Article:

Xu, C. orcid.org/0000-0002-6702-3958, Li, H. orcid.org/0000-0002-5052-2742, Di Lonardo Burr, S. orcid.org/0000-0001-6338-9621 et al. (3 more authors) (2024) We cannot ignore the signs: the development of equivalence and arithmetic for students from grades 3 to 4. Journal of Cognition and Development, 25 (1). pp. 46-65. ISSN 1524-8372

https://doi.org/10.1080/15248372.2023.2245507

Journal of Cognition and Development

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hjcd20

We Cannot Ignore the Signs: The Development of Equivalence and Arithmetic for Students from Grades 3 to 4

Chang Xu, Hongxia Li, Sabrina Di Lonardo Burr, Jiwei Si, Jo-Anne LeFevre & Xinfeng Zhuo

To cite this article: Chang Xu, Hongxia Li, Sabrina Di Lonardo Burr, Jiwei Si, Jo-Anne LeFevre & Xinfeng Zhuo (10 Aug 2023): We Cannot Ignore the Signs: The Development of Equivalence and Arithmetic for Students from Grades 3 to 4, Journal of Cognition and Development, DOI: 10.1080/15248372.2023.2245507

To link to this article: https://doi.org/10.1080/15248372.2023.2245507

JOURNAL OF COGNITION AND DEVELOPMENT
https://doi.org/10.1080/15248372.2023.2245507

We Cannot Ignore the Signs: The Development of Equivalence and Arithmetic for Students from Grades 3 to 4

Chang Xu[b,c], Hongxia Li[a], Sabrina Di Lonardo Burr[d], Jiwei Si[a], Jo-Anne LeFevre[c] and Xinfeng Zhuo[a]

[a]Shandong Normal University, China; [b]Queen's University Belfast, UK; [c]Carleton University, Canada; [d]University of British Columbia, Canada

ABSTRACT

Students' understanding of the meaning of the equal sign develops slowly over the primary grades. In addition to updating their representations of equations to recognize that the equal sign represents an equivalence relation rather than signaling an operation, students need to move beyond full computation to efficiently solve equivalence problems. In this study, we examined the longitudinal relation between arithmetic and equivalence for students who were capable of accurately solving arithmetic problems in different formats. Chinese students ($N = 612$; $M_{age} = 9.0$ years in Grade 3, 57% boys) completed measures of arithmetic fluency and equivalence fluency in Grade 3 and again in Grade 4. They also completed a non-verbal reasoning task in Grade 3. We tested a cross-lagged structural equation model to examine the reciprocal relations between arithmetic and equivalence fluency. We found reciprocal relations between the development of arithmetic and equivalence fluency from Grades 3 to 4, with a greater influence of arithmetic on the development of equivalence than the reverse. Furthermore, non-verbal reasoning predicted the development of equivalence, but not the development of arithmetic. Based on our findings, we conclude that for Chinese students with prior basic understanding of equivalence, flexible access to arithmetic facts supports their development of equivalence fluency.

Equivalence is the concept that the quantities represented by expressions on both sides of the equal sign must be the same (Alibali, 1999; Powell, 2012; Sherman & Bisanz, 2009). Although an understanding of equivalence is fundamental to mathematics, acquisition of this concept is difficult for many elementary school students (Carpenter & Levi, 2000; Kieran, 1981; McNeil & Alibali, 2005; Sherman & Bisanz, 2009; Yang, Huo, & Yan, 2014). In particular, during the early stages of learning arithmetic, a common misconception is that the equal sign is an operational symbol that means "the answer always [comes] right after the equal sign" (Carpenter, Franke, & Levi, 2003, p. 11). Once students accept that the equal sign represents an equivalence relation rather than signaling an operation, they can develop efficient strategies to solve a variety of

CONTACT Chang Xu c.xu@qub.ac.uk School of Psychology, Queen's University Belfast, David Keir Building, 18-30 Malone Road, Belfast BT9 5BN, Northern Ireland; Hongxia, Li lihongxia@sdnu.edu.cn School of Psychology, Shandong Normal University, No.1 University Road, Science Park, Changqing District, Jinan, Shandong 250358, China

This article has been republished with minor changes. These changes do not impact the academic content of the article.

arithmetic problems without relying on full computation (e.g., 1036 + 4 + __ = 1036 + 12; Kindrat & Osana, 2018). We refer to students' ability to solve equivalence problems both efficiently (i.e., quickly and accurately) and flexibly (i.e., applied to equations in different formats) as *equivalence fluency*. Equivalence fluency becomes particularly important when students subsequently learn algebra (Fyfe, Matthews, Amsel, McEldoon, & McNeil, 2018; Kindrat & Osana, 2018; Knuth, Stephens, McNeil, & Alibali, 2006; Matthews & Fuchs, 2020).

Most research on equivalence is focused on students' acquisition of conceptual knowledge and does not assess the fluency with which they use this knowledge to enhance mathematical performance. In contrast, in the present research we examined the reciprocal relations between arithmetic fluency and equivalence fluency. Arithmetic fluency was operationalized as the ability to efficiently solve standard arithmetic problems, in which operations only appear to the left of the equal sign (i.e., a + b = __), whereas equivalence fluency was operationalized as the ability to efficiently solve nonstandard arithmetic problems, in which operations appear on either both sides or to the right of the equal sign (e.g., a = b − __; a + __ = b; a − b = __ + c; Carpenter, Franke, & Levi, 2003; Powell, 2012; Sherman & Bisanz, 2009). We focussed on Chinese students because the Chinese mathematics curriculum introduces the equal sign in relational contexts (i.e., the equal sign means "sameness" and "balance") along with the introduction of addition and subtraction from Grade 1 onwards (Capraro, Ding, Matteson, Capraro, & Li, 2007; Li, Ding, Capraro, & Capraro, 2008). Given the close relation between arithmetic and equivalence, and the simultaneous introduction of and ongoing practice with these concepts for Chinese-educated students, we anticipated that improvements in one would lead to improvements in the other.

The role of equivalence in arithmetic equations

Knowledge of mathematical equivalence is fundamental to students' development of arithmetic and algebraic skills (Alibali, 1999; Fyfe, Matthews, Amsel, McEldoon, & McNeil, 2018; Kieran, 1981; Matthews & Fuchs, 2020; McNeil, 2008; McNeil et al., 2012; McNeil, Hornburg, Devlin, Carrazza, & McKeever, 2019; Robinson, Price, & Demyen, 2018). In 2011, Rittle-Johnson and colleagues proposed a model that summarized the developmental progression of students' understanding of equivalence. According to their model, first, students hold a *rigid operational* view of the equal sign, seeing it as a signal to "do something". Next, students transition to a *flexible operational* view of the equal sign; they still mainly hold an operational view but they will also accept a small portion of atypical (nonstandard) equations that are compatible with the operational view of the equal sign (e.g., c = a + b or a = a). At the next level of understanding, students hold a *basic relational* view of the equal sign, implicitly understanding it as a relational symbol; however, they cannot provide conceptually accurate explanations for such problems. Finally, students develop a *comparative relational* view; they understand the equal sign as a relational symbol and can recognize and explain that performing the same operations on both sides maintains equivalence. These students can use compensatory strategies to solve challenging problems efficiently without needing to perform full computations. For example, given the problem "18 + 35 = 17 + __", they can solve it by reasoning that 18 is 1 more than 17, so the unknown must be 1 more than 35 (Rittle-Johnson, Matthews, Taylor, & McEldoon, 2011). On this view, comparative relational understanding of the equal sign supports students in decomposing and recomposing numbers flexibly when they solve arithmetic problems.

When students solve *standard* equations in which operations appear only on the left side of the equation, they do not need to understand the equal sign as a relational symbol, but only need to recognize it as an operational symbol (e.g., 4 + 9 = __; Sherman & Bisanz, 2009). However, only having an operational understanding of the equal sign may hinder success when students encounter more complex equations that are in *nonstandard* forms (Kieran, 1981; Knuth, Stephens, McNeil, & Alibali, 2006). For example, nonstandard equations may have operations on both sides, or operations on the right side (Carpenter, Franke, & Levi, 2003). Consider the example of 2 + 5 = __ + 1: Students who have only an operational understanding of the equal sign may misinterpret the equal sign as "the answer comes next," and proceed to add the numbers on the left side of the equation and ignore the numbers on the right side (i.e., 2 + 5 = 7). Alternatively, they may incorrectly restructure the equation by adding all three numbers on both sides of the equation (2 + 5 + 1 = 8). These misconceptions about the meaning of the equal sign may limit students' ability to solve nonstandard equations that require reasoning skills (Kindrat & Osana, 2018; Sherman & Bisanz, 2009).

While there are many studies on equivalence in Western countries (e.g., Alibali, 1999; Fyfe, Matthews, Amsel, McEldoon, & McNeil, 2018; Kieran, 1981; Kindrat & Osana, 2018; Knuth, Stephens, McNeil, & Alibali, 2006; McNeil, 2008; McNeil et al., 2012; McNeil, Hornburg, Devlin, Carrazza, & McKeever, 2019; Robinson, Price, & Demyen, 2018; Sherman & Bisanz, 2009), only three studies have examined equivalence knowledge for Chinese students (Jones, Inglis, Gilmore, & Dowens, 2012; Li, Ding, Capraro, & Capraro, 2008; Yang, Huo, & Yan, 2014). Yang, Huo, and Yan (2014) found that most Chinese students in Grades 3 to 5 made correct judgments for equivalent statements (e.g., true or false: 57 + 22 = 58 + 21). However, in Grade 3, most students used a computational strategy, adding numbers on the left (57 + 22 = 79) and right sides of the equation (58 + 21 = 79) to determine if the statement was true. In contrast, in Grades 4 and 5, many students used a compensatory strategy (e.g., 57 is 1 less than 58, and 22 is 1 more than 21, so the statement is true). Use of the compensatory strategy suggests that the older students were more likely to have a relational understanding of the equal sign. However, even in Grade 5, only about one-third of the students were able to state a relational interpretation of these statements (Yang, Huo, & Yan, 2014).

By Grade 6, Li, Ding, Capraro, and Capraro (2008) found that 98% of Chinese students, compared to only 28% of American students, correctly solved and provided conceptually accurate explanations for equivalence problems (e.g., 6 + 9 = __ + 4) and thus demonstrated both procedural and conceptual knowledge of equivalence. Similarly, Jones, Inglis, Gilmore, and Dowens (2012) found that 11- and 12-year-old Chinese students were much more likely to endorse relational interpretations of the equal sign than British students of the same age. Specifically, when asked to rate the "cleverness" of operational, sameness, and substitutive definitions of the equal sign, Chinese students gave higher cleverness ratings to the substitutive definitions, which required relational interpretations of the equal sign (e.g., the two sides can be exchanged) than British students. In contrast, British students gave higher cleverness ratings to the operational definitions (e.g., answer to the problem) than

Chinese students. In summary, these studies suggest that educational experiences dictate students' learning of equivalence, with Chinese-educated students gaining Level 4 knowledge much earlier than students educated in Western countries.

The development of procedural and conceptual understanding of mathematical equivalence is closely tied to the type of education students receive (Li, Ding, Capraro, & Capraro, 2008; Powell, 2012; Rittle-Johnson, Matthews, Taylor, & McEldoon, 2011). Capraro, Ding, Matteson, Capraro, and Li (2007) examined three of the most popular sets of Chinese mathematics textbooks and found that all textbooks introduced the equal sign in relational contexts, for example, students are taught to substitute the words "sameness" and "balance" for the equal sign prior to the introduction of addition and subtraction. In a comparative analysis of teacher guides and student texts in China and the United States, Li, Ding, Capraro, and Capraro (2008) found that Chinese textbooks provided much more direct instruction on equivalence and the relational meaning of the equal sign than American textbooks. More specifically, teachers use concrete pictorial representations and life situations to help students develop an understanding of "the same as" before it is introduced in the arithmetic context (Li, Ding, Capraro, & Capraro, 2008). Subsequently, Chinese students practice many nonstandard problems to enhance students' understanding of the equal sign, including operations without equal signs (e.g., solve 8 + 5), using an arrow to replace the equal sign (e.g., 14–5–2 → __), learning to place numbers in the blank to make simple equations true (e.g., __ = 9, 5 = __) and to fill in missing numbers in more complex nonstandard contexts (e.g., __ + 2 = __; __ + 2 = __). The latter problems have an infinite number of solutions, requiring both relational understanding of the equal sign and flexible understanding of equivalence. The comprehensive instruction on the relational view of the equal sign that Chinese students receive may help them develop a relational understanding of equivalence. Overall, evidence from Chinese curricula and the findings from previous studies indicate that conceptual understanding of equivalence develops slowly over time but it is very sensitive to the curriculum.

The integration of arithmetic and equivalence knowledge

Students' arithmetic skills develop in the early grades of elementary school in China (Ministry of Education, 2011). By Grade 3, Chinese students are expected to be able to fluently solve basic arithmetic problems. Over time, they use their fluency with arithmetic to quickly and accurately solve more complex arithmetic problems. Similarly, equivalence skills are also developing in Grade 3; however, full conceptual understanding does not occur for most students until the later grades of elementary school (Li, Ding, Capraro, & Capraro, 2008; Yang, Huo, & Yan, 2014). To date, studies have separately investigated the development of arithmetic and equivalence, however, to our knowledge none have explored the bidirectional development of these skills. Moreover, the equivalence literature has predominately focused on accuracy (e.g., Alibali, 1999; Fyfe, Matthews, Amsel, McEldoon, & McNeil, 2018; Li, Ding, Capraro, & Capraro, 2008; Mathews & Fuchs, 202 0; McNeil, 2008; Robinson, Price, & Demyen, 2018; Yang, Huo, & Yan, 2014) and thus the development of equivalence fluency in relation to arithmetic fluency remains unknown.

Students with strong equivalence and arithmetic fluency can use these skills in combination to efficiently solve more complex equivalence problems (Osana & Kindrat, 2021). For example, to determine whether 65 + 36 = 67 + 38, instead of relying on full computation, students with equivalence and arithmetic fluency can efficiently recognize that 67 is 2 more than 65 and 38 is 2 more than 36 so it is not possible for this equation to be true (Osana & Kindrat, 2021). Fluency in equivalence and arithmetic is also important when students are challenged by arithmetic problems that take on various formats, such as __ = 5 + 3 and 7 = _ + 3 (McNeil, Fyfe, & Dunwiddie, 2015). Furthermore, fluency allows students to solve more complex arithmetic problems by decomposing and recomposing numbers in a flexible manner (e.g., 29 + 149 = 30 + 150–2; Kindrat & Osana, 2018). Given that both types of fluency play a role in solving arithmetic and equivalence problems, considering their bidirectional development is critical.

The role of reasoning skills in solving arithmetic and equivalence problems

Beyond the relations among different mathematics skills, domain-general skills, such as reasoning, are essential in the development of mathematics (see reviews in Miller Singley & Bunge, 2014; Morsanyi, Prado, & Richland, 2018). Reasoning skills help students to notice relations among numbers and expressions, simplify calculations, extract rules to generate new information, and extend procedures to novel problems (Alexander, Jablansky, Singer, & Dumas, 2016; Jacobs, Franke, Carpenter, Levi, & Battey, 2007). With respect to equivalence, relational reasoning skills may allow students to coordinate quantities in a mathematical expression without full computation by transforming nonstandard problems into equivalent expressions (Kindrat & Osana, 2018). For example, with relational thinking, students can ignore 1938 in the equation 1938 + 5 + __ = 1938 + 11 to quickly determine that the answer is 6 because 5 + 6 = 11. Thus, reasoning skills may be particularly important for the development of equivalence fluency. For standard arithmetic, however, the process of acquiring basic number facts involves discovering, labeling, and internalizing number relations (Baroody, 1985). Development of fluent access to basic arithmetic operations reduces the role of relational thinking in solving arithmetic problems because students no longer need to infer new rules or search for novel number relations.

Current study

The goal of the present study was to examine the reciprocal relations between arithmetic fluency and equivalence fluency for Chinese students in Grade 3 and Grade 4. Consistent with previous studies (Carpenter, Franke, & Levi, 2003; Kieran, 1981; Knuth, Stephens, McNeil, & Alibali, 2006; Li, Ding, Capraro, & Capraro, 2008; McNeil, 2008), we used nonstandard arithmetic problems to measure equivalence. However, novel to the present study, we used speeded tasks to tap into equivalence fluency and arithmetic fluency. We asked two research questions: (a) How are equivalence fluency and arithmetic fluency related from Grade 3 to Grade 4? and (b) Do non-verbal reasoning skills predict individual differences in the development of arithmetic fluency and equivalence fluency?

Based on the assumption that the development of students' arithmetic and equivalence skills are closely related, we hypothesized that there would be reciprocal relations between the two skills between Grades 3 and 4 (Hypothesis 1). Second, based on the assumption that students need arithmetic knowledge to solve problems in nonstandard formats, we predicted that arithmetic would more strongly predict the change in equivalence fluency than equivalence would predict the change in arithmetic fluency (Hypothesis 2). Third, we expect that recognizing relations among numbers can help students choose efficient and flexible solutions to nonstandard problems. Thus, we hypothesized that non-verbal reasoning would predict the change in students' equivalence fluency but would not predict change in their arithmetic fluency (Hypothesis 3).

Method

Participants

Approval from the Institutional Review board at Shandong Normal University and the local school board was obtained, followed by written consent from parents or guardians. Monolingual Chinese students ($N = 612$; $M_{age} = 9.0$ years, $SD = .59$; 57% boys) were recruited near the end of the first semester in Grade 3. The sample was obtained from two public elementary schools (14 classrooms) from a suburban town with an economic level approximately at the national average (National Bureau of Statistics, 2019). One year later, in Grade 4, the same group of students were assessed on the same measures ($M_{age} = 9.8$ years, $SD = .58$). In the present study, we did not collect information about the socioeconomic status of the sample. However, to provide a description of the sample, we estimated parent education levels obtained from other cohorts from the two participating schools (Li et al., 2021). Education levels for parents of the students in these two schools typically ranged from elementary school to a postgraduate degree, with a median education level of a high school degree for both fathers and mothers, representative of low- to middle-socioeconomic status in China.

Procedure

In both Grades 3 and 4, students were assessed in group sessions in their classroom during school hours. Two trained experimenters were present for all testing, with one experimenter focused on reading directions and monitoring time, while the other experimenter circulated through the classroom with the teacher, ensuring that students were following the instructions for the task. In Grade 3, students completed the Raven's progressive matrices test in one 45-minute session. Also, in both Grade 3 and 4, students completed the Chinese-adapted version of the standardized Heidelberg Rechen Test (HRT; Haffner, Baro, & Resch, 2005; adapted from; Wu & Li, 2005), which included the arithmetic and equivalence measures, in a second 45-minute session. All students completed the tasks in the same order: addition, subtraction, multiplication, division, and equivalence. A brief break was provided between each task. In Grade 4, students also completed a number writing speed task at the beginning of the session as a control variable to account for individual differences in the speed with which students could write Arabic digits.

Data availability statement

To promote transparency and openness, anonymized data for the measures analyzed in the current paper and analysis code are freely available for download at [https://osf.io/t7fbj/].

Measures

Non-verbal reasoning

In Grade 3, students completed a paper-and-pencil version of the Raven's standard progressive matrices test (Raven, 1938) as a measure of non-verbal reasoning skills. Students were given 40 minutes to complete five sets of questions (12 questions within each set), with questions ordered by increasing difficulty. Each question consisted of a series of geometric figures with one element missing. Students were asked to find the missing element in a pattern among six options. Scoring was the total number of questions answered correctly (maximum = 60). In the present sample, the items showed good internal reliability, Cronbach's α = .86.

Mathematics tasks

Paper-and pencil tests from the adapted version of the standardized HRT (Haffner, Baro, & Resch, 2005; adapted from; Wu & Li, 2005) were administered. In the present study, we use data from the arithmetic and equivalence measures because they were directly relevant to our research questions. Notably, to ensure the task was appropriate for the population of this study, as recommended by Wu and Li (2005), students were given less time to complete the arithmetic measures and some of the digits, but not the format, were changed for later items. The equivalence measure used was identical to the one used in the original HRT.

Standard arithmetic

Students completed addition, subtraction, multiplication, and division subtests in which operations appear only on the left side of the equation. For each operation, 40 questions were presented in two columns on a single page, with questions ordered by increasing difficulty. Students were given one minute to answer as many questions as possible, in order, beginning with the left column. Scoring was the total number of questions answered correctly (maximum = 40). These subtests have excellent reported test-retest reliabilities based on a large national assessment of Chinese students in Grades 1 through 6 (addition = .89, subtraction = .86, multiplication = .98, division = .94; Wu & Li, 2005).

Addition

The left column consisted of problems with single- and double-digit addends (e.g., 2 + 5 = __, 12 + 8 = __) with no sums greater than 20. The right column consisted of problems with single-, double-, and triple-digit addends (e.g., 7 + 13 = __, 24 + 49 = __, 267 + 432 = __).

Subtraction

The left column consisted of problems with single- and double-digit minuends and subtrahends (e.g., 5–1 = __, 19–6 = __, 17–13 = __), with no minuends greater than 20. The right column consisted of problems with double- and triple-digit minuends, and single-, double- and triple-digit subtrahends (e.g., 17–9 = __, 32–15 = __, 130–28 = __, 732–421 = __).

Multiplication

The left column consisted of problems with single-digit multiplicands and multipliers (e.g., $4 \times 1 =$ __, $9 \times 8 =$ __). The right column consisted of problems with single- and double-digit multiplicands and multipliers, all less than 20 (e.g., $17 \times 11 =$ __, $18 \times 6 =$ __).

Division

The left column consisted of problems with single-and double-digit dividends and single-digit divisors (e.g., $8 \div 2 =$ __, $28 \div 7 =$ __). These problems were complementary to the problems found on a 9×9 multiplication table. The right column consisted of problems with double- and triple-digit dividends and single-digit divisors (e.g., $54 \div 9 =$ __, $450 \div 15 =$ __).

Equivalence

Forty nonstandard addition, subtraction, and mixed operation equivalence questions were presented in two columns on a single page, with questions ordered by increasing difficulty. Questions consisted of a mixture of equations with a missing operand on the right side of the equal sign (e.g., $5 = 8 -$ __; $11 =$ __ $+ 6$), equations with a missing operand on the left side of the equal sign (e.g., $6 +$ __ $= 7$; $12 -$ __ $= 5$), and equations with operations on both sides of the equal sign (e.g., $5–1 =$ __ $+ 2$; $53 + 12 =$ __ $- 30$). Students were given one minute to fill in the missing numbers to make the statements true as quickly as possible, in order, beginning with the left column. Scoring was the total number of equivalence questions answered correctly (maximum = 40). The reported test-retest reliability for the equivalence task from the Chinese HRT was .76 (Wu & Li, 2005).

Number writing speed

Given that the main tasks of interest in our study were all speeded based on paper and pencil tasks, students completed a number writing speed task in Grade 4 as a control variable for individual differences in the speed at which students can write down numbers. Students were given 30 seconds to copy Arabic digits in order as quickly as possible. A maximum of 60 numbers ranging from 0 to 9 were presented in three columns. The score was the total number of digits students copied.

Results

Descriptive statistics

Descriptive statistics and correlations among variables are shown in Tables 1 and 2. Violin plots (see Figure 1) show the distribution of the data in each Grade for each measure. Although all measures were normally distributed, the patterns of distribution varied, with the distribution for multiplication showing less variability than those for addition, subtraction, division, and equivalence. Although retrieval of basic facts is the focus of arithmetic instruction in China, the memorization of the multiplication table is unique, as students are required to orally recite half of the operations up to 9×9 (small-operand-first entries only). Furthermore, most of the items on the test were from this memorized set whereas addition and subtraction included more complex items. For these reasons, the overall variability in multiplication was smaller compared to the other operations.

Table 1. Descriptive statistics and comparisons for measures in grade 3 (time 1) and grade 4 (time 2).

	Grade 3						Grade 4						Comparisons		
	N	*M*	*SD*	Skew	Min	Max	*N*	*M*	*SD*	Skew	Min	Max	*t*	*df*	Cohen's *d*
Non-verbal Reasoning	612	40.54	7.17	−0.65	7	55	-	-	-	-	-	-	-	-	-
Addition	611	24.03	4.41	−0.05	2	40	569	26.32	4.21	−0.11	12	40	15.12	567	0.63
Subtraction	611	24.05	4.63	−0.12	2	40	570	25.18	4.67	−0.39	1	38	6.77	568	0.28
Multiplication	611	28.62	3.37	−1.94	3	40	568	30.87	2.83	−1.79	13	38	15.55	567	0.65
Division	611	22.49	6.07	−0.62	2	34	569	27.69	5.94	−1.03	1	39	26.30	568	1.10
Equivalence	612	20.16	5.62	−0.50	1	36	570	23.18	5.38	−0.72	2	35	15.88	569	0.67
Number Writing	-	-	-	-	-	-	570	44.41	10.76	−0.37	0	60	-	-	-

Maximum possible scores were 40 for the arithmetic measures and 60 for number writing speed. All pairwise *t*-tests comparing grade 3 to grade 4 performance were significant at $p < 0.001$.

Table 2. Correlations among measures in grade 3 (G3) and grade 4 (G4).

	Grade 3						Grade 4				
	1	2	3	4	5	6	7	8	9	10	11
1. Non-verbal Reasoning G3	-										
2. Addition G3	.27	-									
3. Subtraction G3	.22	.77	-								
4. Multiplication G3	.13	.47	.41	-							
5. Division G3	.24	.65	.64	.43	-						
6. Equivalence G3	.33	.49	.42	.38	.55	-					
7. Addition G4	.24	.67	.63	.36	.56	.49	-				
8. Subtraction G4	.29	.62	.66	.33	.55	.54	.71	-			
9. Multiplication G4	.12	.44	.37	.40	.40	.41	.51	.51	-		
10. Division G4	.21	.58	.56	.36	.70	.56	.67	.70	.55	-	
11. Equivalence G4	.40	.57	.53	.32	.53	.66	.63	.66	.47	.65	-
12. Number Writing G4	*.00*	.30	.32	.18	.21	*.07*	.22	.22	.14	.17	.09

Shaded regions indicate within-grade correlations. Nonsignificant correlations are italicized. Except for the correlations of .12 ($p = 0.004$) and .09 ($p = 0.036$), correlations were statistically significant at $p < 0.001$.

A few outliers (i.e., |z-scores| > 3.29), were found for the following tasks: Non-verbal reasoning ($n = 5$), addition in Grade 3 ($n = 2$) and Grade 4 ($n = 1$), subtraction in Grade 3 ($n = 1$) and Grade 4 ($n = 2$), multiplication in Grade 3 ($n = 6$) and Grade 4 ($n = 9$), division in Grade 3 ($n = 2$) Grade 4 ($n = 5$), equivalence in Grade 3 ($n = 1$) and Grade 4 ($n = 5$), and writing speed ($n = 3$). Sensitivity analyses with and without these outliers showed the same patterns of results, and thus all the data were included in the final analyses.

The mean scores (see Table 1) and the interquartile ranges, shown in Figure 1, suggest that most students were capable of solving all types of equivalence problems (i.e., equations with operands to the left of, right of, or on both sides of the equal sign). Thus, the variability in equivalence fluency likely reflects how flexibly students applied their arithmetic skills to solve these problems efficiently.

A few significant gender differences were present. In Grade 3, compared to girls, boys solved more addition (24.4 versus 23.6, $t = 2.35$, $p = 0.019$, $d = .19$) subtraction (24.4 versus 23.6, $t = 2.32$, $p = 0.021$, $d = .19$), and multiplication (28.9 versus 28.3, $t = 1.96$, $p = 0.051$, $d = .16$) problems correctly. In Grade 4, boys solved more subtraction (25.5 versus 24.8, $t = 1.96$, $p = 0.050$, $d = .17$) and multiplication (31.2 versus 30.4, $t = 3.50$, $p = 0.001$, $d = .30$)

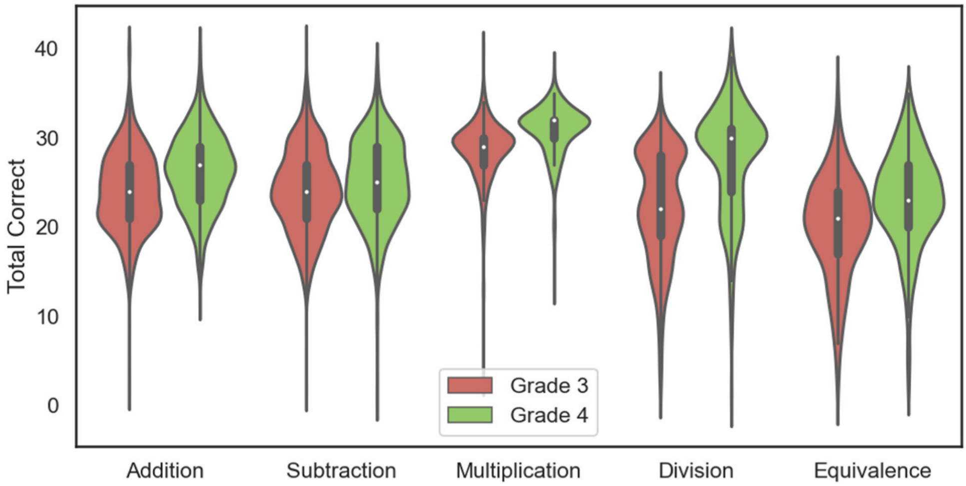

Figure 1. Violin plots for standard arithmetic and equivalence in grade 3 (time 1) and grade 4 (time 2). The white dot is the median, and the black bar in the center of the plot shows the interquartile range.

problems correctly. There were no other significant gender differences (*p*s > 0.05). We note that these differences are trivial (i.e., small effect sizes, mean differences < 1 point) and that the statistical significance reflects the large sample size. Nevertheless, for further analyses, we controlled for gender.

Development of arithmetic and equivalence

All of the measures were significantly correlated, except for number writing speed with non-verbal reasoning and equivalence (see Table 2). Moreover, students improved from Grades 3 to 4 on the arithmetic and equivalence measures (see Table 1 for pairwise comparisons). Based on the significant correlations between standard arithmetic and equivalence and the improvement from Grades 3 to 4, we proceeded with cross-lagged analyses to further investigate the longitudinal relations between these measures.

Multilevel cross-lagged structural equation modelling

The goal of the present study was to investigate the development of standard arithmetic fluency (i.e., addition, subtraction, multiplication, division) and equivalence fluency from Grade 3 to Grade 4. Given that students were from 14 classrooms, we tested a multilevel cross-lagged structural equation model to account for the hierarchical nature of the dataset using M*plus* (Muthén & Muthén, 1998). Model fit was examined using a combination of the chi-square goodness of fit test ($p > 0.05$), comparative fit index (CFI > 0.90), root mean square error of approximation (RMSEA < 0.06), Tucker-Lewis index (TLI > 0.90), and standardized root mean square residual (SRMR <.08; Hu & Bentler, 1999).

In Grade 3, for each variable, a low percentage of data (<0.2%) were missing (see Table 1). In Grade 4, 44 students dropped out of the study for personal reasons. Given the low attrition rate (7.2%) relative to the sample size, these missing cases were unlikely to influence the interpretation of the results even if they were not missing at random (Enders, 2010). First, sensitivity analyses were conducted to compare the results for analyses with all students ($n = 612$) versus only those with complete data ($n = 569$). The pattern of results was the same in both analyses and thus the final model was estimated by a full information maximum likelihood method where all available information is used in all observations to find the optimal combination of estimates for the missing parameters (Enders, 2010).

We examined the longitudinal relations between standard arithmetic and equivalence fluency from Grades 3 to 4. We hypothesized that there would be bidirectional relations between standard arithmetic and equivalence fluency (Hypothesis 1), controlling for students' number writing speed. We tested a multilevel cross-lagged structural equation model, specifying classroom as a random effect. First, an intercept-only model was fit which contained a classroom variable. The intra-class correlation coefficients were .10 (addition), .14 (subtraction), .13 (multiplication), .06 (division), .15 (equivalence) and .05 (non-verbal reasoning), indicating modest variability in performance in Grade 3 among the classrooms.

Second, we added a confirmatory factor analysis to the intercept-only model to test factor loadings for each of the arithmetic operations on latent standard arithmetic constructs in Grades 3 and 4. The model had good fit to the data, $\chi^2(15) = 33.37$, $p = 0.004$, SRMR = .021, CFI = .989, TLI = .976, RMSEA = .045. The factor loadings of each type of arithmetic operation on the latent variable were high (see Figure 2). The slightly lower factor loadings

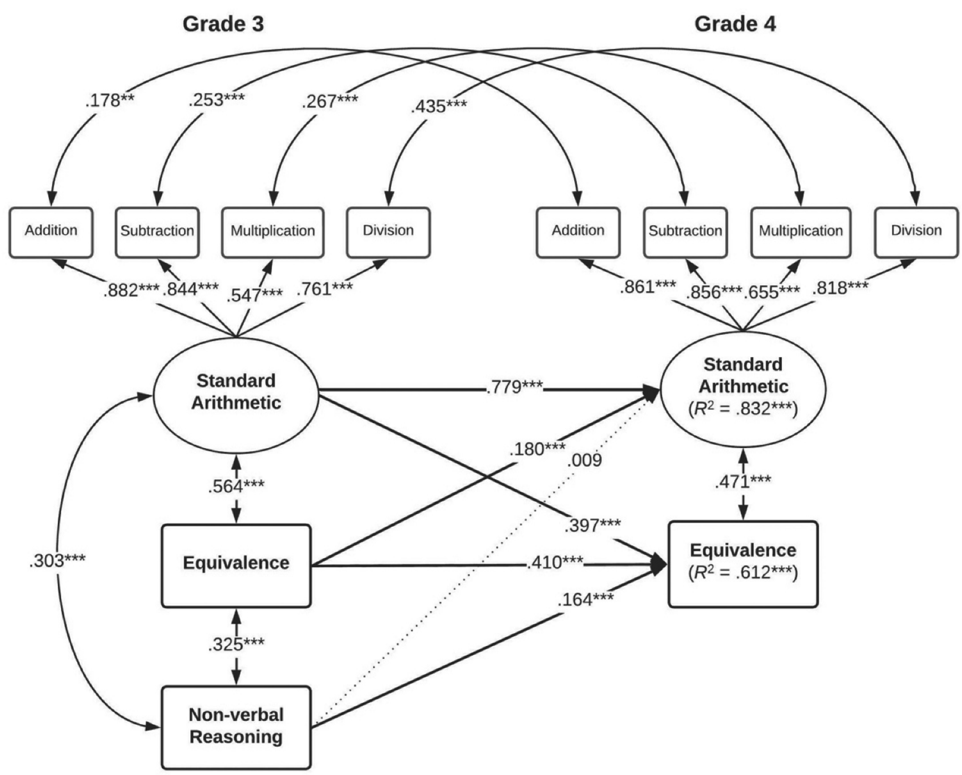

Figure 2. Multilevel cross-lagged structural equation model for standard arithmetic and equivalence fluency for students in grade 3 (Time 1) and grade 4 (Time 2).

for multiplication reflect that there was less variability in this operation. These latent variables were subsequently labeled *standard arithmetic*.

Then, we constructed a cross-lagged path model between students' standard arithmetic and equivalence fluency in Grades 3 and 4, controlling for number writing speed. Non-verbal reasoning in Grade 3 was modeled as a correlate and predictor of both standard arithmetic and equivalence variables in Grades 3 and 4, respectively. We compared models in which the two cross-lagged paths were either constrained to be equal or were freely estimated using the Satorra-Bentler scaled chi-square difference test (Satorra & Bentler, 2010). The constrained model did not fit well, $\chi^2(54) = 179.13$, $p < 0.001$, SRMR = .057, CFI = .942, TLI = .906, RMSEA = .062. In contrast, the unconstrained model had acceptable fit to the data, $\chi^2(53) = 120.47$, $p < 0.001$, SRMR = .034, CFI = .969, TLI = .948, RMSEA = .046. Moreover, the cross-lagged path coefficients were significantly different between the constrained and unconstrained models, $\Delta\chi^2(1) = 1923.75$, $p < 0.001$, implying that the path from standard arithmetic in Grade 3 to equivalence in Grade 4 was significantly stronger than the path from equivalence in Grade 3 to standard arithmetic in Grade 4. Thus, the unconstrained model was retained for further interpretation (see Figure 2).

Consistent with Hypothesis 1, standard arithmetic in Grade 3 significantly predicted the improvement of equivalence from Grades 3 to 4. Equivalence in Grade 3 also predicted the improvement in standard arithmetic from Grades 3 to 4. Consistent with Hypothesis 2, the strength of the path coefficient from standard arithmetic to equivalence was stronger than the path coefficient from equivalence to standard arithmetic. Finally, consistent with Hypothesis 3, non-verbal reasoning was related to the development of equivalence fluency, but not to the development of standard arithmetic fluency, highlighting the specific role of non-verbal reasoning in solving equivalence problems.

Because the equivalence task only included additive operations (i.e., addition, subtraction, and mixed addition and subtraction), and considering the conceptual distinction between additive and multiplicative operations (e.g., Harel & Confrey, 1994; Robinson, 2017; Steffe, 1992), in additional exploratory analyses, we used multi-level SEM to separately test the bidirectional relations between the additive operations and equivalence, and between the multiplicative operations and equivalence. Similar to the original model, for the additive skills model we found reciprocal relations between the development of additive operations and equivalence fluency from Grades 3 to 4, with a greater influence of the additive operations on the development of equivalence than the reverse (see Figure 3). In contrast, for the multiplicative skills model, equivalence fluency in Grade 3 did not predict multiplicative skills in Grade 4 (see Figure 4). For both the additive and multiplicative models, non-verbal reasoning was related to the development of equivalence fluency. Overall, these additional analyses provide further evidence supporting the importance of students' fluent access to arithmetic skills as a predictor of the development of equivalence fluency but also show that equivalence fluency may be operation specific.

Discussion

Understanding mathematical equivalence is fundamental to students' mathematical development (Fyfe, Matthews, Amsel, McEldoon, & McNeil, 2018; Kieran, 1981; Knuth, Stephens, McNeil, & Alibali, 2006; McNeil, 2008). Once students understand that the equal sign represents an equivalence relation, they can use this understanding to develop

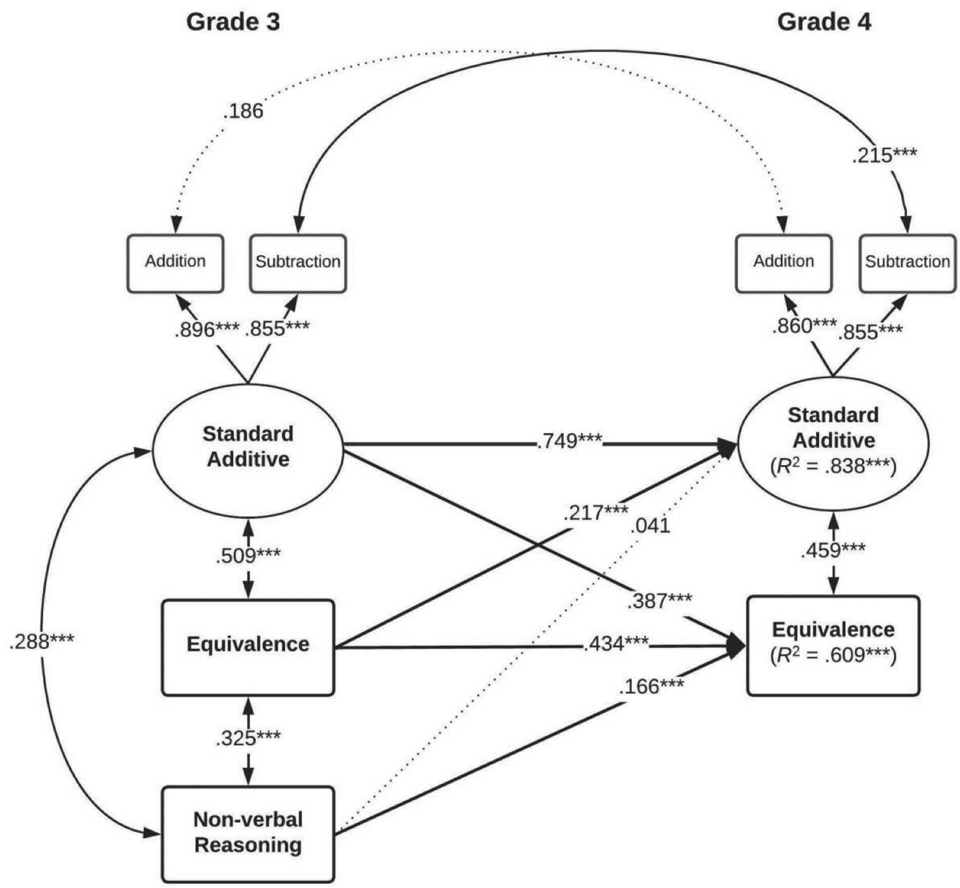

Figure 3. Multilevel cross-lagged SEM model for standard additive and equivalence fluency for students in grade 3 (Time 1) and grade 4 (Time 2). The model had excellent fit to the data, $\chi^2(17) = 23.271$, $p = 0.141$, SRMR = 0.034, CFI = 0.994, TLI = 0.986, RMSEA = 0.025. Numbers on the arrows are standardized coefficients. ***$p < 0.001$. Dashed lines represent non-significant paths. Number writing speed and gender were controlled for both the standard additive operations ($\beta_{Speed} = .071$, $p = 0.033$; $\beta_{Gender} = -.009$, $p = 0.706$) and equivalence ($\beta_{Speed} = -.065$, $p = 0.051$; $\beta_{Gender} = -.020$, $p = 0.298$) in grade 4. R^2 values for each outcome include variance predicted by the control measures.

flexible strategies to efficiently solve nonstandard equivalence problems. In the present study, we investigated the reciprocal relations and the role of reasoning skills in the development of arithmetic and equivalence fluency for Chinese students from Grades 3 to 4.

Using a cross-lagged analytical framework, we found that arithmetic fluency in Grade 3 predicted the change in Chinese students' equivalence fluency from Grades 3 to 4 and vice versa, demonstrating that these two types of knowledge are reciprocally related. Consistent with previous work showing that, from Grades 3 to 4, Chinese students are continuing to develop equivalence knowledge (Yang, Huo, & Yan, 2014), we found bidirectional relations between arithmetic and equivalence. Thus, knowledge of one skill facilitates knowledge of the other.

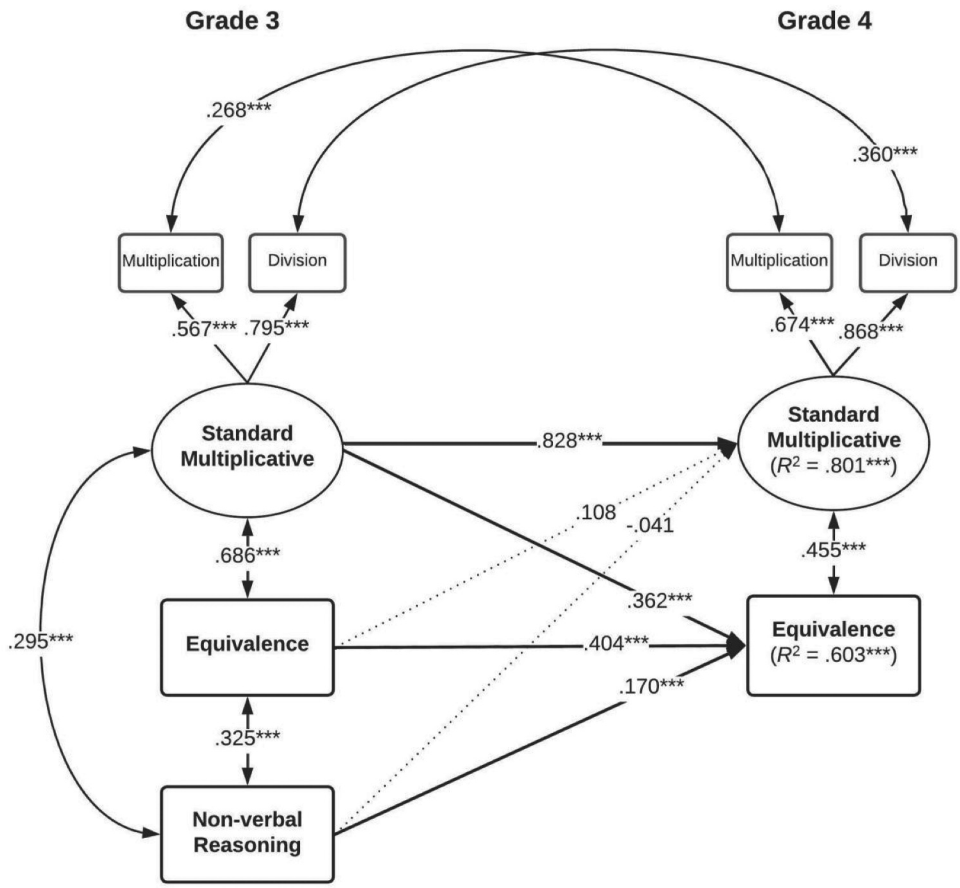

Figure 4. Multilevel cross-lagged SEM model for standard multiplicative and equivalence fluency for students in grade 3 (Time 1) and grade 4 (Time 2). The model had good fit to the data, $\chi^2(17) = 25.885$, $p = .077$, SRMR = 0.035, CFI = 0.990, TLI =0 .977, RMSEA = 0.029. Numbers on the arrows are standardized coefficients. ***$p < .001$. Dashed lines represent non-significant paths. Number writing speed and gender were controlled for both the standard multiplicative ($\beta_{Speed} = .003$, $p = .944$; $\beta_{Gender} = -.032$, $p = .097$) and equivalence ($\beta_{Speed} = -.034$, $p = .439$; $\beta_{Gender} = -.040$, $p = .072$) in grade 4. R^2 values for each outcome include variance predicted by the control measures.

Although the two skills develop together, the size of the relation between arithmetic in Grade 3 and equivalence in Grade 4 was larger than between equivalence in Grade 3 and arithmetic in Grade 4. In the present study, equivalence fluency was assessed using a nonstandard arithmetic task that included addition, subtraction, and mixed addition and subtraction problems. Thus, proficiency in standard additive arithmetic skills (addition and subtraction) serves as a prerequisite for solving nonstandard equivalence problems. Moreover, beyond operation-specific relations, our additional models revealed that both additive and multiplicative arithmetic skills significantly predicted the development of equivalence fluency. The process of solving mathematical equivalence problems involves various subskills, such as encoding numbers and operators and carrying out calculations (McNeil, Hornburg, Fuhs, & Connor, 2017). When students have fluent access to basic arithmetic facts, they can free up space in their working memory, allowing them to apply efficient strategies when performing calculations (Anderson, 2002). Moreover, students who generate a greater variety of equations that are equal to a given value (e.g., 4 can be represented as “3 + 1”, “2 + 2”, ’5–1’, “2 × 2”, “8 ÷ 2” and so on) show a better understanding of equivalence than children who generate fewer equations (Chesney et al., 2014). Establishing stronger interconnections among numbers may therefore enable students to use their knowledge of arithmetic facts in novel situations, leading to efficient solutions for nonstandard arithmetic problems presented in various formats (Chesney et al., 2014; Chesney, McNeil, Petersen, & Dunwiddie, 2018; McNeil et al., 2012). In the present study, the arithmetic fluency task consisted of problems of varying levels of difficulty. As such, performance on this task not only reflects students’ ability to efficiently retrieve number facts (e.g., 2 + 5 = __) but also their ability to use effective strategies to solve complex problems (e.g., 24 + 49 = __). Consistent with Chesney et al. (2014), our findings suggest that the strong predictive link between arithmetic and equivalence fluency may be attributed to students’ proficiency in retrieving or computing arithmetic facts when tackling equivalence problems.

Interestingly, although arithmetic fluency in Grade 3 predicted equivalence fluency in Grade 4 for the additive and multiplicative models, equivalence fluency in Grade 3 only predicted arithmetic in Grade 4 in the additive model. In the present study, equivalence fluency was assessed using a nonstandard arithmetic task that required only addition and subtraction knowledge. Thus, one possibility is that the predictive link from equivalence to arithmetic is specific to the operations included in the equivalence task. This possibility is consistent with the view that specific arithmetic operations play an important role in students’ development of understanding arithmetic concepts (Robinson, Price, & Demyen, 2018). Another possibility is that the multiplicative model reflects the limited variability in multiplication fluency in both Grades 3 and 4. In China, students are expected to acquire multiplication (i.e., memorize the multiplication table up to 9 × 9) by Grade 3. Thus, because many students had high scores on the multiplication task, as evidenced by the skewed distribution, there may be little change in arithmetic fluency to predict in the multiplicative model. To gain further insights into the bidirectional relations between different types of arithmetic operations and equivalence problems, future research should incorporate both additive and multiplicative operations within the equivalence task. This approach will allow for a more comprehensive understanding of the specific relations between equivalence knowledge and arithmetic knowledge for each type of arithmetic operation.

We also found that non-verbal reasoning predicted the development of equivalence fluency from Grades 3 to 4, supporting the view that development of equivalence knowledge involves reasoning skills (Miller Singley & Bunge, 2014; Morsanyi, Prado, & Richland, 2018). Reasoning skills may help students notice the connections between expressions on each side of the equation and assist them in mentally transforming the expressions to simplify calculation (Kindrat & Osana, 2018). In contrast, non-verbal reasoning did not predict the development of standard arithmetic fluency from Grades 3 to 4 (Peng et al., 2016), even though there were concurrent relations between reasoning and arithmetic fluency (Fuchs et al., 2006; Xu et al., 2021). In summary, these results highlight the specific role of non-verbal reasoning in the development of students’ ability to solve arithmetic problems presented in nonstandard formats.

Educational implications

Our research is necessarily tied to the educational experiences of students in China. The differences in educational experiences for students in China versus Western countries are pervasive across the materials and methods used to teach arithmetic (Li & Huang, 2013). For example, in contrast to the mathematics textbooks in Western countries, Chinese mathematics textbooks introduce the equal sign in a context of number relations and teach students to interpret the equal sign as balancing two sides of the equation starting in Grade 1 (Capraro, Ding, Matteson, Capraro, & Li, 2007). Beyond textbooks, Chinese teachers provide more instruction on equivalence and Chinese students have more exposure to nonstandard equations that promote the relational understanding of the equal sign compared to students in Western countries (Li, Ding, Capraro, & Capraro, 2008; Powell, 2012). These differences are critical for the development of students' knowledge of equivalence (Knuth, Stephens, McNeil, & Alibali, 2006; Li, Ding, Capraro, & Capraro, 2008; Powell, 2012). Nevertheless, although Chinese students are exposed to intensive instruction and practice on equivalence as young as Grade 1, the comparative relational understanding of equivalence continues to develop throughout the elementary school years, such that most students do not have full knowledge of equivalence until Grade 6 (Jones, Inglis, Gilmore, & Dowens, 2012; Yang, Huo, & Yan, 2014).

One way to assist students in their development of challenging concepts such as equivalence is to ensure that they develop fluent access to related core knowledge, such as basic arithmetic facts. Mathematics education in China emphasizes the importance of developing an interconnected understanding of mathematics concepts and skills through repetition (Ministry of Education, 2011). Notably, the purpose of repetition is not to encourage students to memorize number facts mechanically through rehearsal strategies (i.e., rote learning). Instead, the purpose of repetition is to promote deepened understanding by focusing on different aspects of the same question each time through problem variations (Dahlin & Watkins, 2000; Marton, Dall' Alba, & Tse, 1996).

The present results support the view that flexible access to arithmetic facts will support students' knowledge of equivalence (Kindrat & Osana, 2018). Conversely, developing equivalence fluency may also be important for selection of efficient procedures when students are solving arithmetic problems. Equivalence fluency could be developed through consistent exposure to various types of nonstandard equations that provide students with ample practice in solving nonstandard problems (Powell, 2012). For example, students could be exposed to operations without equal signs (e.g., a + b), simple equations (e.g., __ = a, b = __), and equations that are more complex in nonstandard contexts (e.g., a + __ = b + c), to enhance their understanding of the equal sign. Furthermore, emphasizing the connections between closely related concepts will lead to students more frequently using flexible solutions on mathematics problems (Baroody, 1999; Kindrat & Osana, 2018; Ma, 2010; Robinson, 2017).

Our finding that students' reasoning skills predicted their ability to efficiently solve mathematics equivalence problems has educational implications. Because understanding of equivalence goes beyond applying algorithms to the problems by rote, educators can encourage students to use reasoning skills to identify relations among numbers and number operations (Jacobs, Franke, Carpenter, Levi, & Battey, 2007). For example, instead of immediately trying to perform an operation, students can consider the magnitude of values on each side of the equal sign and determine whether there is a simpler solution (e.g., solving 181−25 = __ − 24 by recognizing that 25 and 24 are only one unit apart; Kindrat & Osana, 2018). Recognizing these relations may help students to move past the equal sign as an indicator that an answer must be provided and think of the equal sign as representing the balance between two sides of an equation.

Limitations and future research

In the present study, we assumed that Chinese students were using more efficient strategies to solve equivalence problems in Grade 4 than Grade 3, based on the total number of problems solved in a limited time period. However, to fully assess this claim, research is needed in which students solve various types of complex equivalence problems while both accuracy and response time on each individual trial are recorded and strategy use data is collected (Robinson, Dubé, & Beatch, 2017). Moreover, in the present study, only addition and subtraction problems were used to index students' knowledge of equivalence. To improve construct validity, future studies should use multiple measures for measuring equivalence to determine if all the measures tap into the same underlying construct (e.g., a task that measures multiplication and division equivalence knowledge; Robinson, Price, & Demyen, 2018).

Conclusion

Despite a large corpus of research investigating the development of arithmetic and equivalence, our study is the first to find reciprocal relations between these two types of mathematics knowledge from Grades 3 to 4, with a greater influence of arithmetic on the development of equivalence than the reverse for Chinese students. Our findings suggest that in addition to strong domain-general reasoning skills, students need to acquire fluent access to basic arithmetic associations to facilitate their development of equivalence fluency.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China under Grant 31971010 and Natural Sciences and Engineering Research Council of Canada under Grant RGPIN-2017-06355.

ORCID

Chang Xu http://orcid.org/0000-0002-6702-3958
Hongxia Li http://orcid.org/0000-0002-5052-2742
Sabrina Di Lonardo Burr http://orcid.org/0000-0001-6338-9621
Jiwei Si http://orcid.org/0000-0003-1342-7348
Jo-Anne LeFevre http://orcid.org/0000-0002-1927-7734
Xinfeng Zhuo http://orcid.org/0009-0008-6492-2457

Data availability statement

To promote transparency and openness, anonymized data for the measures analyzed in the current paper and analysis code are freely available for download at [https://osf.io/t7fbj/?view_only=e9d9cfacb69b413aab4183734cc33fb7].

Open scholarship

This article has earned the Center for Open Science badge for Open Data. The data are openly accessible at https://osf.io/t7fbj/?view_only=e9d9cfacb69b413aab4183734cc33fb7.

References

Alexander, P. A., Jablansky, S., Singer, L. M., & Dumas, D. (2016). Relational reasoning: What we know and why it matters. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 36–44. doi:10.1177/2372732215622029

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*(1), 127–145. doi:10.1037/0012-1649.35.1.127

Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, *26*(1), 85–112. doi:10.1207/s15516709cog2601_3

Baroody, A. J. (1985). Mastery of basic number combinations: Internalization of relationships or facts? *Journal for Research in Mathematics Education*, *16*(2), 83–98. doi:10.2307/748366

Baroody, A. J. (1999). Children's relational knowledge of addition and subtraction. *Cognition and Instruction*, *17*(2), 137–175. doi:10.1207/S1532690XCI170201

Capraro, M. M., Ding, M., Matteson, S., Capraro, R. M., & Li, X. (2007). Representational implications for understanding equivalence. *School Science and Mathematics*, *107*(3), 86–88. doi:10.1111/j.1949-8594.2007.tb17773.x

Carpenter, T. P., Franke, L. P., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic & algebra in elementary school*. Portsmouth, NH: Heinemann.

Carpenter, T. P., & Levi, L. (2000). Developing conceptions of algebraic reasoning in the primary grades (Research Report #002). Madison, WI: National Center for Improving Student Learning and Achievement in Mathematics and Science. www.wcer.wisc.edu/ncisla/publications/index.html

Chesney, D. L., McNeil, N. M., Matthews, P. G., Byrd, C. E., Petersen, L. A. . . . Dunwiddie, A. E. (2014). Organization matters: Mental organization of addition knowledge relates to understanding math equivalence in symbolic form. *Cognitive Development*, *30*, 30–46. doi:10.1016/j.cogdev.2014.01.001

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. doi:10.1016/j.lindif.2018.04.013

Dahlin, B., & Watkins, D. (2000). The role of repetition in the processes of memorising and understanding: A comparison of the views of German and Chinese secondary school students in Hong Kong. *British Journal of Educational Psychology*, *70*(1), 65–84. doi:10.1348/000709900157976

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M. . . . Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*(1), 29–43. doi:10.1037/0022-0663.98.1.29

Fyfe, E. R., Matthews, P. G., Amsel, E., McEldoon, K. L., & McNeil, N. M. (2018). Assessing formal knowledge of math equivalence among algebra and pre-algebra students. *Journal of Educational Psychology*, *110*(1), 87–101. doi:10.1037/edu0000208

Haffner, J., Baro, K., & Resch, F. (2005). *Heidelberger Rechentest (HRT 1-4) erfassung mathematischer basiskompetenzen im grundschulalter [The Heidelberg Mathematics Test (HRT 1-4): Assessing mathematics at primary school age]*. Göttingen: Hogrefe.

Harel, G., & Confrey, J. (1994). *The Development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi:10.1080/10705519909540118

Jacobs, V. R., Franke, M. L., Carpenter, T., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38*(3), 258–288. doi:10.2307/30034868

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

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, *12*(3), 317–326. doi:10.1007/BF00311062

Kindrat, A. N., & Osana, H. P. (2018). The relationship between mental computation and relational thinking in the seventh grade. *Fields Mathematics Education Journal*, *3*(1), 1–22. doi:10.1186/s40928-018-0011-4

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*, 297–312. doi:10.2307/30034852

Li, X., Ding, M., Capraro, M. M., & Capraro, R. M. (2008). Sources of differences in children's understandings of mathematical equality: Comparative analysis of teacher guides and student texts in China and the United States. *Cognition and Instruction*, *26*(2), 195–217. doi:10.1080/07370000801980845

Li, Y., & Huang, R. (2013). *How Chinese teach mathematics and improve teaching*. Routledge. doi:10.4324/9780203110119

Li, H. X., Zhang, A. X., Zhang, M. L., Huang, B. J., Zhao, X. M., Gao, J., & Si, J. W. (2021). Concurrent and longitudinal associations between parental educational involvement, teacher support, and math anxiety: The role of math learning involvement in elementary school children. *Contemporary Educational Psychology*, *66*, 101984. doi:10.1016/j.cedpsych.2021.101984

Ma, L. (2010). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Albany, NY: Routledge.

Marton, F., Dall' Alba, G., & Tse, K. T. (1996). Memorising and understanding: The keys to the paradox? In D. Watkins & J. B. Biggs (Eds.), *The Chinese learner: Cultural, psychological, and contextual influences* (pp. 69–83). Hong Kong: Comparative Education Research Centre and Australian Council for Educational Research.

Matthews, P. G., & Fuchs, L. S. (2020). Keys to the gate? Equal sign knowledge at second grade predicts fourth-grade algebra competence. *Child Development*, *91*(1), e14–e28. doi:10.1111/cdev.13144

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development*, *79*(5), 1524–1537. doi:10.1111/j.1467-8624.2008.01203.x

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. doi:10.1111/j.1467-8624.2005.00884.x

McNeil, N. M., Chesney, D. L., Matthews, P. G., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Wheeler, M. C. (2012). It pays to be organized: Organizing arithmetic practice around equivalent values facilitates understanding of math equivalence. *Journal of Educational Psychology*, *104*(4), 1109–1121. doi:10.1037/a0028997

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of mathematical equivalence. *Journal of Educational Psychology*, *107*(2), 423–436. doi:10.1037/a0037687

McNeil, N. M., Hornburg, C. B., Devlin, B. L., Carrazza, C., & McKeever, M. O. (2019). Consequences of individual differences in children's formal understanding of mathematical equivalence. *Child Development*, *90*(3), 940–956. doi:10.1111/cdev.12948

McNeil, N. M., Hornburg, C. B., Fuhs, M. W., & Connor, D. O. (2017). Understanding children's difficulties with mathematical equivalence. In *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 167–195). Academic Press. doi:10.1016/B978-0-12-805086-6.00008-4

Miller Singley, A. T., & Bunge, S. A. (2014). Neurodevelopment of relational reasoning: Implications for mathematical pedagogy. *Trends in Neuroscience and Education*, *3*(2), 33–37. doi:10.1016/j.tine.2014.03.001

Ministry of Education. (2011). *Curriculum standards for school mathematics of nine-year compulsory education*. Beijing: Beijing Normal University Press.

Morsanyi, K., Prado, J., & Richland, L. E. (2018). Editorial: The role of reasoning in mathematical thinking. *Thinking & Reasoning*, *24*(2), 129–137. doi:10.1080/13546783.2018.1435425

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. New York: Muthén & Muthén.

National Bureau of Statistics. (2019). *Annual of statistics about Chinese cities*. Beijing: Statistics Publishing House of China.

Osana, H. P., & Kindrat, A. N. (2021). Examining the impact of a mental computation classroom intervention on the relational thinking of seventh-grade students. *Journal of Mathematics Education*, *14*(1), 47–70. doi:10.26711/007577152790063

Peng, P., Namkung, J. M., Fuchs, D., Fuchs, L. S., Patton, S. . . . Hamlett, C. (2016). A longitudinal study on predictors of early calculation development among young children at risk for learning difficulties. *Journal of Experimental Child Psychology*, *152*, 221–241. doi:10.1016/j.jecp.2016.07.017

Powell, S. R. (2012). Equations and the equal sign in elementary mathematics textbooks. *The Elementary School Journal*, *112*(4), 627–648. doi:10.1086/665009

Raven, J. C. (1938). *Raven's progressive matrices (1938): Sets* A, B, C, D, E. Melbourne: Australian Council for Educational Research.

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, *103*(1), 85–104. doi:10.1037/a0021334

Robinson, K. M. (2017). The understanding of additive and multiplicative arithmetic concepts. In D. Geary, D. Berch, R. Ochsendorf, & Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 21–46). Elsevier. doi:10.1016/B978-0-12-805086-6.00002-3

Robinson, K. M., Dubé, A. K., & Beatch, J.-A. (2017). Children's understanding of additive concepts. *Journal of Experimental Child Psychology*, *156*, 16–28. doi:10.1016/j.jecp.2016.11.009

Robinson, K. M., Price, J. A. B., & Demyen, B. (2018). Understanding arithmetic concepts: Does operation matter? *Journal of Experimental Child Psychology*, *166*, 421–436. doi:10.1016/j.jecp.2017.09.003

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference Chi-square test statistic. *Psychometrika*, *75*(2), 243–248. doi:10.1007/s11336-009-9135-y

Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and nonsymbolic contexts: Benefits of solving problems with manipulatives. *Journal of Educational Psychology*, *101*(1), 88–100. doi:10.1037/a0013156

Steffe, L. P. (1992). Schemes of action and operation involving composite units. *Learning and Individual Differences*, *4*(3), 259–309. doi:10.1016/1041-6080(92)90005-Y

Wu, H., & Li, L. (2005). Development of Chinese rating scale of pupil's mathematic abilities and study on its reliability and validity. *China Health*, *21*(4), 473–475.

Xu, C., LeFevre, J.-A., Skwarchuk, S., Di Lonardo Burr, S., Lafay, A. . . . Simms, V. (2021). Individual differences in the development of children's arithmetic fluency from grades 2 to 3. *Developmental Psychology*, *57*(7), 1067–1079. doi:10.1037/dev0001220

Yang, X., Huo, Y., & Yan, Y. (2014). Primary school students' understanding of equation structure and the meaning of equal Sign: A Chinese sample study. *Research in Mathematical Education*, *18*(4), 237–256. doi:10.7468/jksmed.2014.18.4.237