versus short-term disadvantages for teacher-directed learning similar to that found in our self-learning study (Hypothesis 3). Similar to our results in the self-learning study, we observed this opposite effect in the immediate learning results, which favored the sequential group in both immediate learning measures, indicating that sequential learning is convenient for short-term memorization. This pattern of delayed benefits corresponds to the retention effects observed in interleaved or mixed-concept learning (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Rohrer & Pashler, 2010). Delayed gains indicate more demanding processing during the training phase (E. Bjork & Bjork, 2011). Our hypothesis that teacher's explanations reduce cognitive load and will therefore increase contrast students' immediate results to the same level as those of sequential students was not supported. Our results show that it is not enough to reduce cognitive load via direct instruction. Rather, contrasting itself makes a difference and positively affects the development of student knowledge, although the effects only appear delayed.

In both this study and a previous study (Ziegler & Stern, 2014), we have demonstrated that although contrasting is more effective, it is more demanding than sequential learning for similar but structurally different algebra principles in secondary school. These findings underline our claim for an optimal level of the cognitive load of learning materials in the form of desirable difficulty and interconnection. However, mixing might not be better under all circumstances. Particularly when concepts are dissimilar, processing only one type of problem may be more helpful in learning its characteristics (Goldstone, 1996). Based on the results of our algebra material, we maintain that contrasting different concepts is useful in introducing similar and therefore easily confused concepts or strategies. However, this result may not hold for inherently difficult content. In their cognitive load theory, Sweller and Chandler (1994) emphasized that concepts with an intrinsically high degree of element interactivity often must be taught in isolation. Mixing different types of problems based on such concepts may overtax learners. In our study, it was not the difficulty of the concepts or the complexity of the strategies that may have impeded learning but the danger of confusion. Therefore, contrasted comparison was a useful method in the long run in learning to distinguish between two algebraic principles.

Therefore, we cannot support the claim that younger and less skilled students require separate practice or instruction before materials can be mixed (Dunlosky et al., 2013). In general, our results with algebra novices demonstrate that this recommendation is not valid. In the long run, students gained from a mixed introduction to two contrasting principles, illustrating that it helps learners to attend to two contrasting principles from the beginning. Our study demonstrates that the benefits of contrasted comparison only appear with delay, which may cause these benefits to be overlooked.

To determine whether textbooks should be changed to introduce similar concepts in a contrasted format, further controlled studies are needed to replicate the present results for various content areas. The following concepts could be considered for mathematics: addition versus multiplication with decimal numbers or fractions, proportionality versus reverse proportionality, Venn versus Carroll diagrams, and differential versus integral calculus. Additionally, in science, one can find numerous suitable candidates for contrasted introductions of similar concepts. For example, in biology, possibilities include mitosis versus meiosis, monohybrid versus dihybrid inheritance, transcription versus translation in genetics, and the metabolism of autotrophic versus heterotrophic unicellular organisms. In studies of contrasted comparison in these areas, it would be necessary to control for the similarity and complexity of the concepts (Dunlosky et al., 2013).

The challenge in investigating comparison and sequential settings in an experimental design is to parallelize the instructional materials. In our study, the juxtaposition both changed the introductory order of problems and impeded the use of identical verbal instructions. However, we designed the instructions to keep the wording in the two conditions almost identical. A potential conflict of interest may arise from the role of the research leader in administering the blackboard instruction. However, we videotaped the lessons and checked for equal engagement of the instructor and cooperation of the students.

Although we demonstrated the superiority of a simultaneous and contrasted introduction over a sequential introduction, we cannot exclude the possibility that there are additional methods of designing learning material. The contrast learners learned the algebra material more effectively than the sequential learners, although they did not learn it perfectly. The question then arises of how to further improve knowledge following a contrasted introduction: with more contrast training, with separate practice for each principle, or with a combination and alternation of comparison and separate training sessions. Further investigations must determine whether more contrast training or additional combined or alternated training would be advantageous and if so, how long the respective comparison or sequenced sessions should last. Another worthwhile line of research that could be combined with studies of contrasting and mixed training would involve beginning with an overview of all of the concepts or principles to be introduced, i.e., a global or holistic approach that precedes the stepwise introduction of concepts or principles (Calcaterra, Antonietti, & Underwood, 2005; Munro, 1993; Pask, 1976). If providing a global perspective on materials combined with a comparison of individual concepts is successful, it would be broadly applicable in many academic fields.

By replicating the contrasting effects gained in a self-learning setting under a teacher-directed instruction format, we have demonstrated that comparisons are operative under both methods. It would be interesting to directly compare a concept introduction under a self-study condition with its introduction under a direct instruction condition. Further studies should be conducted to determine the optimal conditions for comparison instructions, that is, to find additional optimal levels of difficulty and to improve the adequacy of the cognitive load associated with the learning material. Nonetheless, we maintain that given the current scientific evidence, the time is favorable for a broad application of contrasted comparisons. We have contributed to the discussion with two experimentally controlled intervention studies performed under realistic classroom conditions over four days and that included both immediate and follow-up measures. Because our contrast and sequential programs were identical in content and differed only in the presentation order of the examples and trial tasks, it would be relatively easy to adapt the learning materials and books to a contrasted instruction format.

Acknowledgments

We are grateful to Sara Ziegler for her assistance in the implementation of the Experiments in the school classes, and Sara Ziegler and Adrienne Suter for their assistance in the evaluation of the test materials.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.learninstruc.2015.09.006.

References

Alfieri, L., Nokes, T. J., & Schunn, C. D. (2013). Learning through case comparisons: a metaanalytic review. *Educational Psychologist, 48*(2), 87–113. http://dx.doi.org/10.1080/00461520.2013.775712.

Bjork, R. (1994). Institutional impediments to effective training. In D. Druckman, & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 295–306). Washington, DC: National Academy Press.

Bjork, E., & Bjork, R. (2011). Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth Publishers.

Calcaterra, A., Antonietti, A., & Underwood, J. (2005). Cognitive style, hypermedia navigation and learning. *Computers & Education, 44*(4), 441–457. http://dx.doi.org/10.1016/j.compedu.2004.04.007.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology Learning Memory and Cognition, 15*(6), 1147–1156. http://dx.doi.org/10.1037/0278-7393.15.6.1147.

Chase, C. C., Shemwell, J. T., & Schwartz, D. L. (2010). Explaining across contrasting cases for deep understanding in science: an example using interactive simulations. In *Paper presented at the proceedings of the 9th International Conference of the learning Sciences*.

Christie, S., & Gentner, D. (2010). Where hypotheses come from: learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373. http://dx.doi.org/10.1080/15248371003700015.

Cummins, D. D. (1992). Role of analogical reasoning in the induction of problem categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 1103–1124. http://dx.doi.org/10.1037/0278-7393.18.5.1103.

Day, S., & Goldstone, R. L. (2010). The effects of similarity and individual differences on comparison and transfer. In *Paper presented at the thirty-second Annual Conference of the cognitive science society*. Portland, Oregon: Cognitive Science Society.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. http://dx.doi.org/10.1177/1529100612453266.

Gadgil, S., Nokes, T. J., & Chi, M. T. H. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction, 22*(1), 47–61. http://dx.doi.org/10.1016/j.learninstruc.2011.06.002.

Gentner, D. (2010). Bootstrapping the mind: analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775. http://dx.doi.org/10.1111/j.1551-6709.2010.01114.x.

Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. D. (2009). Reviving inert knowledge: analogical abstraction supports relational retrieval of past events. *Cognitive Science, 33*(8), 1343–1382. http://dx.doi.org/10.1111/j.1551-6709.2009.01070.x.

Gerjets, P., Scheiter, K., & Schuh, J. (2008). Information comparisons in example-based hypermedia environments: supporting learners with processing prompts and an interactive comparison tool. *Education Technology Research and Development, 56*(1), 73–92. http://dx.doi.org/10.1007/s11423–007–9068-z.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory and Cognition, 24*(5), 608–628. http://dx.doi.org/10.3758/BF03201087.

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*(3), 280–290. http://dx.doi.org/10.1016/j.jecp.2010.04.017.

Hardy, I., Jonen, A., Moller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology, 98*(2), 307–326. http://dx.doi.org/10.1037/0022-0663.98.2.307.

Hattie, J. A. (2009). *Visible learning for teachers. A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Horn, W. (1983). *Leistungsprüfsystem LPS. 2. Auflage*. Göttingen: Hogrefe.

Kang, S. H. K., & Pashler, H. (2012). Learning painting styles: spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology, 26*(1), 97–103. http://dx.doi.org/10.1002/acp.1801.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences, 21*(1), 45–83. http://dx.doi.org/10.1080/10508406.2011.591717.

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. http://dx.doi.org/10.1207/s15326985ep4102_1.

Kurtz, K. J., Miao, C. H., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences, 10*(4), 417–446. http://dx.doi.org/10.1207/S15327809JLS1004new_2.

Liem, G. A., & Martin, A. J. (2013). Direct instruction. In J. A. Hattie, & E. M. Anderman (Eds.), *International guide to student achievement*. New York: Routledge.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: close comparisons facilitate far mappings. *Journal of Cognition and Development, 2*(2), 189–219. http://dx.doi.org/10.1207/S15327647JCD0202_4.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*(4), 586–597. http://dx.doi.org/10.3758/BF03212967.

Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical learning in negotiation teams: comparing cases promotes learning and transfer. *Academy of Management Learning and Education, 2*(2), 119–127. http://dx.doi.org/10.5465/AMLE.2003.9901663.

Markman, A. B., & Gentner, D. (1993). Splitting the differences - a structural alignment view of similarity. *Journal of Memory and Language, 32*(4), 517–535. http://dx.doi.org/10.1006/jmla.1993.1027.

Mason, L. (2004). Fostering understanding by structural alignment as a route to analogical learning. *Instructional Science, 32*(4), 293–318. http://dx.doi.org/10.1023/B: TRUC.0000026512.88700.32.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning. *American Psychologist, 59*, 14–19. http://dx.doi.org/10.1037/0003-066X.59.1.14.

Michael, A. L., Klee, T., Bransford, J. D., & Warren, S. F. (1993). The transition from theory to therapy: test of two instructional methods. *Applied Cognitive Psychology, 7*(2), 139–153. http://dx.doi.org/10.1002/acp.2350070206.

Munro, J. (1993). *Mathematics teaching: Matching individual learning preferences*. Melbourne: MERA.

Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology, 46*(2), 128–148. http://dx.doi.org/10.1111/j.2044-8279.1976.tb02305.x.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science, 316*(5828), 1128–1129. http://dx.doi.org/10.1126/science.1142103.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural, knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*(3), 561–574. http://dx.doi.org/10.1037/0022-0663.99.3.561.

Rittle-Johnson, B., & Star, J. R. (2011). The power of comparison in learning and instruction: learning outcomes supported by different types of comparisons. *Psychology of Learning and Motivation: Cognition in Education, 55*, 199–225. http://dx.doi.org/10.1016/B978-0-12-387691-1.00007-7.

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39*(5), 406–412. http://dx.doi.org/10.3102/0013189X10374770.

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376–392). New York: Macmillan.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–522. http://dx.doi.org/10.1207/s1532690xci1604_4.

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. http://dx.doi.org/10.1016/j.jecp.2008.11.004.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.), *The psychology of learning and motivation* (vol. 43, pp. 215–266). San Diego, CA: Academic Press.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185–233. http://dx.doi.org/10.1207/s1532690xci1203_1.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. http://dx.doi.org/10.1023/A:1022193728205.

VanderStoep, S. W., & Seifert, C. M. (1993). Learning "how" versus learning "when": Improving transfer of problem-solving principles. *Journal of the Learning Sciences, 3*(1), 93–111. http://dx.doi.org/10.1207/s15327809jls0301_3.

Ziegler, E., & Stern, E. (2014). Delayed benefits of learning elementary algebraic transformations through contrasted comparisons. *Learning and Instruction, 33*(1), 131–146.

Ziegler, E. & Stern, E. (submitted). The benefit of explicit comparisons for meaningful concept learning in mathematics.