ORIGINAL ARTICLE

Simple pictorial mathematics problems for children: locating sources of cognitive load and how to reduce it

Ernest C. D. M. van Lieshout[1] 🔴 · Iro Xenidou-Dervou[2] 🔴

Abstract

Pictorial representations are often used to help children understand the situation described in a given number-sentence scheme. These static pictorial problems essentially attempt to depict a dynamic situation (e.g., one bird flies away while there are three birds still sitting on the fence). Previous research suggested that such pictorial decrease problems impose higher cognitive load on children than the corresponding increase problems, even though both are solved with addition. However, the source of this cognitive load is unclear. It could be the direction of the depicted change or the position of the unknown (start vs. end set). To address this question and disentangle the sources of the load, we presented the problems in two different formats: (1) the conventional static one-picture problems and, (2) an adapted three-picture problem-format, which depicted the dynamic change in sequential steps. We also examined whether the three-picture problem-format makes the decrease problems easier. Seventy-nine first-graders participated in this study. Results showed that, overall, problems with the position of the unknown at the end were easier to solve than the ones in which the unknown was at the start. Furthermore, three-picture decrease problems with the unknown in the last position were easier than the one-picture decrease problems, and therefore appear to be a meaningful way to make such problems easier for children to understand.

Keywords Arithmetic · Pictorial problems · Animation · Modality effect · Mathematics education · Primary school

1 Introduction

Word problems are widely used in mathematics education. Research during the past decades has shed light on the factors that influence children's performance in these problems (e.g., De Corte and Verschaffel 1981, 1987; Hiebert 1982). In the meantime, the use of such problems has also encountered much criticism, especially from adherents of the Realistic Mathematics Education movement which originated in the Netherlands (Gravermeijer 1997). Although word problems were intended to be descriptive representations of real-life situations, they often were not. This situation led many to believe that perhaps they were not optimal for mathematics education because the answers produced in

such problems were unrealistic or were not useful for real-life (e.g., Dewolf et al. 2016; Greer 1997; Verschaffel et al. 1994, 2000). So, the construction of more authentic real-life problems was deemed necessary. One way of achieving this goal was to use pictures instead of (only) words; these are known as pictorial problems.

Recently, Hoogland and colleagues (2018) addressed the question of whether pictorial problems are more beneficial compared to traditional word problems in primary and secondary education. Their results were in favour of the pictorial problems; however, the effect size was very small ($d = .02$). Large sample sizes, such as the one in their study ($N = 31,842$; Hoogland et al. 2018), increase the chances of a relatively weak effect being statistically significant. This low impact of pictorial problems on students' performance is not surprising. Pictures in problems can indeed be helpful, but they can also be confusing, depending on the way the information is presented. Berends and van Lieshout (2009) found that pictorial problems can impose high cognitive load (Sweller 1988) on primary school children. Therefore, we need more theory-driven fine-grained research on how the characteristics of pictorial information contained

✉ Ernest C. D. M. van Lieshout
E.C.D.M.van.Lieshout@vu.nl

1 Faculty of Behavioural and Movement Sciences, Section Educational Studies, Vrije Universiteit Amsterdam, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

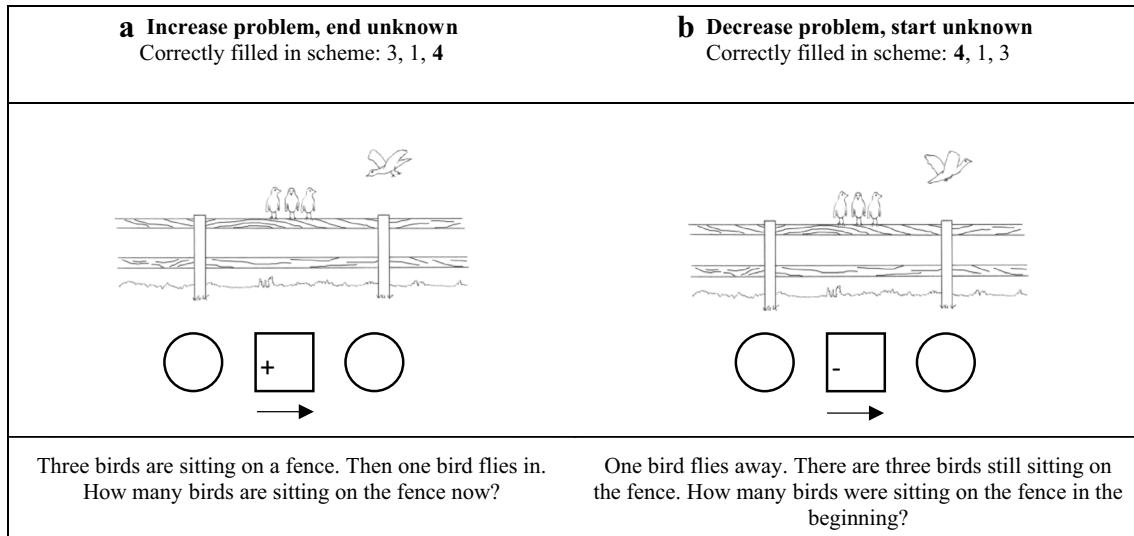2 Mathematics Education Centre, Loughborough University, Loughborough LE11 3TU, UK

Ⓐ Springer

a Increase problem, end unknown
Correctly filled in scheme: 3, 1, **4**

b Decrease problem, start unknown
Correctly filled in scheme: **4**, 1, 3

Three birds are sitting on a fence. Then one bird flies in. How many birds are sitting on the fence now?

One bird flies away. There are three birds still sitting on the fence. How many birds were sitting on the fence in the beginning?

Fig. 1 Examples of the experimental problems: a pictorial increase with an unknown end situation (**a**) and a decrease problem with an unknown start situation (**b**). The examples also contain the required numbers with the correct solution in bold. The text, which the experi- menter read to the child in the auditory support condition, is given below the dashed line (translation from Dutch). The child never saw this text—only heard it

in mathematics problems affect children's performance. In other words, we need to identify the characteristics of the optimal representation of mathematical problems, which could foster and enhance children's mathematics learning. The present study focused on a particular type of static pictorial mathematics problem used in Dutch curricula, which is used to depict increase and decrease situations (Fig. 1). Our overarching goal was to identify elements in these problems that impose difficulties on children's learning and performance, and optimal ways of alleviating this load.

2 Theoretical background

Problems can cause difficulties when they impose cognitive (over)load. *Cognitive load theory* (Sweller 2010) describes several types of cognitive load depending on the type of effort imposed on a learner. One of these is intrinsic load, which refers to the degree of element interactivity involved in a problem. Element interactivity can be seen as the degree to which the elements of a task are connected in a compli- cated way.

Certain types of pictorial mathematics problems involve high element interactivity and can therefore impose unneces- sary intrinsic cognitive load on young learners (Van Lieshout and Xenidou-Dervou 2018). Figure 1 depicts two types of pictorial problems that are often found in Dutch mathemat- ics curricula. One of them depicts an increase of an amount whereas the other one depicts a decrease. Each problem is accompanied by an empty number-sentence scheme, which the child has to fill in. In the increase problems the scheme

contains a plus sign, whereas in the decrease there is a minus sign. Van Lieshout and Xenidou-Dervou (2018) showed that the decrease problems were harder to solve than the increase problems even though both problems had to be solved with a simple addition.

This difference in difficulty was attributed to the higher cognitive load that the decrease problems imposed on the child. This conclusion derived from the finding that chil- dren's performance in the decrease problems improved when they were accompanied by the corresponding word prob- lem presented auditorily. The beneficial effect of combining pictorial and auditory information is commonly known as the "modality effect", which concerns the reduction of cog- nitive load when information is simultaneously processed in—and thus shared across—two different modalities (e.g., vision and hearing) instead of just one (Sweller et al. 1998). Combining the pictorial decrease problems with auditory support demonstrated a bigger modality effect, resulting in larger reduction in errors, than in the case of the increase problems. But the source of the cognitive load in this type of problems was as yet unclear. Specifically, which elements involved in these decrease problems lead to the increased intrinsic cognitive load?

To unravel the sources of cognitive load, it is important to clarify the elements involved: Looking at Fig. 1, it is clear that the increase situation shows an addend (the incoming birds) that is about to be joined with the augend (the birds on the fence). The elements of these static pictures trying to depict a dynamic situation correspond one-to-one with the computational actions needed to fill in the empty number- sentence scheme: $3 + 1 = 4$. In contrast, the picture of the

decrease situation shows a subtrahend (the outgoing bird) and the resulting difference of a subtraction (three birds). This dynamic change now does not directly correspond with the computational action needed to fill in the empty number-sentence scheme (again $3 + 1 = 4$) in the same direction as the pictorially suggested change (decrease). Instead, the child had to reconstruct the unknown minuend, which can cost more cognitive effort than solving the increase problem. In other words, the child has to mentally 'play back', the situation until the flying away bird is back on the fence. Meanwhile, the minus sign in the scheme below the decrease problem can be misleading; the child has to realise that an addition is needed, not a subtraction. Thus, the direction of change is clearly a potential source of the observed cognitive load.

Another source of cognitive load that arises from the aforementioned description is the position of the unknown. In the increase problems the unknown is at the end (the sum), whereas in the decrease it's at the start (the minuend). Past research has demonstrated that in bare problems, i.e., when there are no pictures or text (e.g., $3 + 2 = ?$ or $? + 2 = 5$), the problems with the unknown at the start are harder to solve than the ones with the unknown at the end (De Corte and Verschaffel 1981; Hiebert 1982). The same holds for word problems (De Corte and Verschaffel 1987; García et al. 2006; Hiebert 1982). But it is unclear whether this result holds for pictorial problems too.

Van Lieshout and Xenidou-Dervou's (2018) design did not allow them to pinpoint the source of the cognitive load, as the effect of the position of the unknown could not be separated from the effect of the direction of change. Naturally, these are intertwined in a static representation of a situation that is, in fact, dynamic: The entire flow of actions is compressed into only one snapshot, which leaves no room for more than two combinations of the position of the unknown and the change, i.e., either increase with an unknown end, or decrease with an unknown start.

To disentangle the effects of the direction of change from the position of the unknown, in the present study we replaced the static pictures with dynamic representations of the change in quantities. This dynamic representation included three pictures: (1) a picture of the start situation, (2) the change and, (3) the end situation. Combining the two possibilities for the location of the unknown (start vs. end) and the two possibilities for the direction of the change (increase vs. decrease) led to four conditions (see Fig. 2 for an example): (a) increase with an unknown end situation ($a + b = ?$), (b) decrease with an unknown end situation ($a - b = ?$), (c) increase with an unknown start situation ($? + a = b$) and, finally, (d) decrease with an unknown start situation ($? - a = b$). In these four conditions, the pictorial problem was presented in three sequential steps depicting the start situation, the change and the end situation one by one, thereby forming a kind of animated sequence. To make a distinction between the different problems that we used in our study, in what follows we call the original problems used by van Lieshout and Xenidou-Dervou (2018), with only one picture, 1-picture problems (Fig. 1); in contrast, the problems consisting of a sequence of three problems are called 3-picture problems (Fig. 2).
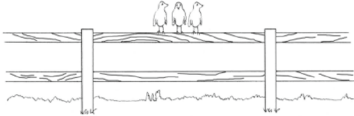
The 3-picture problems, besides being useful from the perspective of the aforementioned experimental design, could also constitute a condition that reduces the element interactivity of the problem, and thus the cognitive load imposed on young learners. It could be argued that if one wants young children to practice understanding and writing subtraction number sentences ($a - b = c$), it may perhaps be better not to use the relatively difficult 1-picture decrease problems ($x - a = b$) but rather semi-animated 3-picture decrease problems with the unknown at the end ($a - b = x$, Fig. 2b). In this problem the child can follow the consecutive steps of a real subtraction (i.e., the online separation) and map them onto the number-sentence scheme. Such 3-picture subtraction problems could reduce the intrinsic cognitive load by putting more explicit focus on the narrative of subsequent events. Research has shown that breaking up a task into successive parts can decrease its intrinsic load (de Jong 2010). Furthermore, multimedia research has shown that animations give rise to better understanding than static pictures (Höffler and Leutner 2007; Yung and Paas 2015).
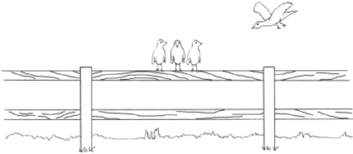
3 Research questions and hypotheses

The aims of the present study were three-fold. Firstly, we aimed to replicate van Lieshout and Xenidou-Dervou's (2018) findings regarding the 1-picture problems. Similarly to our goals in that study, we wanted to examine (a) if 1-picture decrease problems are harder than their increase counterparts, (b) if auditory support increases performance because of the modality effect and, (c) if the modality effect is more pronounced for the decrease rather than the increase problems.

Secondly, we addressed the following questions: Is it the location of the unknown and not, or to a lesser degree, the direction of the change or the modality (with or without auditory support) that which affects performance? Because of the aforementioned theoretical reasoning, we expected that performance in the 3-picture problems with an unknown start situation would be lower than in the case of 3-picture problems with an unknown end situation and auditory support. Although 1-picture decrease problems are harder to solve than 1-picture increase problems, we had no prediction concerning the effect of the direction of change in the case of the 3-picture problems. That is because in the case

a Increase – end unknown
Correctly filled in scheme: 3, 1, **4**
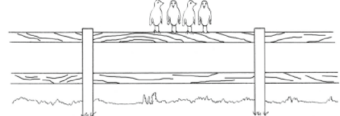
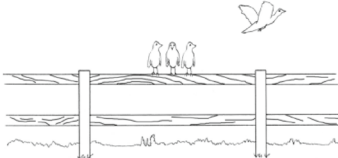Three birds are sitting on the fence.

Then one bird flies in.

How many birds are sitting on the fence now?

b Decrease – end unknown
Correctly filled in scheme: 4, 1, **3**

Four birds are sitting on the fence.

One bird flies away.

How many birds are sitting on the fence now?

c Increase – start unknown
Correctly filled in scheme: **3**, 1, 4

A number of birds are sitting on the fence.

Then one bird flies in.

You see four birds sitting on the fence. How many birds were sitting on the fence in the beginning?
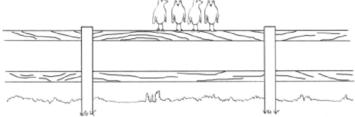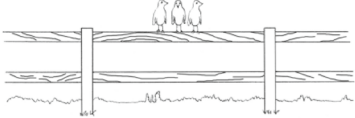
d Decrease – start unknown
Correctly filled in scheme: **4**, 1, 3

A number of birds are sitting on the fence.

One bird flies away.

You see three birds sitting on the fence. How many birds were sitting on the fence in the beginning?

◄**Fig. 2** Examples of the experimental problems of each of the four combinations of the direction of change (increase or decrease) and the location of the unknown (end or start): **a** increase, end unknown, **b** decrease, end unknown, **c** increase, start unknown, and **d** decrease, start unknown. The examples also contain the required number-sentence scheme with the correct solution in bold and the text in the case of the auditory support condition (translation from Dutch). The child never saw this text only heard it in the auditory support condition. As in the 1-picture problems, the figure contained empty number-sentences schemes (not shown here). These empty number-sentence schemes were placed below each of the 12 pictures. On the computer screen, the real size of the pictures was the same as the pictures in Fig. 1

of 1-picture problems, the effect of the direction of change was confounded by the effect of the position of the unknown.

Lastly, we aimed at examining whether we could improve performance on the relatively ambiguous and difficult 1-picture decrease problems by presenting the problems in a form of animation (Fig. 2b) or by providing auditory support (modality effect). In the 1-picture static decrease problems (Fig. 1b), the children could find it difficult to understand the relationship between what is depicted in the dynamic problem and the three-term number scheme, because all information was presented simultaneously. Therefore, we hypothesized that the 3-picture sequential presentation with auditory support of the decrease problem with the unknown at the end would make it easier to map the numerical information given in the pictures on the empty number scheme, compared to the static 1-picture problems.

4 Method

4.1 Participants

Eighty-first-grade children (41 boys and 39 girls) from three Dutch rural primary schools participated in the study after their parents had given written informed consent. Two children with ADHD (as informed by their teachers) were not included in this group. The participating children had to score at least 75% on an addition and subtraction speed test (see Sect. 4.2.1), which all of them did, except for one child, who was thus excluded. The final number of children whose data were used was 79. The mean ages in years of the boys ($n=40$) and girls ($n=39$) were respectively 6.64 ($SD=.30$) and 6.78 ($SD=.34$).

4.2 Material

4.2.1 Arithmetic speed tests

The present study focused on the children's ability to understand the meaning of mathematics problems depicted in the form of pictures; not their arithmetic skills. Thus, we aimed

at including participants who had good simple arithmetic skills. To achieve this aim, we selected our participants on the basis of their performance on an addition and a subtraction speed test. Each consisted of 60 problems and each problem consisted of a one one-digit number below 10 and an addend or subtrahend that was either 1 or 2. The sum of the addition problems ranged from 2 up to 10. The difference between minuend and subtrahend was at least 0 and at the most 7. The order of presentation of problems within each test was randomized. Both tests were presented in paper and pencil format. The children had to solve as many problems as possible within 1 min. On average, the children completed 13.73 ($SD=3.45$) addition problems, out of which 13.58 ($SD=3.43$) were solved correctly. Also, they completed on average 10.39 ($SD=3.02$) subtraction problems, out of which 9.65 ($SD=3.00$) were solved correctly. Only one participant was excluded based on our exclusion criterion.

4.2.2 Experimental problems

The experimental problems depicted either increase or decrease situations and were presented either as 1-picture or 3-picture problems. There were 16 one-picture and 32 three-picture problems. All problems were presented with or without auditory support. All experimental problems consisted of a picture or pictures representing birds sitting on or flying to or away from a fence. They also contained an empty number-sentence scheme below the picture (Fig. 1). All problems are available in (https://osf.io/e28tm/). The problems were modelled after the examples we encountered in Dutch practice booklets, which meant that we used the same elements, such as the square, circle, operation symbol and arrow presented in the same way as in those booklets.

In the case of the 1-picture problems these pictures represented either an increase situation with a known start set of a number of birds on the fence, a known increase and an unknown end situation (Fig. 1a), or a decrease situation with an unknown start set, a known decrease and a known set size in the end situation (Fig. 1b). Note that, as explained in the introduction, an increase problem always entailed an unknown end set whereas the decrease set always entailed an unknown start set. We used some of the pictures from van Lieshout and Xenidou-Dervou's (2018) study, which were based on Dutch mathematics curricula.

The 3-picture problems were a kind of animated version of the 1-picture problems. This condition separated the direction of the change, i.e., increase or decrease, from the position of the unknown, i.e., at the start or at the end, which leads to four conditions (Fig. 2). The child saw one picture at a time. All problems were accompanied with an empty number-sentence scheme such as the one depicted in Fig. 1. In the case of the two versions in which the unknown was at the start (Fig. 2c, d), the first picture could not give any

information to the child about what number it should fill in in the empty number-sentence scheme, conveying the fact that there was an unknown quantity in the number-sentence scheme. To be consistent, we did the same for the unknown quantity also when it was at the end of the number-sentence scheme.

In all the decrease problems, the number of birds that were left on the fence was always larger than the number of birds that flew away. The reverse situation (e.g., five birds fly away while two remain) could hint to the child that filling in the number-sentence scheme with, e.g., 2–5 would be incorrect. In these types of pictures, the child is expected to infer the direction of the change by the direction of the movement of the birds.

We used four triplets of numbers to construct the problems: (3, 4, 1), (5, 6, 1), (5, 7, 2) and (6, 8, 2). The last number of each triplet represented the change (either increase or decrease) in the problem. The triplet (3, 4, 1), for example, formed the addition $3 + 1 = 4$ or the subtraction $4 - 1 = 3$. Only the first or the second number in the triplet could be the unknown. The number triplets were pseudo randomly allocated to the conditions within a session. We used only these small addends and subtrahends because we were not interested in the arithmetical accuracy of the calculations themselves. Instead, we were interested in examining the children's understanding of the relationship between the quantities in the pictorial problem and the corresponding number-sentence scheme; not their arithmetic skills. Also, perhaps one could think that a realistic answer to our bird problem would consist of saying that when a bird flies to a group of sitting birds, one or more birds could perhaps fly away, or the landing bird could change its course. It was, however, not our intention to study the occurrence of such answers.

4.3 Procedure

The children participated in four sessions. Table 1 shows the complete experimental design. Half of the participants started with two sessions with 3-picture problems and then received two sessions with 1-picture problems. The other half received the reversed order. Within each of these two orders of presentation, half of the problems were accompanied with auditory support and half without support, and this alternated between the sessions. A session of 3-picture problems consisted of the four combinations (conditions) of the place of the unknown and the direction of the change. In order to counterbalance order effects, these four conditions, consisting of one problem each, were rotated four times in a Latin square design (e.g., ABCD—BCDA—CDAB—DABC) for each participant, which amounted to 16 problems per participant. The same kind of rotation design was used across all participants. This rotation meant that the next

participant started with BCDA and so on, until all conditions were rotated with each fourth participant. During the second session with 3-picture problems, the counterbalancing design was repeated, but the order of each set of four conditions for each participant was reversed (e.g., ABCD became DCBA) and rotated again. By reversing the orders, all pairs of successive conditions (for example AB and BA) occurred just as often in the complete counterbalancing design. In sum, the factors direction of change and location of the unknown were counterbalanced both in a between-subjects and a within-subjects design. The factor auditory support was counterbalanced only in a between-subjects design.

This counterbalancing design has a limitation: It cannot control for possible carry-over effects, e.g., that the participant learns something from being exposed to a certain condition and applies what has been learnt to the subsequent condition. Using a between-subjects design with a separate group of participants for each combination of factor levels would avoid this problem. We did not opt for this possibility because a within-subjects design is more economical in terms of the number of participants than a between-subjects design.

As mentioned above, half of the participants started with two sessions with 3-picture problems, whereas the other half started with two sessions with 1-picture problems in order to control for order effects of the sessions. Therefore, the whole design required a multiple of eight participants. The data of one participant in such a group of eight had to be excluded from the analyses as described in the Sect. 4.1. We decided not to exclude the data of the other seven participants because the counterbalancing design was quite robust in relation to the hypotheses and therefore such an exclusion would not lead to noticeably biased results.

The counterbalancing design for the 1-picture problems was simpler because there were only two conditions: increase versus decrease. The same within-participants counterbalancing design was used as in the case of the 3-picture problems. This amounted to four 1-picture problems per condition and eight problems per session.

Each session started with two practice trials in Power-Point during which the experimenter gave general information about the pictorial problems (birds that were flying to and away from the fence) and explained the relation between the picture(s) and the number-sentence scheme. The first practice trial concerned an increase problem and the second a decrease problem. Both practice problems contained the number triplet (2, 3, 1) (the last number being the size of the increase or decrease); this triplet was not used in the experimental problems. During this instructional stage of the practice trials the experimenter showed the child the correct number-sentence scheme and explained its relation to the picture(s). In the case of the 1-picture problems, by pressing the spacebar, the experimenter could show each stage of the

Table 1 Experimental design with counterbalancing for every eight participants

Participant	Session											
	1			2			3			4		
	Auditory support	Pictures	Order of conditions	Auditory support	Pictures	Order of conditions	Auditory support	Pictures	Order of conditions	Auditory support	Pictures	Order of conditions
1	No	3	ABCD BCDA CDAB DABC	Yes	3	DCBA CBAD BADC ADCB	No	1	DA AD AD AD	Yes	1	AD DA DA DA
2	No	3	BCDA CDAB DABC ABCD	Yes	3	CBAD BADC ADCB DCBA	No	1	AD AD AD DA	Yes	1	DA DA DA AD
3	Yes	3	CDAB DABC ABCD BCDA	No	3	BADC ADCB DCBA CBAD	Yes	1	AD AD DA AD	No	1	DA DA AD DA
4	Yes	3	DABC ABCD BCDA CDAB	No	3	ADCB DCBA CBAD BADC	Yes	1	AD DA AD AD	No	1	DA AD DA DA
5	No	1	AD DA AD AD	Yes	1	DA AD DA DA	No	3	DABC ABCD BCDA CDAB	Yes	3	ADCB DCBA CBAD BADC
6	No	1	AD AD DA AD	Yes	1	DA DA AD DA	No	3	CDAB DABC ABCD BCDA	Yes	3	BADC ADCB DCBA CBAD
7	Yes	1	AD AD AD DA	No	1	DA DA DA AD	Yes	3	BCDA CDAB DABC ABCD	No	3	CBAD BADC ADCB DCBA
8	Yes	1	DA AD AD AD	No	1	AD DA DA DA	Yes	3	ABCD BCDA CDAB DABC	No	3	DCBA CBAD BADC ADCB

The number in the column pictures refers to 3-pict or 1-pict problems. The meaning of the letters in the column of Order of Conditions is: *A* start unknown, decrease, *B* start unknown, increase, *C* end unknown, decrease; *D* end unknown, increase. Each letter represents one pictorial problem. As explained in the text, the conditions B and C were not possible for 1-picture problems. Spaces between groups of 2 or 4 letters are intended only to make it easier for the reader to see the structure of the series

development of the number-sentence scheme, i.e., the appearance of a number in the scheme, but the picture of the fence with the birds did not change. In the case of the 3-picture problems, also the picture itself changed. For example, during the first stage of an increase problem, the picture showed only the start set (e.g., two birds on the fence). By pressing the spacebar again, the number two appeared in the left most empty place of the number-sentence scheme. With the next press, a bird seemed to be heading towards the fence. The subsequent press showed the number one behind the plus sign in the second empty place. Then the next press highlighted the third empty place and the back of a head that seemed to hide the result of the addition. Next, the number three was filled in in the third empty place and finally the head disappeared and the three birds on the fence were visible.

The experimenter sat next to the child, both looking at the pictorial problem presented in PowerPoint on a computer screen. A cross in the middle of the screen indicated the start of a new problem. The presentation of the problems was self-paced: By pressing the spacebar, the child could start a new problem. In the case of the 3-picture problems, pressing the spacebar showed also the next one of the three pictures. As mentioned before, in the 3-picture conditions, the child could never see all three pictures at the same time on the screen. This method of self-paced presenting, i.e., that the participants decided themselves when they wanted to see the next picture, leads to higher performance than computer-paced animations, i.e., when the computer determines the appearance of the next picture (Leahy and Sweller 2011).

In the auditory support conditions, the experimenter gave oral support by reading out the corresponding word problems[1] during the presentation of a pictorial problem. This support consisted of describing to the child what kind of event the picture(s) represented (Figs. 1 and 2).

The children had to fill in the numbers in the right places on an answer sheet with the same empty number-sentence schemes as the ones presented on the computer screen. The right place meant that, depending on whether it concerned an addition or subtraction number-sentence scheme, the augend or minuend were filled in in the leftmost position, the addend or subtrahend in the square and the sum or the difference in the rightmost location.

4.4 Analysis

We used two outcome variables: The proportion of correctly solved unknowns (x) and the proportion of correctly filled in complete number-sentence schemes ($x \pm a = b$ or $a \pm b = x$).

For example, children got a score of 1 if they found correctly the unknown in a problem and 1 when they had also filled in correctly the entire number-sentence scheme. Mathematically, the order of the two addends in an increase problem is irrelevant. However, we wanted to see whether the child used the sequence of the events in the problem situation as the basis for filling in the number scheme from left to right. Therefore, a response was recorded as correct only if the child filled in the first circle the first addend, and the second in the box (Fig. 1). The children were free to choose the moment when they wanted to write down the numbers. So, they could do this during the presentation of the problem or afterwards.

Although one can often see the application of ANOVAs for such dichotomous responses, the comparison of mean proportions that differ in their distance to the extremes of the distribution (0 or 1) can invalidate the outcome of ANOVAs. The reason is that the variance of the scores is smaller the closer the mean is to one of the extremes. This violates the assumption of homogeneous variances, especially when the to be compared means are smaller than .3 or higher than .7 (Agresti 2002, p. 120). Indeed, several mean proportions in our study were larger than .7. An alternative way to analyse these proportions is to use a Generalized Estimating Equations (GEE) procedure (Jaeger 2008) with a logit link. Whereas the original proportion score varies by definition between 0 and 1, the transformed logit function allows the transformed score to vary between minus infinity and plus infinity (with zero as midpoint). All analyses concerned GEEs with repeated measures on all factors.

We used the odds ratio (OR) as our measure of effect size. We labelled these effect sizes according to the categories proposed by Cohen (1988) and Sawilowsky (2009) by first converting the OR-values into d-values[2] (Borenstein et al. 2009). The OR gives the ratio between the odds of one treatment condition compared to the odds of another. In order to calculate ORs that are not confounded by an interaction between the constituting factors, we reran the analyses without the interaction term when the interaction effect was not significant. In the case that the interaction was significant, we would have tested the simple effects and calculated the corresponding ORs. However, significant interactions did not occur.

5 Results

5.1 One-picture problems: replication

The first research question was intended to investigate whether we could replicate the findings of our previous study (van

---

[1] A mathematical word or story problem is a verbal description of a quantitative problem. The two texts in the lower part of Fig. 1 are examples of standard word problems.

---

[2] $d = \log_e OR \times \sqrt{3}/\pi$.

Lieshout and Xenidou-Dervou 2018). We found that 1-picture decrease problems with the unknown start position ("? − b = ?") were more difficult than 1-picture increase problem with an unknown at the end position ("a + b = ?"). For ease of reading, we shorten the name of the combination to just Direction. Furthermore, we found a positive effect of auditory support. We used a 2 (direction: decrease vs. increase) by 2 (auditory support: with vs. without) GEE analysis of the proportion of correctly filled in number-sentence schemes. Contrary to our expectations, there was no significant interaction effect, Wald $\chi^2(1, N = 79) = .16$, $p = .689$. As described in the analysis section, we then ran the analysis without the interaction term. Similarly to our previous findings (van Lieshout and Xenidou-Dervou 2018), the effects of the direction of change were significant with, as expected, a higher proportion correctly filled in number-sentence schemes in the increase problems than in the decrease problems, Wald $\chi^2(1, N = 79) = 36.15$, $p < .001$, $OR = 11.70$, 95% CI [5.25, 26.07], $M_{increase} = .99$ ($SE = .01$), $M_{decrease} = .86$ ($SE = .02$). Also, contrary to our prediction the effect of auditory support was not significant, Wald $\chi^2(1, N = 79) = 1.97$, $p = .160$, $OR = 1.45$, 95% CI [.86, 2.44], $M_{with} = .96$ ($SE = .01$), $M_{without} = .94$ ($SE = .01$). In terms of Cohen's (1988) d and Sawilowsky's (2009) effect size categories, the effects were, respectively, $d_{direction} = 1.36$ (very large), $d_{auditory\ support} = .21$ (small). As described in Sect. 4.4, we scored the filled-in number-sentence schemes as correct only when the position of the numbers from left to right corresponded with the course of events. If we had relaxed this criterion in the case of the increase problems, the significant performance difference between the increase and decrease problems would perhaps have been bigger, but without a change in the statistical significance. Thus, in the present study we could replicate van Lieshout and Xenidou-Dervou's (2018) finding regarding the decrease problems being more difficult than the increase problems. However, we could not replicate the finding regarding the auditory support being helpful and that it would especially be more helpful for the decrease problems with the unknown in the start position than for to the increase problem with the unknown at the end.

5.2 Three-picture problems: Finding the source of cognitive load

The second research question addressed the issue of whether the direction of the change of the set or the position of the unknown could explain the findings concerning the 1-picture problems. In contrast to the earlier reported analyses with the 1-picture problems, in this analysis the effect of the factor direction of the change and position of the unknown could be estimated independently from each other due to the 3-picture experimental design. We give an overview of the means and standard errors in Table 2. To analyse the effects on the accuracy of filling-in number-sentence schemes with

Table 2 Means, standard errors of the factors location of the unknown (end vs. start), direction of the change (increase vs. decrease), and oral support (with vs. without)

	End		Start	
	M	SE	M	SE
With auditory support				
Increase	1.00	.00	.84	.03
Decrease	.98	.01	.74	.04
Without auditory support				
Increase	.98	.01	.73	.04
Decrease	.98	.01	.62	.04

3-picture problems, we ran a 2 (Auditory Support: with vs. without) × 2 (Position of the Unknown: end vs. start) × 2 (Size Change direction: increase versus decrease) GEE. However, it turned out that this analysis design was not feasible due to a singular Hessian matrix, caused by a zero SE (and the maximum score of 1) in one of the factor combinations (Table 2). Because this analysis was not feasible, we decided to remove the three-way interaction of the analysis design. Removing this interaction did not limit the possibility of testing our predictions because we had no hypothesis concerning this particular interaction.

The adapted analysis demonstrated that none of the two-way interaction effects were significant, Position of the Unknown by Direction of the Change, Wald $\chi^2(1, N = 79) = .49$, $p = .484$; Position of the Unknown by Oral Support, Wald $\chi^2(1, N = 79) = .18$, $p = .892$; and lastly, Auditory Support by Direction of the Change, Wald $\chi^2(1, N = 79) = 1.40$, $p = .237$. To calculate the odds ratios of the main effects, we ran an analysis in which also the two-way interaction terms were removed. The main effect of the position of the unknown had a significant effect, Wald $\chi^2(1, N = 79) = 156.95$, $p < .001$, $OR = 27.72$, 95% CI [16.48, 46.61]. The mean proportions of accurately completed number-sentence schemes were, respectively $M_{end} = .99$ ($SE = .00$), $M_{begin} = .74$ ($SE = .03$) and as predicted they were in favour of the end position of the unknown. The main effect of auditory support appeared to be significant too, Wald $\chi^2(1, N = 79) = 13.23$, $p < .001$, $OR = 1.80$, 95% CI [1.31, 2.46], $M_{with} = .95$ ($SE = .01$), $M_{without} = .92$ ($SE = .01$). Finally, also the main effect of the direction of the change, Wald $\chi^2(1, N = 79) = 11.12$, $p = .001$, $OR = 1.80$, 95% CI [1.27, 2.53], $M_{increase} = .95$ ($SE = .01$), $M_{decrease} = .92$ ($SE = .01$) was significant. In sum, all main effects were significant, whereby the strongest effect *(OR)* occurred with the position of the unknown, followed by the smaller effect sizes of auditory support and the direction of the size change. In terms of Cohen's (1988) d and Sawilowsky's (2009) categories, the effects were, respectively, $d_{position\ unknown} = 1.83$ (very large), $d_{auditory\ support} = .32$ (small) and $d_{direction\ of\ the\ change} = .32$ (small).

Note that in the two 3-picture problem analyses, auditory support did lead to higher performance, whereas we could not show this effect in the analysis with the 1-picture problems.

5.3 Improving performance in the decrease problems: one vs. three-picture problems

Finally, the third research question concerned whether decrease problems with three sequentially presented pictures (the 3-picture problems) and an unknown end set (Fig. 2b) would make it easier for the children to fill in the empty number-sentence scheme than in the case of 1-picture decrease problems, which necessarily have at the same time the unknown in the start position. We compared the effects of both factors within a 2 (Number of Pictures: three versus one) $\times$ 2 (Auditory Support: with vs. without) GEE analysis. The Number of Pictures by 2 Auditory Support interaction did not show a significant effect, Wald $\chi^2(1, N = 79) = .24$, $p = .623$. We again repeated the analysis without interaction term. As expected, the children completed more number-sentence schemes correctly while solving 3-picture decrease problems with the unknown at the end than in the case of 1-picture decrease problems, Wald $\chi^2(1, N = 79) = 39.68$, $p < .001$, $OR = 8.71$, 95% CI [4.47, 17.08], (very large effect), $M_{\text{3-pict/end}} = .98$ ($SE = .01$), $M_{\text{1-pict/begin}} = .86$ ($SE = .02$). The factor auditory support had no significant effect on the performance this time, Wald $\chi^2(1, N = 79) = 1.07$, $p = .301$, $OR = 1.34$, 95% CI [.77, 2.32], (small effect), $M_{\text{with}} = .95$ ($SE = .01$), $M_{\text{without}} = .94$ ($SE = .01$). In terms of Cohen's (1988) d and Sawilowsky's (2009) categories, the effects were, respectively, $d_{\text{no of pictures}} = 1.19$ (large) and $d_{\text{auditory support}} = .16$ (very small).

As mentioned earlier, we had two outcome variables. We analysed the proportion of correct solutions of the unknown in the same way as the proportion correctly filled in number-sentence schemes. The scores on the former variable were somewhat higher than the scores on the latter. That is because a correctly filled in number-sentence scheme meant that all numbers had to be filled in correctly but also in the right place; thus, this variable was more stringent compared to just reporting the unknown correctly. Importantly, analyses with both outcome variables demonstrated similar results.

6 Discussion

Developers of mathematics curricula have been constructing realistic pictorial problems in the hope of overcoming difficulties occurring from abstract word problems (e.g.,

Hoogland et al. 2018). However, research has suggested that such pictorial problems can also sometimes carry their own shortcomings (Berends and van Lieshout 2009; Dewolf et al. 2015, 2016; van Lieshout and Xenidou-Dervou 2018). Van Lieshout and Xenidou-Dervou (2018) demonstrated that simple decrease pictorial problems impose high cognitive load on primary school children. The purpose of the present study was to replicate and extend this study by uncovering the sources of this cognitive load and exploring ways of reducing it.

Firstly, our results partially replicated past findings. As expected, writing a number-sentence scheme as a representation for pictorial decrease problems was harder than for increase problems (van Lieshout and Xenidou-Dervou 2018). However, in the present study auditory support did not improve children's performance in the difficult decrease problems. Secondly, we found that, as hypothesised, the source of the evident difficulty in the pictorial decrease problems appeared to be the position of the unknown: The unknown at the start of the depicted situation made the child's task more difficult than in the case of the unknown at the end of the situation. Evidently, the position of the unknown had a larger impact compared to the direction of the depicted change (increase versus decrease). Finally, we showed that sequentially presented pictures of a decrease problem with the unknown at the end position, made it easier for the child to fill in the corresponding number-sentence scheme compared to the traditional static representation of the decrease problem in just a single picture. In the next paragraphs we elaborate on these findings on the basis of each research question.

Our first research question was whether we were able to replicate key past findings in this domain (van Lieshout and Xenidou-Dervou 2018). The focus of our previous study (van Lieshout and Xenidou-Dervou 2018) was on a specific type of pictorial problems that represented in a static picture a dynamic situation where a change of an amount takes place. These problems can be found in common Dutch exercise and testing booklets. However, it is not easy to clearly represent a dynamic subtraction situation in a single static representation, because it is difficult to simultaneously picture the start situation (the minuend) as well as the decrease (the subtrahend). Ekeblad (1993) also recognized this representational difficulty in the case of subtraction. If one shows the decrease (e.g., birds flying away from a group of birds), then the only option left is to show the result of the subtraction (Fig. 1b). This means that the child has to counterintuitively add the subtrahend and the difference in order to reconstruct the minuend. Because an addition is needed to solve such problems, van Lieshout and Xenidou-Dervou (2018) did not use the term "subtraction problems" but instead named them "decrease problems" to avoid giving the impression

that these problems should be solved with a subtraction. Following the reasoning above, the authors expected and found that the decrease problems were harder than the increase problems for 7-year-old children. The difficulty of the decrease problem is probably aggravated by the presence of an empty number-sentence scheme with a minus sign, even though performing a subtraction would be incorrect. Our present study replicated van Lieshout and Xenidou-Dervou's (2018) original finding: a different group of 7-year-old children performed worse in the decrease problems compared to the increase ones (Fig. 1).

We also tried to replicate the finding that auditory support in these difficult decrease problems improves performance by causing a modality effect, which led van Lieshout and Xenidou-Dervou (2018) to conclude that the decrease problems impose higher cognitive load than the increase ones. Their conclusion was based on both accuracy and reaction time. The present study used only accuracy as dependent variable. We were not able to replicate the effect on this variable, most probably because of the ceiling effects evidenced in both the increase as well as the decrease problems. It would have been optimal if we had also collected RT data. Moreover, van Lieshout and Xenidou-Dervou's (2018) study included 36 one-picture problems, whereas the present study included only 16 one-picture problems due to its design. Psychometrically, a larger test length increases reliability (Gulliksen 1950) and therefore the chance of finding a significant result. Additionally, the number of participants may have been too small to confirm the earlier findings of van Lieshout and Xenidou-Dervou, although they were able to demonstrate this effect with 60 instead of the present study's sample of 80 children.

The second research question concerned locating the source of the cognitive load imposed by the 1-picture decrease problems: Was it the direction of the change (increase or decrease) or the location of the unknown (start vs. end)? To address this question, we developed 3-picture problems where we could vary the place of the unknown and the direction of the change. This resulted in four types of problems where the three pictures were presented sequentially (see Fig. 2a–d). Our results showed that the position of the unknown was the most influential factor. When the unknown was in the first position, the child's task of filling in the number-sentence scheme was clearly more difficult than when the unknown was in the last position. Therefore, we can conclude that the 1-picture decrease problems are more difficult than the 1-picture increase problems primarily because of the location of the unknown. So, our study demonstrated once again that 1-picture decrease problems with their noncanonical mathematical structure ($x - a = b$) are more difficult than their increase counterparts ($a + b = x$). Therefore, our findings verify and extend the literature on the importance of the position of the unknown. Past research has

demonstrated that bare or word problems with the unknown at the start are harder than corresponding problems with the unknown at the end (De Corte and Verschaffel 1981, 1987; García et al. 2006; Hiebert 1982). Our results show that the same occurs in pictorial problems. The position of the unknown appears to play a fundamental role irrespective of the mode of presentation of the problem.

The importance of the position of the unknown in arithmetic problems reflects the well-known fact that children's arithmetic errors—especially in Western countries—are often attributed to their tendency to think of arithmetic problems in an operational manner, i.e., a prompt to carry out the operation and fill in the unknown result at the end. That is because arithmetic problems are mostly taught in the form of $a + b = c$ (McNeil and Alibali 2005), which leads to a superficial understanding of the meaning of the equal sign: Rather than viewing it as a relational symbol, i.e., that both sides of the equal sign should be equal, students mostly see it as an operational symbol, i.e., an impetus to carry out an operation with the result at the right side of the equation (e.g., Jones et al. 2012; Powell et al. 2016; Simsek et al. 2019). This literature highlights the importance of teaching the relational meaning of the equal sign (e.g., Chesney et al. 2018). Our problems did not entail the equal sign per se, nevertheless we evidenced children's preference for viewing also the pictorial problems in an operational manner: problems were easier when the unknown was at the end.

One could think that perhaps the developers of the pictorial decrease problems such as the one depicted in Fig. 1b aimed at devising pre-algebra exercises in which the child would realise the need to mentally rearrange the equation to be able to fill in a number-sentence scheme where the unknown is at the start. This is a meaningful aim. However, many children have difficulty understanding this rearrangement because unfortunately most material that they encounter has the unknown at the end of the arithmetic problems. As described earlier, the fact that we found that pictorial problems with the unknown at the start are more difficult than the ones with the unknown at the end, is in line with the assumption that children are driven by an operational conception of the equal sign, which negatively affects their algebraic performance (McNeil et al. 2010). However, it should be noted we nevertheless believe that arithmetic problems with the unknown at the start should still be used in mathematics curricula. Material should be challenging for meaningful learning to take place. Thus, arithmetic problems should not be presented only in the (pictorial) form of '$a \pm b = c$' because that promotes operational thinking (Chesney et al. 2018 add also McNeil and Alibali 2005).

Our third research question concerned whether children's performance in the decrease situation would improve in the semi-animated 3-picture presentation with the unknown at the end compared to the 1-picture static condition. As argued

earlier, such a condition would be more realistic than the static 1-picture problems, which cannot really depict the sequence of events. As expected, the 3-picture condition led to a higher proportion of correctly filled in number schemes. The fact that the 3-picture problems included a more explicit focus on the narrative of the subsequent events could have lowered the learners' intrinsic cognitive load, i.e., the level of difficulty derived from the element interactivity included in the problem (Sweller 2010; Sweller et al. 1998).

Notably, the 3-picture problems had the unknown at the end; contrary to the 1-picture decrease problems. In light of our findings for our second research question, it is highly possible that the reason why the 3-picture problems were easier was also because of the position of the unknown. The position of the unknown (start or end) and the format (1-picture and 3-picture) are in fact confounded in this case and this should be taken into account when considering our present findings. Future research should separate and differentiate the potential effect of these factors. Lastly, we are not saying that the kind of 1-picture pictorial problems that we studied should not be used in mathematics education. Rather, we think that these problems should be used in the light of the goals that the developer has and the phase of mathematical development in which the children are.

It should also be noted that in the present study we found a positive effect of auditory support for the 3-picture decrease problems, but not for the 1-picture problems, contrary to our previous findings (van Lieshout and Xenidou-Dervou 2018). The positive effect of auditory support could be attributed to the modality effect: splitting the information across the two modalities—vision and hearing—reduced the cognitive load. However, the fact that this was observed only for the 3-picture problems and not the 1-picture problems could have been due to the fact that in our study there were double the number of problems in the 3-picture compared to the 1-picture condition. Future research should address this limitation.

6.1 Concluding remarks

Studies concerning the role of pictorial problems in the lower grades of primary school are quite scarce. As we put forward in the introduction, we need theory-driven fine-grained experiments to find out more about which characteristics of pictorial problems are profitable and which are not. For example, recently Bennet and colleagues (2019) showed that teaching children in the lower primary school grades to connect representations of numerosities to number symbols was more successful when the elements of the set were abstract (e.g., dots) than when they were represented with concrete objects (e.g., drawings of fish). The authors discuss that perhaps the use of abstract objects reduces the

working memory load involved. More research is rendered necessary to examine whether children have difficulties in solving pictorial mathematics problems because of (age-related) cognitive limitations or because of the used education material, for example the one-picture problems in our study, which are used for children who have not yet had the relevant experience or the instruction to understand the tasks. This consideration would also be important for developers of educational material. The present study's findings suggest that using a semi-animated presentation of realistic mathematics problems could be a useful tool for overcoming difficulties inherent in presenting dynamic situations in static pictures. One may argue that three-step semi-animated problems such as the ones depicted in Fig. 2 cannot be used in paper and pencil mathematics curricula. Perhaps one could consider a sequential illustration in comic-strip fashion (Elia et al. 2007; although see Berends and Van Lieshout 2009). Also, in this day and age where even young children use smartphones, computers, and tablets even for educational purposes, animated mathematics problems should be feasible and could provide a good platform for improving children's early arithmetic learning.

Compliance with ethical standards

References

Agresti, A. (2002). *Categorial data analysis* (2nd ed.). Hoboken: Wiley.

Bennett, A., Inglis, M., & Gilmore, C. (2019). The cost of multiple representations: Learning number symbols with abstract and concrete representations. *Journal of Educational Psychology, 111*(5), 847–860.

Berends, I. E., & van Lieshout, E. C. D. M. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction*. https://doi.org/10.1016/j.learninstruc.2008.06.012.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis* (2nd ed.). NJ: Hoboken.

Chesney, D. L., McNeil, N. M., Petersen, L. A., & Dunwiddie, A. E. (2018). Arithmetic practice that includes relational words

promotes understanding of symbolic equations. *Learning and Individual Differences, 64,* 104–112. https://doi.org/10.1016/j.lindif.2018.04.013.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, vol. 91* (2nd ed.). New York: Erlbaum.

De Corte, E., & Verschaffel, L. (1981). Children's solution processes in elementary arithmetic problems: Analysis and improvement. *Journal of Educational Psychology, 73*(6), 765–779. https://doi.org/10.1037/0022-0663.73.6.765.

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. *Journal for Research in Mathematics Education.* https://doi.org/10.2307/749085.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science, 38,* 105–134. https://doi.org/10.1007/s11251-009-9110-0.

Dewolf, T., Van Dooren, W., Hermens, F., & Verschaffel, L. (2015). Do students attend to representational illustrations of non-standard mathematical word problems, and if so, how helpful are they? *Instructional Science, 43*(1), 147–171. https://doi.org/10.1007/s11251-014-9332-7.

Dewolf, T., Van Dooren, W., & Verschaffel, L. (2016). Can visual aids in representational illustrations help pupils to solve mathematical word problems more realistically? *European Journal of Psychology of Education.* https://doi.org/10.1007/s10212-016-0308-7.

Ekeblad, E. (1993). Arithmetic without reading: Children interpreting representations of part-whole relations. *Computers in Human Behavior, 9*(2–3), 213–225. https://doi.org/10.1016/0747-5632(93)90008-G.

Elia, I., Gagatsis, A., & Demetriou, A. (2007). The effects of different modes of representation on the solution of one-step additive problems. *Learning and Instruction, 17,* 658–672. https://doi.org/10.1016/j.learninstruc.2007.09.011.

García, Ana, Jiménez, I. J. E., & Hess, S. (2006). Solving arithmetic word problems: An analysis of classifications as a function of difficulty in children with and without arithmetic LD. *Journal of Learning Disabilities, 39*(3), 270–281. https://doi.org/10.1177/00222194060390030601.

Gravermeijer, K. (1997). Commentary on solving word problems: A case study of modeling? *Learning and Instruction, 7*(4), 389–397.

Greer, B. (1997). Modelling reality in mathematics classrooms: The case of word problems. *Learning and Instruction, 7*(4), 293–307. https://doi.org/10.1016/S0959-4752(97)00006-6.

Gulliksen, H. (1950). *Theory of mental tests*. Hoboken: Wiley.

Hiebert, J. (1982). The position of the unknown set and children's solutions of verbal arithmetic problems. *Journal for Research in Mathematics Education, 13*(5), 341–349.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*(6), 722–738. https://doi.org/10.1016/j.learninstruc.2007.09.013.

Hoogland, K., de Koning, J., Bakker, A., Pepin, B. E. U., & Gravemeijer, K. (2018). Changing representation in contextual mathematical problems from descriptive to depictive: The effect on students' performance. *Studies in Educational Evaluation, 58,* 122–131. https://doi.org/10.1016/j.stueduc.2018.06.004.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007.

Jones, I., Inglis, M., Gilmore, C., & Dowens, M. (2012). Substitution and sameness: Two components of a relational conception of the equals sign. *Journal of Experimental Child Psychology, 113*(1), 166–176. https://doi.org/10.1016/j.jecp.2012.05.003.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*(6), 943–951. https://doi.org/10.1002/acp.1787.

Mcneil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*(4), 883–899.

McNeil, N. M., Rittle-Johnson, B., Hattikudur, S., & Petersen, L. A. (2010). Continuity in representation between children and adults: Arithmetic knowledge hinders undergraduates' algebraic problem solving. *Journal of Cognition and Development, 11*(4), 437–457. https://doi.org/10.1080/15248372.2010.516421.

Powell, S. R., Kearns, D. M., & Driver, M. K. (2016). Exploring the connection between arithmetic and prealgebraic reasoning at first and second grade. *Journal of Educational Psychology, 108*(7), 943–959. https://doi.org/10.1037/edu0000112.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8*(2), 597–599. https://doi.org/10.22237/jmasm/1257035100.

Simsek, E., Xenidou-dervou, I., Karadeniz, I., & Jones, I. (2019). The conception of substitution of the equals sign plays a unique role in students' algebra performance. *Journal of Numerical Cognition, 5*(1), 24–37.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychological Review, 22,* 123–138. https://doi.org/10.1007/s10648-010-9128-5.

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

van Lieshout, E. C. D. M., & Xenidou-Dervou, I. (2018). Pictorial representations of simple arithmetic problems are not always helpful: a cognitive load perspective. *Educational Studies in Mathematics, 98*(1), 39–55. https://doi.org/10.1007/s10649-017-9802-3.

Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction, 4*(4), 273–294. https://doi.org/10.1016/0959-4752(94)90002-7.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Lisse: Swets & Zeitlinger.

Yung, H. I., & Paas, F. (2015). Effects of computer-based visual representation on mathematics learning and cognitive load. *Educational Technology and Society, 18*(4), 70–77.