

This is a repository copy of *Two-digit number writing and arithmetic in Year 1 children : does number word inversion matter?*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/166862/>

Version: Published Version

---

**Article:**

Clayton, Francina Jane, Copper, Clare, Steiner, Anna et al. (4 more authors) (2020) Two-digit number writing and arithmetic in Year 1 children : does number word inversion matter? *Cognitive Development*. 100967. ISSN 0885-2014

<https://doi.org/10.1016/j.cogdev.2020.100967>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Cognitive Development

journal homepage: [www.elsevier.com/locate/cogdev](http://www.elsevier.com/locate/cogdev)

## Two-digit number writing and arithmetic in Year 1 children: Does number word inversion matter?

Francina J. Clayton<sup>a</sup>, Clare Copper<sup>a</sup>, Anna F. Steiner<sup>b</sup>, Chiara Banfi<sup>b</sup>,  
Sabrina Finke<sup>b</sup>, Karin Landerl<sup>b,c</sup>, Silke M. Göbel<sup>a,d,\*</sup>

<sup>a</sup> Department of Psychology, University of York, UK

<sup>b</sup> Institute of Psychology, University of Graz, Austria

<sup>c</sup> Department of Cognitive Science, Macquarie University, Sydney, Australia

<sup>d</sup> Department of Special Needs Education, University of Oslo, Norway

### ARTICLE INFO

#### Keywords:

Transcoding  
Arithmetic  
Cross-linguistic  
Number writing  
Mathematical development  
Number word inversion

### ABSTRACT

Early mathematical development relies upon the ability to translate between spoken number words and Arabic digits. Few studies have investigated whether differences in number word structure influence transcoding and its relationship with mathematics. We tested number writing and arithmetic in 177 German-speaking and 309 English-speaking Year 1 children. In English number words the order of tens and units (e.g., twenty-five) follows the written order of the Arabic digits (e.g., 25), whereas German number words are inverted (e.g., 'fünfundzwanzig', five-and-twenty). Transcoding at the item level was strongly influenced by number word inversion. German-speaking children made more errors transcoding two-digit numbers. However, English-speaking children made more errors writing teens, which are exceptional in English because of their inverted structure. Nevertheless, number writing was a significant predictor of arithmetic in both languages highlighting a significant relationship between transcoding ability and arithmetic for languages with and without number word inversion.

### 1. Introduction

Transcoding between spoken number words and Arabic digits works seamlessly in most adults. Young children, however, take years to become proficient transcoders. Individual differences in transcoding are associated with concurrent arithmetic performance (Geary, Hoard, & Hamson, 1999; Simmons, Willis, & Adams, 2012; Sowinski et al., 2015); however, the mechanisms underlying this relationship are not yet clear. Language-specific morphology such as number word inversion might play a role. The aim of the current study was to investigate the influence of language, specifically number word inversion, upon transcoding and its relationship with arithmetic in early primary school.

#### 1.1. Transcoding: from number words to Arabic digits

When typically developing children enter school, their production and comprehension of single Arabic digits is near perfect (Geary et al., 1999; Geary, Hamson, & Hoard, 2000; Johansson, 2005; Zuber, Pixner, Moeller, & Nuerk, 2009). However, learning the

\* Corresponding author at: Department of Psychology, University of York, UK.  
E-mail address: [silke.goebel@york.ac.uk](mailto:silke.goebel@york.ac.uk) (S.M. Göbel).

<https://doi.org/10.1016/j.cogdev.2020.100967>

Received 15 April 2020; Received in revised form 6 October 2020; Accepted 13 October 2020

Available online 1 November 2020

0885-2014/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

correspondences between single digits and number words is only a first step in the acquisition of transcoding between verbal numerical and Arabic-visual codes. Children need to quickly shift from using spoken number words (e.g., 'twenty-three') to being able to write down multi-digit Arabic strings (e.g., '23') upon hearing spoken number words. Without formal instruction, 4–6 year old children, have some rudimentary understanding about three-digit Arabic numerals (Yuan, Prather, Mix, & Smith, 2019): they know that the first digit is written on the left, the second written in the middle and the third mentioned written on the right (Byrge, Smith, & Mix, 2014). Despite this rudimentary understanding, children in early primary school are still learning number words for many two-digits numbers. Barrouillet, Thevenot, and Fayol (2010) reported that 5-year-old French-speaking children on average only knew the number words up to thirty-four (34) and none of the 49 children tested knew the numbers larger than seventy-nine (79). Our study explicitly focuses on the ability of children in early primary school to successfully transcode two-digit numbers.

For accurate number transcoding children need to acquire the complexities of the place-value system and the language-specific rules of multi-digit numbers. The Arabic number system for multi-digits is relatively simple: it consists of 10 elements (0–9) and one principle (place-value system). In contrast, verbal number systems are more complex. Typically they consist of a limited lexicon with entries in separate lexical classes (so-called lexical primitives) such as units ('one' to 'nine'), decades ('ten' to 'ninety'), hundreds, thousands etc. Elements of these lexical classes can be combined following a syntax. This syntax has additive (e.g., '349' as *three hundred (and) forty-nine*:  $300 + 49$ ) and multiplicative rules (e.g., '500' as *five hundred*:  $5 \times 100$ ). In many languages, the lexicon also contains particulars such as 'eleven' and 'twelve'. Mapping from spoken number words to Arabic digits is not straightforward. For example, zeros are absent in spoken multi-digit numbers, but essential in the Arabic notation (e.g., 'four hundred and three': 403).

There are two broad classes of models of multi-digit number processing: semantic and asemantic models. The main assumption in semantic models (e.g., McCloskey, 1992; McCloskey, Caramazza, & Basili, 1985; Power & Dal Martello, 1990) is that number words are first transformed into a semantic magnitude representation and then into its constituent Arabic digits. These models predict number transcoding to be harder for larger numbers even if they contain the same number of digits (e.g., harder for 8 than 2, or 87 than 23, a numerical size effect). While there is little evidence for this from errors committed in number transcoding (Barrouillet, Camos, Peruchet, & Seron, 2004), research on planning time for transcoding hints at the existence of a semantic route early in the development of number transcoding (van Loosbroek, Dirx, Hulstijn, & Janssen, 2009). Imbo, Vanden Bulcke, De Brauwer, and Fias (2014) suggested that the selection between semantic and asemantic routes depends on children's mathematical skills with less-skilled transcoders using semantic rather than asemantic transcoding routes.

In contrast to semantic models, asemantic models do not propose that activation of the number magnitude of the digits is a necessary part of the transcoding process (e.g., Verguts & Fias, 2006). For number writing the most influential asemantic model to date is ADAPT (A Developmental, Asemantic, and Procedural model for Transcoding from verbal to Arabic numerals; Barrouillet et al., 2004). In this model the verbal input is stored as a phonological sequence. Parsing mechanisms then subdivide this sequence into smaller units that are processed by a production system. The initial role of the production system is retrieval of existing (learned) Arabic forms from long-term memory (LTM), managing the number of slots for the digit chain and filling all slots. It uses so-called separators such as thousand and hundred to identify the number of slots. Once each verbal unit is translated into its digit form, placed in the chain and all the slots are filled, the number is transcribed. Clearly, LTM as well as verbal and visuo-spatial working memory play an important role in this model. The model is explicitly developmental: it predicts that experience leads to 1) a larger number of Arabic forms that can be directly retrieved from LTM and 2) an improvement of transcoding rules.

## 1.2. Syntactic differences between languages: the case of number word inversion

None of the existing models of number transcoding account for syntactic differences between number words in different languages. This is surprising, because, while the Arabic notational system is used worldwide, the syntactic structure of number words differs between languages. It clearly influences number processing and arithmetic (e.g., Bahnmüller, Nuerk, & Moeller, 2018; Dowker & Nuerk, 2016) and differences between languages in number word morphology lead to distinct errors in number writing (e.g., Moeller, Zuber, Olsen, Nuerk, & Willmes, 2015; Seron & Fayol, 1994; Van Rinsveld & Schiltz, 2016). When number words are not transparent with respect to the Arabic place-value system, successful integration of tens and units into the place-value structure is particularly difficult.

Many studies on number writing in primary school children have focused on a particular syntactic feature in number words: inversion. In languages with number word inversion the spoken order of basic lexical elements is inverted with respect to the Arabic notation. While in left-to-right reading societies the first number in two-digit Arabic numbers is the decade digit and the second number is the unit digit (e.g., 25), in inverted number words the unit word is spoken first followed by the decade word (e.g., in German: 'fünfundzwanzig', literally five-and-twenty). There are several languages with full number word inversion (see Comrie, 2005): Arabic, Dutch, Danish, German, Malagasy, Maltese, and some languages use both inverted and non-inverted number words (e.g., Czech and Norwegian). Existing transcoding models do not fully explain the cognitive processes involved in transcoding in languages with number word inversion.

This, however, is an important goal for a developmental model of transcoding because number word inversion is a major obstacle for number writing in early primary school. Inversion errors, i.e., writing 52 for 'twenty-five' instead of 25, are much more common in children with a language background with number word inversion. We use the term non-inversion errors here for all other errors. Typical non-inversion errors are, for example, writing down a wrong digit (e.g., writing 28 for 'twenty-five') or not overwriting the zero (e.g., writing 205 for 'twenty-five'). Zuber et al. (2009) analysed number writing errors of 7-year-old Austrian (German-speaking) children in Year 1. Half of all errors were inversion errors. This is in stark contrast to children from a language-background without number word inversion (e.g., Krinzinger et al., 2011; Moeller et al., 2015). Directly comparing 7-year-old Dutch- and French-speaking

children in Year 2 in Belgium, Imbo et al. (2014) found a significantly higher percentage of inversion errors in Dutch- (with number word inversion) than in French- (without number word inversion) speaking children. Czech has two number word systems, an inverted and a non-inverted system, so in Pixner et al. (2011) the same 7-year-old children were tested on number writing in both systems. When the inverted system was used, around 50 % of errors were inversion-related, while the same children showed hardly any inversion errors when numbers were dictated using the non-inverted number word system. Interestingly, a study on a large sample of Dutch children from kindergarten (mean age 5–6 years) to the end of primary school (van der Ven, Klaiber, & van der Maas, 2016) showed that while the frequency of inversion errors declined with age, 57 % of children still made inversion errors in Year 6 (mean age 10–11 years). Clearly, number word inversion presents a challenge and this challenge may persist throughout primary school.

Does number word inversion lead to a general increase in transcoding errors, i.e., also in non-inversion errors, or does it specifically affect only the number of inversion errors? The additional complexity of number word inversion might tax children's available resources and thus could impair the efficiency of executing other transcoding rules (Camos, 2008; Imbo et al., 2014; Pixner et al., 2011). If this is the case, a larger number of non-inversion errors would also be predicted in languages with number word inversion than without number inversion. Alternatively, number word inversion might not influence other transcoding rules, in that case the number of non-inversion errors should not be affected by number-word inversion. Imbo et al. (2014) and Pixner et al. (2011) reported a smaller number of non-inversion errors for inverted than non-inverted languages. This suggests that number word inversion might not lead to a larger number of non-inversion errors.

While number word inversion in numbers larger than twenty has been researched extensively, number writing of teens, which are learned earlier, has received much less attention. In most languages teen number words consist of two types: particulars (e.g., 'eleven' in English) and combinations of a single-digit number word with a teen marker (e.g., 'sixteen' in English). Interestingly, these combination number words are often inverted with respect to the order of digits in the Arabic digit string, even in languages with otherwise non-inverted number words. This leads to a novel hypothesis: if it is the specific syntactic structure of the particular number word, rather than the general formation of number words, that leads to inversion errors, we expect inversion errors for teen numbers also in languages that are non-inverted for number words larger 20, e.g., in English. Furthermore, one might expect more inversion errors for teens in a language where number word inversion is specific to (most) teens than in a language where number word inversion is the norm for (nearly) all two-digit numbers. In a language with only inverted teen numbers, the application of an inversion rule is not needed for the transcoding of two-digit numbers. Overgeneralisation of the typical transcoding procedure could lead to more errors for teens in languages where the majority of spoken number words are not inverted. However, it is also possible that number word inversion operates only at the item level. If so, one would expect inversion errors for teens regardless of how many other number words are inverted in a particular language.

### 1.3. Individual differences in number writing as predictor of arithmetic

While errors in number writing have been investigated for many years, only recently has number writing been proposed as a foundational skill for arithmetic development (e.g., Göbel, Watson, Lervåg, & Hulme, 2014; Habermann, Donlan, Göbel, & Hulme, 2020; Malone, Burgoyne, & Hulme, 2019; Sowinski et al., 2015; for an overview of studies see Appendix A). Earlier studies (e.g., Johansson, 2005) reported correlations between number writing and arithmetic, but concluded that there is no direct link from numeral knowledge to arithmetic. However, there is some evidence in the literature that children with low mathematical performance (i.e., in the bottom 25–35 %) perform worse on place-value understanding and number writing. Hanich, Jordan, Kaplan, and Dick (2001), for example, showed that children with low mathematical performance perform significantly worse than typically developing children on positional knowledge and digit correspondence. Six- and seven-year-old children with both low mathematical and low reading performance in Year 1 and Year 2 also have difficulties in number writing (Geary et al., 1999, 2000). Moura et al. (2015) found significantly more errors in number writing in Year 1 and Year 2 children with low mathematical performance than in control children. It has been suggested that children with low mathematical performance might struggle to establish an Arabic numerical lexicon and to acquire more complex transcoding rules (Moura et al., 2013).

Recently, performance on number system knowledge tasks, including multi-digit Arabic numbers and place-value understanding, has been shown to be a predictor of individual differences in arithmetic and calculation (e.g., Göbel, Moeller, Kaufmann, Pixner, & Nuerk, 2014; Göbel, Watson et al., 2014; Habermann et al., 2020; Malone et al., 2019; Sowinski et al., 2015). These findings suggest that the ability to integrate two exact symbolic number systems (number words and Arabic digits) is an important predictor of growth of arithmetic skills at the beginning of primary school.

To our knowledge, only three studies have directly investigated number writing as a predictor of arithmetic. While all three studies used two-digit target numbers, the majority of items were three-digit numbers and overall transcoding accuracy was used as predictor. In a study with 7-year-old Dutch-speaking children from Belgium, more skilled transcoders achieved significantly higher mathematics scores than less skilled transcoders (Imbo et al., 2014). van der Ven et al. (2016) also reported a significant concurrent relationship between number writing ability and arithmetic performance in a large sample of Dutch children. In a longitudinal study with German-speaking children, Moeller, Pixner, Zuber, Kaufmann, and Nuerk (2011) found that 7-year-old children who committed more inversion errors in first grade made significantly more addition errors two years later. These studies clearly show an association between number writing and arithmetic. However, all three studies focused on children from a language with number word inversion. Number writing might have been a predictor of arithmetic in those studies because of the added syntactic complexity of number writing in languages with number word inversion. Inverted number words lead to more inversion errors in number transcoding in children and it is possible that this particular linguistic feature was driving the association. Indeed, in the longitudinal study by Moeller et al. (2011) inversion-related transcoding errors assessed in Year 1 were the only reliable predictor of mathematics grades two years

later. It is thus crucial to investigate the predictive power of number writing of two-digit numbers in a language without number word inversion (e.g., English).

Recent results by [Malone et al. \(2019\)](#) suggest that number writing might also be a predictor for English-speaking children. In their Australian sample children's performance on a latent factor 'number knowledge' which included children's performance in number writing of six numbers (including two two-digit numbers) at school entry (mean age 5 years) was a significant predictor of their arithmetic development over their first year in school.

#### 1.4. The current study

In the current study, we measured number writing ability and concurrent arithmetic performance in 5–7-year-old English- and German-speaking children in the first year of primary school. We chose to compare English and German because the particulars (e.g., number words from 1 to 12, decade numbers like *twenty - zwanzig*) and the construction of number words are highly similar in both languages and both have inverted teen number words from 13 to 19. However, number words larger than twenty are inverted in German but not in English [i.e., 'twenty-five' in English, but 'fünf-und-zwanzig' ('five-and-twenty') in German].

Our first aim was to shed more light onto the influence of language on transcoding of two-digit numbers. We predicted that the inverted structure of teen number words in German and English would lead to inversion errors in both language groups. In line with previous research, we predicted that German-speaking children would show more inversion errors for two-digit numbers larger than 20 and fewer non-inversion errors for two-digit numbers larger than 20, than English-speaking children. We also expected the German-speaking children to make fewer inversion errors for teen number words than English-speaking children, because number word inversion is exceptional in English, but the norm for two-digit numbers in German.

Our second aim was to evaluate the relationship between number writing and arithmetic in more detail. In particular, we aimed to compare the relative contribution of two-digit number writing as a concurrent predictor of arithmetic across the two languages, German with inverted number words and English with non-inverted number words. This will enable us to isolate whether two-digit number writing per se is a predictor of arithmetic or whether two-digit number writing is only a predictor in languages with the added syntactic complexity of number word inversion. We predicted that number writing of two-digit numbers would be a significant predictor of concurrent performance in arithmetic in both languages.

Domain-general skills contribute towards number writing performance and arithmetic (e.g., [Bull, Espy, & Wiebe, 2008](#); [De Smedt et al., 2009](#); [Raghubar, Barnes, & Hecht, 2010](#)), and thus could drive any association between them. Therefore, we will also investigate the contribution of domain-general skills to number writing in the current study for both language groups. Previous studies have reported an association between number writing and visuo-spatial working memory (WM) ([Simmons et al., 2012](#); [van der Ven et al., 2016](#); [Zuber et al., 2009](#)), verbal WM (e.g., [Simmons et al., 2012](#)) and the central executive ([Camos, 2008](#); [Imbo et al., 2014](#); [Moura et al., 2013](#); [Pixner et al., 2011](#); [van der Ven et al., 2016](#); [Zuber et al., 2009](#)). It is therefore crucial to control for variation in those domain-general skills when investigating whether number writing is a predictor of arithmetic performance. Thus, our second prediction was that two-digit number writing would be a significant predictor of concurrent performance in arithmetic in both languages, even when we control for domain-general predictors of arithmetic and number writing.

## 2. Material and methods

### 2.1. Participants

Four hundred and eighty-six children in Year 1 (177 German-speaking, 83 females; 309 English-speaking children, 142 females) were recruited from primary schools across Yorkshire (UK) and in Graz (Austria). Children in Austria came from a middle-income urban school district. Children in the UK came from four urban, three town and four rural schools, with a mean deprivation index decile score of 8 (indicating the 30 % of least deprived neighbourhoods, [Department for Communities & Local Government, 2015](#)) and an average of 11 % of free school meals.

In both countries we tested children (mean age of German-speaking children: 7 years and 1 month,  $SD = 3.46$  months; mean age of English-speaking children: 6 years and 2 months,  $SD = 3.70$  months) towards the end of Year 1 (from April to July). The groups were matched on duration of formal education. In both countries children are taught explicitly about writing double-digit numbers in Year 1. In the United Kingdom, by the end of Year 1 children are expected to be able to count, read and write numbers to 100 and to add and subtract one- and two-digit numbers to 20 ([Department for Education, 2013](#)). The curriculum in Austria ([Bundesministerium Bildung, Wissenschaft und Forschung, 2012](#)) does not specify separate learning goals for Year 1 and Year 2, but states that by the end of Year 2 children are expected to be able to count, read and write numbers to 100 and perform calculations up to 100. Most of the mathematical textbooks used in Year 1 in Austria include numbers up to 100.

The University of York and the University of Graz Psychology Department Ethics Committees granted ethical approval for the study. Head teachers and parents provided consent for children to take part.

### 2.2. Tasks and stimuli

Children were assessed on number writing, arithmetic and non-verbal reasoning as part of a larger battery of measures administered in a fixed order in a classroom setting in two 1-h sessions. Working memory was tested individually. All tests were administered in English for the English-speaking and in German for the German-speaking sample by researchers and trained research assistants.

### 2.2.1. Number writing

Number writing performance for teens and two-digit numbers larger than 20 (referred to as ‘XX numbers’ hereafter) excluding decade numbers was measured as part of a number writing task with a total of 52 items distributed evenly across four subtests. Subtest 1 and 2 were given in session 1, subtest 3 and 4 in session 2. This measure included four teen numbers and sixteen XX numbers (in bold): Subtest 1: 1, **11**, 80, **73**, **42**, **34**, **81**, **32**, **89**, **53**, 700, 203, 1300; Subtest 2: 7, **15**, 40, **64**, 300, 560, 340, 107, 242, 349, 3791, 1002, 1060; Subtest 3: 4, **16**, 70, **25**, **68**, **56**, **91**, **48**, **27**, **79**, 200, 304, 8000; Subtest 4: 6, **13**, 30, **47**, 600, 190, 220, 109, 123, 643, 2150, 1015, 2609.

Children were instructed to write Arabic digits to dictation, one item per line, using consecutive lines. In the current study we only analysed responses to two-digit numbers. Items were scored correct or incorrect (1, 0) and a total correct score for XX and teen numbers was calculated. Data for one item (11) was excluded from the teen data, as this is an exceptional number word in both English and German and cannot be inverted. Parallel forms reliability was .88 for the English-speaking and .89 for the German-speaking sample.

### 2.2.2. Arithmetic

Children completed the Numerical Operations subtest from the Wechsler Individual Achievement Test (WIAT-II UK; Wechsler, 2005) adapted for group use. This measure includes six items that involve identifying and writing Arabic digits and counting dots. The remaining nine items are standard arithmetic calculations (addition, subtraction and multiplication) increasing in difficulty. The researcher dictated the first six items and children were allowed 15 min to work through the remaining nine items. The format of questions 7–15 was changed from vertical to horizontal to reflect the format more commonly used in primary school. The number of correct calculations was scored. Responses to the first six items were excluded from the total score; the total correct score for items 7–15 provides a conventional measure of arithmetic performance. Cronbach’s alpha was .71 for the English-speaking and .60 for the German-speaking sample.

### 2.2.3. Non-verbal reasoning

Children completed sets A–C of Ravens Standard Progressive Matrices Plus (Raven, Raven, & Court, 1998) adapted for group use. The task comprised of two practice items followed by 34 test items increasing in difficulty. On each item, children were required to study a picture and to select the correct missing piece from a selection of six or eight possible answers. Following the practice trials completed as a class, children were allowed 15 min to complete the test items. The number of correct test items was scored (maximum = 34). Cronbach’s alpha was .76 for the English-speaking and .77 for the German-speaking sample.

### 2.2.4. Working memory

The Forward and Backward Digit Recall and Forward Block Recall subtests from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001) were administered individually. An additional Backward Block Recall task was administered; this task used the items from the Forward Block Recall task presented in a different order. Each subtest began with three practice trials to determine the beginning span of items (one, two or three). In the Forward Digit subtest, children were asked to repeat, in the same order, a verbal number string dictated by the researcher. In the Backward Digit subtest, children were asked to repeat the number string in a backwards order. In the Block Recall subtest, children were presented with a CORSI block arrangement with nine blocks and asked to watch the researcher tap a sequence of blocks, and then to tap the block sequence in the same order. In the Backward Block Recall task, children were asked to tap the sequence in a backwards order. For all subtests, children continued to the next span if they correctly recalled at least two out of four consecutive items in a given span and testing was discontinued following three incorrect trials in a given span.

The number of total test items correct was calculated for each subtest (maximum = 54 for Digit Forward, 36 for Digit Backward, 54 for Block Forward and 48 for Block Backward). The total score for the Forward Digit subtest provided a measure of verbal working memory, and the total score for Block Recall provided a measure of visuo-spatial working memory. A measure of central executive function was calculated by averaging z-scores (standardised within language group) from the Backward Digit and Backward Block

**Table 1**  
Descriptive statistics for each language group.

Variable	German-speaking		English-speaking		Group differences				
	Mean	Std. Dev.	Mean	Std. Dev.		p-value*	Effect of language with age controlled**		
Gender (% female)		46 %		46 %	$\chi^2(1) = .026$	.872			
Age (months)	177	85.72	3.47	309	74.87	3.70	3.02	< .001	
Arithmetic	177	5.45	1.41	309	4.10	2.01	.78	< .001	.313
Non-verbal reasoning (IQ)	177	17.94	4.74	307	14.50	4.47	.75	< .001	.788
Forward digit recall (verbal WM)	176	25.49	3.61	309	25.69	4.05	.05	.591	.758
Block recall (visuo-spatial WM)	176	22.45	3.35	309	20.08	4.28	.62	< .001	.142
Backward digit recall	176	10.01	2.64	309	8.19	2.99	.65	< .001	.100
Backward block recall	176	13.48	4.66	307	10.48	4.58	.65	.001	.212

\* p-value associated with independent samples t-test (except for gender where it is the p-value for the Chi-Square test), Bonferroni corrected p-value is  $p < .007$ .

\*\* p-value associated with language coefficient in simultaneous regression analysis with age controlled.

Recall tasks (e.g., Göbel, Moeller et al., 2014; Göbel, Watson et al., 2014; Moeller et al., 2011; Rasmussen & Bisanz, 2005; Zuber et al., 2009).

For the English-speaking children Cronbach's  $\alpha$  was .87 for the Forward Digit recall, .82 for the Backward Digit recall, .86 for the Forward Block recall and .87 for Backward Block recall. Item-level data was not available for the German-speaking children. Test-retest reliability reported for this age group in the test manual (Pickering & Gathercole, 2001) is .81 for the Forward Digit recall, .53 for the Backward Digit recall and .53 for the Forward Block recall.

### 3. Results

The data for this study are available at <https://reshare.ukdataservice.ac.uk/854398/>. German-speaking children were significantly older than English-speaking children which is reflected in overall higher performance in German-speaking children (see Table 1). After statistically controlling for differences in age between the two language groups, these group differences on measures of arithmetic, non-verbal reasoning and working memory were no longer significant. Correlation matrices are reported in Appendices B and C. The distribution of all variables - with the exception of two-digit number writing accuracy in the English-speaking sample - did not differ substantially from normality.

#### 3.1. Number writing accuracy

Overall, accuracy was high. German-speaking children transcribed 83.71 % = 25.30 % ( $SD = 25.30$  %) of the two-digit numbers correctly while the overall accuracy of English-speaking children was 93.71 % ( $SD = 13.56$  %). Fig. 1 shows the accuracy for teens and two-digit numbers larger than 20 (XX numbers) for both language groups. The English-speaking group demonstrated higher accuracy for XX numbers relative to teen numbers whereas the German-speaking group showed the opposite pattern.

A repeated-measures ANCOVA<sup>1</sup> was performed on the average percentage correct for number writing with number type (teens, XX) as the within-subject factor, language (German-speaking, English-speaking) as the between-subject factor and age entered as a covariate. Results revealed a significant main effect of language ( $F(1,483) = 12.76, p < .001, \eta_p^2 = .026$ ), indicating lower accuracy in the German-speaking children, and no significant main effect of number type ( $F(1,483) = 2.71, p = .101, \eta_p^2 = .006$ ). There was a significant language by number type interaction ( $F(1,483) = 13.90, p < .001, \eta_p^2 = .028$ ). German-speaking children showed significantly higher accuracy for writing teen numbers ( $M = .95, SD = .18$ ) than XX numbers ( $M = .82, SD = .29$ ),  $t(176) = 6.11, p < .001$ , *Cohen's d* = .92. In contrast, English-speaking children showed significantly higher accuracy for writing XX numbers ( $M = .95, SD = .13$ ) than for teens ( $M = .87, SD = .25$ ),  $t(308) = -6.24, p < .001$ , *Cohen's d* = .71.

#### 3.2. Inversion errors

Incorrect teen and XX items were coded as either inversion or non-inversion errors. For inversion errors, the verbal order of the decade and unit is reversed in symbolic notation, such as writing 54 for the item "forty-five" or 61 for the item "sixteen". Overall, children 78 numbers, made 506 inversion errors and 390 non-inversion errors.

Overall, the occurrence of inversion errors as a proportion of all errors made in writing two-digit numbers differed significantly across the language groups,  $\chi^2(1) = 325.99, p < .001$  (see Table 2). There were 430 inversion errors (81 %) compared to 98 non-inversion errors (19 %) in the German-speaking group, whereas in the English-speaking group there were 76 inversion errors (21 %) compared to 292 non-inversion errors (79 %). Typical examples of non-inversion errors include lexical errors (e.g., 45 for "twenty-five", 53.6 % of non-inversion errors [English-speaking: 58.9 %; German-speaking: 37.7 %]), syntactic errors (e.g., 302 for "thirty two", 33.1 % of non-inversion errors [English-speaking: 31.2 %; German-speaking: 38.8 %]) combination and other errors (e.g., 347 for "forty seven", 13.3 % of non-inversion errors [English-speaking: 9.9 %; German-speaking: 23.5 %]).

However, when investigating inversion errors separately for teen and XX items, the proportion of inversion errors for teen items did not differ significantly across the language groups,  $\chi^2(1) = .272, p = .602$ . By contrast, the proportion of inversion errors differed significantly for XX items across language groups,  $\chi^2(1) = 335.77, p < .001$ . German-speaking children made more inversion errors than non-inversion errors, whereas the English-speaking children made more non-inversion errors than inversion errors.

The German-speaking children had a higher percentage of inversion errors for XX numbers than for teens and a higher percentage of non-inversion errors for teens than for XX numbers,  $\chi^2(1) = 53.77, p < .001$ . The English-speaking group showed the opposite pattern; children had a higher percentage of inversion errors for teens than XX numbers and a higher percentage of non-inversion errors for XX numbers than teens,  $\chi^2(1) = 13.55, p < .001$ .

<sup>1</sup> A sensitivity analysis was performed in G\*Power (Faul, Erdfelder, Buchner, & Lang, 2009) to evaluate whether our sample size was sufficiently large to detect the interaction of interest in the ANCOVA on number type. We computed the minimal detectable effect size given an alpha level of .05, power of .95 and the sample size of 486 participants. Results yielded a minimal detectable effect size  $f^2 = .16$  which corresponds to  $\eta_p^2 = .026$ . This effect size is smaller than the one found in our analysis. This sensitivity analysis therefore shows that our sample size was adequately large to detect the main result of interest for this study.

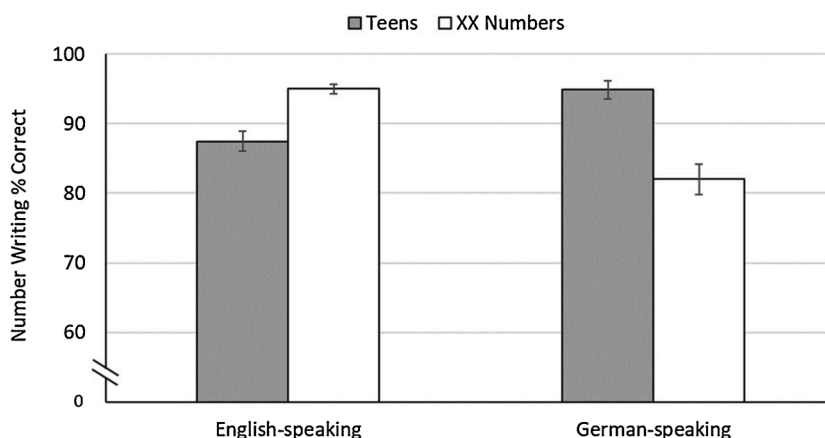


Fig. 1. Number writing accuracy for teen and other two-digit (XX) numbers across language groups (error bars represent standard errors).

Table 2

Number (and percentage) of errors for teen (1X) and double-digit (XX) items by language group.

	Teen items (1X)			Double-digit items (XX)			Total (1X + XX)		
	Inversion	Non-inversion	Total	Inversion	Non-inversion	Total	Inversion	Non-inversion	Total
English-speaking	37 (32 %)	78 (68 %)	115	39 (15 %)	214 (85 %)	253	76 (21 %)	292 (79 %)	368
German-speaking	7 (27 %)	19 (73 %)	26	423 (84 %)	79 (16 %)	502	430 (81 %)	98 (19 %)	528
Total	44 (31 %)	97 (69 %)	141	462 (61 %)	293 (39 %)	755	506 (56 %)	390 (44 %)	896

### 3.3. Predictors of transcoding accuracy for two-digit number writing

Number writing accuracy for teen and XX items were combined and children were categorised into two groups depending on whether they answered all number writing items correctly or not. A binomial logistic regression<sup>2</sup> was used to ascertain the effects of age, non-verbal reasoning, verbal WM, visuo-spatial WM and central executive on the likelihood of accurate two-digit number writing in each language group (see Table 3). Data from children with missing data (4 English-speaking children, 1 German-speaking child) were excluded from all our regression analyses.

The logistic regression model was statistically significant for both German- ( $\chi^2(5) = 24.48, p < .001$ ) and English- ( $\chi^2(5) = 30.91, p < .001$ ) speaking children. For the German-speaking children the model explained 17.3 % (Nagelkerke  $R^2$ ) of the variance in number writing accuracy and correctly classified 67.0 % of cases. For the English-speaking children the model explained 13.1 % (Nagelkerke  $R^2$ ) of the variance in number writing accuracy and correctly classified 66.2 % of cases. In both language groups, increased CE skills were associated with an increased likelihood of accurate two-digit number writing. In addition, German-speaking children with higher non-verbal reasoning also had an increased likelihood of accurate two-digit number writing. Follow-up z-tests using pooled standard error terms (Brame, Paternoster, Mazerolle, & Piquero, 1998; Clogg, Petkova, & Haritou, 1995) indicated that there was neither a significant difference in the beta weight for central executive predicting number writing across the two language groups ( $z = 1.20, p = .770$ ) nor for non-verbal reasoning ( $z = 1.21, p = .774$ ).

### 3.4. Predictors of inversion errors in two-digit number writing

The number of inversion errors for teen and XX items were combined. Because the average number of inversion errors per child was low and many children made no inversion errors, a Poisson regression was used to predict the number of inversion errors in two-digit numbers based upon age, non-verbal reasoning, verbal and visuo-spatial working memory and central executive (see Table 4). We use the vce (robust) option in Stata to obtain robust standard errors for the parameter estimates as recommended by Cameron and Trivedi (2009) to control for mild violation of underlying assumptions.

<sup>2</sup> A sensitivity analysis was conducted in G\*Power to compute the required effect size (in odds ratio), given alpha = .05, power = .95 and sample size. Note that the probability of an event under the null hypothesis ( $p_1$ ) was calculated based on the intercept of the regression models without predictors, which corresponds to the number of children who wrote all numbers correct divided by the sample size ( $p_1$  in the German-speaking sample = .49;  $p_1$  in the English-speaking sample = .63). Results yielded a minimal detectable odds ratio of 1.90 for the German-speaking sample and of 1.61 for the English-speaking sample, which correspond to betas of .64 and .48, respectively. These results show that the sample size in the German-speaking sample was large enough to detect the contribution of the central executive (but not of non-verbal reasoning), whereas the required effect size for logistic regression in the English-speaking was just minimally larger than the one observed for the central executive.



**Table 3**

Binary logistic regression analysis predicting two-digit number writing accuracy for the two language groups.

Predictors of Number Writing (1X and XX)	German-speaking			English-speaking		
	$\beta$	SE( $\beta$ )	Exp( $\beta$ )	$\beta$	SE( $\beta$ )	Exp( $\beta$ )
Age	.04	.05	1.05	.06	.04	1.06
Non-verbal reasoning	.08*	.04	1.09	.02	.03	1.02
Verbal WM	.004	.05	1.00	.06	.04	1.06
Visuo-spatial WM	-.09	.06	.91	.05	.04	1.05
Central executive	.79**	.25	2.20	.42*	.18	1.52

\*  $p < .05$ .\*\*  $p < .01$ .

The incidence rate of inversion errors was higher for children with lower CE skills in both language groups. The incidence rate of inversion errors also increased for German-speaking with lower non-verbal reasoning and for English-speaking children with lower visuo-spatial WM.

### 3.5. The relationship between two-digit number writing and arithmetic

Number writing accuracy for teen and XX items were combined to evaluate the overall contribution of number writing. Hierarchical linear regression models<sup>3</sup> were used to compare predictors of arithmetic in each language group (see Table 5).

Domain-general predictors (age, non-verbal reasoning, verbal WM, visuo-spatial WM and central executive) were entered in the first step of the model. Together these predictors accounted for 20 % of the variance (adjusted  $R^2 = .17$ ) in arithmetic in the German-speaking group ( $F(5, 170) = 8.31, p < .001$ ) and 35 % of the variance (adjusted  $R^2 = .34$ ) in arithmetic in the English-speaking group ( $F(5, 299) = 31.82, p < .001$ ). In both language groups, children with better non-verbal reasoning and better CE performed better on arithmetic. In addition, English-speaking children with higher verbal and visuo-spatial WM also performed with higher accuracy. Follow-up z-tests using pooled standard error terms indicated that there were no significant differences in the beta weights for non-verbal reasoning, visuo-spatial or verbal working memory predicting arithmetic across the two language groups ( $z = .98, p = .164$ ;  $z = -.42, p = .337$ ;  $z = -1.13, p = .129$  respectively). There was a significant difference in the beta weights for CE predicting arithmetic across the two language groups ( $z = -1.69, p = .046$ ).

Entering two-digit number writing as a second step resulted in significant change in  $R^2$  for both the German ( $\Delta R^2 = 0.03, p = .019$ ) and English-speaking groups ( $\Delta R^2 = 0.02, p < .001$ ). Follow-up z-tests using pooled standard error terms indicated that there was no significant difference in the beta weight for number writing predicting arithmetic across the two language groups ( $z = -.54, p = .295$ ).

## 4. Discussion

This study investigated the writing of two-digit numbers in 5–7 year-old English- and German-speaking children during their first year of formal education. We compared these two language groups because German number words for 13–99 are inverted with respect to the Arabic notation. In contrast, in English, only teens (13–19) are inverted, but number words larger than twenty are not inverted.

First, we were interested in the influence of number word inversion at the item level on number writing. As in previous studies, we found that German-speaking children made more inversion errors for numbers larger than 20 than English-speaking children. In addition, as far as we know for the first time, we report inversion errors in English-speaking children. These errors occurred more often for number words that are inverted at the item level, i.e., teen numbers. Our study shows that an inverted spoken number word leads to more inversion errors, even in a language where most number words are not inverted.

Our second aim was to investigate the concurrent relationship between the ability to write two-digit numbers correctly from dictation and arithmetic performance in English- and German-speaking children. We investigated this relationship in both language groups in order to determine whether the relationship between transcoding ability and arithmetic reported in previous studies can be attributed to syntactic complexity, i.e., number word inversion. We found two-digit number writing to be a unique concurrent predictor of arithmetic in German and in English, showing that the relationship between transcoding ability and arithmetic is not specifically driven by number word inversion.

<sup>3</sup> The minimal detectable effect size for the contribution of two-digit number writing to arithmetic on the linear multiple regression model was calculated with a sensitivity analysis in G\*Power (alpha level = .05 and power = .95). Results yielded a minimal detectable effect size  $f^2 = .07$  for the regression on German-speaking children, which corresponds to  $R^2 = .07$ , and a minimal detectable effect size  $f^2 = .04$  for the regression on English-speaking children, which corresponds to  $R^2 = .04$ . The minimal detectable effect size for the regression on English-speaking children is comparable to the one reported in Table 5. The sample size of English-speaking children was thus adequately large to detect a significant  $\Delta R^2$ . The significant  $\Delta R^2$  found in the regression on German-speaking children, however, is slightly lower than the minimal detectable effect size found in our sensitivity analysis. These findings indicate that the regression on arithmetic in the German-speaking sample is not optimally powered.

**Table 4**

Poisson regression analysis predicting number of inversion errors in two-digit number writing for the two language groups.

Predictors of inversion errors	German-speaking group		English speaking group	
	Poisson regression coefficient	z	Poisson regression coefficient	z
Age	-.02	-.67	-.07	-1.71
Non-verbal reasoning	-.06	-2.16*	-.06	-1.55
Verbal WM	.01	.39	-.03	-.67
Visuo-spatial WM	-.03	-.71	-.07	-2.19*
Central executive	-.39	-2.10*	-.46	-2.44*
Wald $\chi^2(5)$	28.85***		63.99***	
N	176		305	

\* p &lt; .05.

\*\*\* p &lt; .001.

**Table 5**

Hierarchical multiple regression analyses predicting arithmetic from domain-general control variables and two-digit number writing accuracy for the two language groups.

Predictors of Arithmetic	German-speaking			English-speaking		
	$\Delta R^2$	$\beta$	t	$\Delta R^2$	$\beta$	t
<b>Step 1*</b>						
Age		.10	1.39		.08	1.59
Non-verbal reasoning		.24	3.06**		.15	2.79**
Verbal WM		.07	.91		.10	2.09*
Visuo-spatial WM		.004	.06		.11	2.07*
Central executive		.23	2.79**		.39	7.18***
<b>Step 2</b>						
Number writing	.03*	.17	2.37*	.04***	.22	4.23**
Total R <sup>2</sup>	.22***			.38***		
Total adjusted R <sup>2</sup>	.20			.37		
N	176			305		

\*  $\beta$ - and t-values reported for step 1 are from the first model.

\*\* p &lt; .01.

\*\*\* p &lt; .001.

#### 4.1. Inversion errors are item-level specific

Overall, German-speaking children transcribed less accurately and made more inversion errors than English-speaking children. This replicates numerous previous findings showing that in languages with number word inversion children struggle with this extra hurdle when writing numbers (e.g., Imbo et al., 2014; Krinzing et al., 2011; Moeller et al., 2015; Pixner et al., 2011; van der Ven et al., 2016; Zuber et al., 2009). Our findings suggest that more explicit instruction and practice even for two-digit number writing might be beneficial for children's mathematical development, in particular for languages with number word inversion. In our study, German-speaking children made more inversion errors for two-digit numbers larger than twenty than English-speaking children.

Previous cross-linguistic studies did not specifically investigate number writing of teens. Thus, our study is the first to compare directly number writing of teens in a language with number word inversion for all numbers 13–99 (German) to a language with no inversion except for teen items 13–19 (English). We tested children's number writing of nearly half of those relevant teen items, but the number of relevant teen items is low. While our findings are thus based on the writing of only three teen items per child, and a replication with a large number of teen items would be desirable, we partially compensate for this by having a large sample size. Interestingly, the number of inversion errors did not differ significantly between the German- and English-speaking children for teen numbers, which are inverted in English and German. English-speaking children, however, made significantly more errors when writing teen numbers compared to when they were writing two-digit numbers larger than twenty, which are not inverted in English. At first glance this finding may seem surprising because teens are learned earlier than numbers larger than twenty and by the end of Year 1 children in England are expected to be able to count, read, add and subtract two-digit numbers up to 20 (Department for Education, 2013). Our findings demonstrate clearly that the specific syntactic structure at the item level, in this case inverted teen number words in English, rather than the general rules for number word formation, is related to inversion errors.

This has important educational implications. Several previous studies have shown that number word inversion does not only affect number writing adversely but also arithmetic performance, even when a purely Arabic digit format is used (Göbel, Moeller et al., 2014; Lewis, Bahnmüller, Wesierska, Moeller, & Göbel, 2020; Lonnemann & Yan, 2015). It is possible that even in an otherwise non-inverted language such as English, number word inversion at the item-level, e.g., for teens, could adversely affect calculation performance. Indeed, in line with this prediction, Dowker, Bala, and Lloyd (2008) reported indirect effects of the teen number word structure on arithmetic performance in children. They compared number skills of Tamil- and/or English-speaking children living in

London. In Tamil number words for 11–19 are formed regularly and are non-inverted. Children speaking Tamil performed better than English-speaking children on written calculation items mainly involving sums between 11 and 19. The authors suggest that in their study the language differences for the teens might have led to the lower arithmetic performance in English-speaking than Tamil-speaking children for teen items. Our results support this suggestion and provide more direct evidence that it is indeed the structure of teen items that affects numerical performance of English-speaking children. Furthermore, Lewis et al. (2020) showed that even English-speaking adults are slowed down at the item-level by the inversion in teen numbers when they are performing an arithmetic verification task in purely Arabic format.

Thus, we suggest that English-speaking children would benefit from increased explicit teaching of teen numbers. The incorrect writing of these relatively low numbers suggest that some children are not yet secure in their knowledge of the place value system. Being able to accurately record numbers and having a solid place-value understanding are fundamental to mathematical success and therefore an increased emphasis on understanding this element of the number system might support children's future mathematical development.

In contrast to our prediction, while German-speaking children made overall less errors on teen items, the percentage of inversion errors for teen numbers did not differ significantly between German- and English-speaking children. This also highlights the importance of the number word structure at the item level. However, it is also possible that due to the relatively small number of errors for teen items in the German-speaking group we lack the power to detect between language differences for inversion errors for teens. The difference in accuracy for teen items between German- and English-speaking children could have been influenced by even finer grained differences in the number word structure. For example, in both English and German the number word for 13 is inverted, but in German the unit ('three') and decade constituents ('ten') are transparent and unchanged ('dreizehn', literally translated threeten) while in English the original lexical units are transformed ('thirteen'). Transparency in number word formation is often associated with better performance (e.g., Dowker et al., 2008; Siegler & Mu, 2008). Thus, it might be worthwhile to further investigate the effect of the transparency of the number word constituents in inverted number words in future studies with more teen items.

Moving on to non-inversion errors, as predicted, German-speaking children made significantly fewer non-inversion errors than English-speaking children for numbers larger than 20. This is in line with findings by Imbo et al. (2014) and Pixner et al. (2011) who also reported a smaller number of non-inversion errors for inverted than non-inverted languages. Number word inversion specifically affected only the number of inversion errors and did not lead to a general increase in non-inversion errors. This suggests that number word inversion might not influence other transcoding rules.

#### 4.2. Predictors of inversion errors

In both language groups children with lower Central Executive (CE) capacity made more inversion errors. This replicates and extends previous findings. The CE component of WM is crucial for manipulating content in WM (Baddeley & Hitch, 1974; Baddeley, 2002). While several studies report an association between better transcoding performance and higher capacity (Imbo et al., 2014; Pixner et al., 2011), as far as we are aware there is only one published study specifically investigating predictors of inversion errors in number writing. Zuber et al. (2009) investigated the contribution of different working memory components towards inversion errors in German-speaking children and found that children with lower CE made more inversion errors. Our results show that children with lower CE skills also made more inversion errors in a language where only teen numbers are inverted. It is possible that children with lower CE capacity struggle to simultaneously memorise the lexical elements of the number word and their sequence in the spoken number word while manipulating this sequence (i.e., inverting it). In order to compensate for lower capacity children might be writing down the lexical elements of the number word as soon as possible, i.e., in the order in which they appear in the spoken number word. This strategy would be clearly maladaptive in the case of inverted number words. We did not record planning and response times in the current study, so this suggestion is speculative. Future research should investigate this further.

In addition, in our study lower visuo-spatial WM was also a significant predictor of more inversion errors for the English-speaking children only. In contrast, for the German-speaking group, children with lower non-verbal reasoning skills made more inversion errors. Thus, our data strongly highlights the importance of individual differences in the central executive for children's ability to deal with inversion and hints at possible contributions from visuo-spatial working memory and non-verbal reasoning.

#### 4.3. Number writing accuracy

Individual differences in central executive skills also predicted the likelihood of accurate two-digit number writing. This is in line with several previous studies showing that children with better central executive skills make less mistakes in number writing (e.g., Pixner et al., 2011; Zuber et al., 2009). However, our findings are in contrast with Simmons et al. (2012), the only previous study to investigate predictors of transcoding in English-speaking children. In their study individual differences in central executive skills did not predict number writing accuracy when visuo-spatial WM was also included as predictor. However, individual differences in CE were significantly correlated with number writing accuracy as well as with individual differences in visuo-spatial WM. Given their small sample of Year 1 children ( $N = 41$ ) it is possible that their study lacked the power to detect significant independent contributions of CE skill.

Central executive skills predicted the likelihood of accurate number writing in our study also for English-speaking children, in a language where most two-digit numbers are not inverted. This shows that even towards the end of Year 1 in primary school children still require their central executive skills for writing two-digit numbers. Also, in our study the contribution of CE to the likelihood of accurate number writing was not significantly larger in German- than English-speaking children. This suggests, that at least for two-

digit numbers, the extra step involved in inversion does not lead to differences in CE demands. This result mirrors an earlier finding by Imbo et al. (2014) who found that children speaking an inverted (Dutch) or non-inverted (French) language relied on CE to a similar degree during transcoding. Thus, in sum, there is overwhelming evidence of the importance of central executive skills for number writing, even just for writing two-digit numbers.

Clearly, there might have been other factors influencing children's number writing performance that could potentially have differed between children and between our two language samples that were not measured [e.g., home numeracy (LeFevre et al., 2009), children's exposure to numbers, task familiarity].

#### 4.4. Implications for models of transcoding

Our results have at least two new implications for models of transcoding. First, in their ADAPT model Barrouillet et al. (2004) proposed that upon hearing the same spoken number word repeatedly children will quickly become able to retrieve the corresponding Arabic digit string from Long-term memory (LTM). Given that children in our study were at the end of Year 1 they should thus be able to retrieve the corresponding Arabic digit string from LTM at least for the teen numbers and perhaps also for many other two-digit numbers. However, children in our study still made mistakes in two-digit number writing, including teen items. They might have used retrieval but may not have been fully proficient yet at retrieving the correct Arabic digit strings from LTM or they could have stored representations incorrectly. However, given that many mistakes, even for two-digit items, were syntactic, our data thus show that at least on some items children still used production rules for two-digit number writing and did not retrieve the corresponding Arabic digit from LTM.

Furthermore, as pointed out in previous research on transcoding and number word inversion, existing transcoding models need to be adapted to be able to account for number word inversion. Previously it was highlighted that number word inversion is a feature of several languages (e.g., Arabic, Dutch, Danish, German, Malagasy, Maltese, and Czech; Comrie, 2005). The particular contribution of our study is to show that accounting for number word inversion in models for transcoding is not only essential for languages with general number word inversion but also for other mainly non-inverted number languages such as English because many languages nevertheless have a minority of items with number word inversion (often the teens). In sum, there is clearly a need to develop a hybrid model of number transcoding that can account for inversion errors. This model should also include a proposal for how number word inversion might affect the development of number writing.

#### 4.5. Number writing as predictor of arithmetic performance

Number writing is also an important foundation of arithmetic performance. This is the case even when domain-general factors are taken into account. In line with previous studies (e.g., Raghobar et al., 2010; Simmons et al., 2012), domain-general factors predicted a significant amount of variance in concurrent arithmetic performance for both German- and English-speaking children in our study. Higher performance in arithmetic in 5–7 year-olds in Year 1 was predicted by higher central executive skills and better non-verbal reasoning for both samples. In addition, stronger visuo-spatial and verbal WM was also related to better arithmetic performance in English-speaking children. It is possible that visuo-spatial and verbal WM are more important for arithmetic performance in younger children (e.g., Bull et al., 2008; De Smedt et al., 2009) and that could account for the differences between the language groups. However, when we directly compared the weight of the influence of visuo-spatial and verbal WM on arithmetic performance between the two language groups, no significant group differences emerged. Overall, this replicates (e.g., Bull et al., 2008; Raghobar et al., 2010) and highlights the importance of working memory for arithmetic performance.

In addition, overall accuracy in two-digit number writing was a unique concurrent predictor of arithmetic performance in both German- and English-speaking children. Our study is the first to show that number writing predicts additional unique variance in arithmetic in both inverted and non-inverted languages, even after controlling for a range of domain-general skills that contribute towards both number writing and arithmetic (e.g., Lopes-Silva, Moura, Julio-Costa, Haase, & Wood, 2014; Simmons et al., 2012; van der Ven et al., 2016; Zuber et al., 2009). This is a strong test of the contribution of number writing because – as discussed above – number writing itself is predicted by those domain-general factors and highlights the domain-specific contribution of number writing.

Evidence of a relationship of similar size between number writing and arithmetic in both German and in English, a language with very few inverted number words, also strongly suggests that the relationship between number writing and arithmetic is not driven by the additional processing demands of an inverted number word system. Our study thus extends previous research, which has until now been limited to children who speak a language with number word inversion (Imbo et al., 2014; Moeller et al., 2011; van der Ven et al., 2016).

Our study adds to the emerging evidence that the ability to translate between verbal number words and Arabic digits provides an important foundation in numerical development (e.g., Göbel, Moeller et al., 2014; Göbel, Watson et al., 2014; Malone et al., 2019; Purpura, Baroody, & Lonigan, 2013). While this study has focused solely on one transcoding direction, from verbal number words to Arabic digits, it is very likely that similar results would have been found had we investigated number naming. Performance in number naming and number writing are highly correlated in early primary school and accuracy patterns of number naming show similar effects of number structure (and inversion) as number writing (Steiner et al., 2020). Our findings also suggest that the mechanism underlying the relationship between transcoding and arithmetic at least in part independent from general ability and working memory. Potential candidate explanations for this mechanism are single-digit knowledge and place-value understanding.

For reading acquisition, letter knowledge and the association between single letters and sounds are crucial in predicting reading success (Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012). Analogous to this, Arabic digits (0–9) are the basic building blocks

for all Arabic multi-digit strings, and so variation in early single-digit knowledge might similarly underlie arithmetic development (Knudsen, Fischer, & Aschersleben, 2015). There is clear evidence that the efficiency of matching between spoken number words and Arabic digits is still developing in children in primary school and varies between children within the same school year (Lyons, Price, Vaessen, Blomert, & Ansari, 2014), but the empirical support for a relationship with arithmetic is mixed. Sasanguie and Reynvoet (2014) showed that faster association of single-digit number words with Arabic digits is related to arithmetic performance in adults. However, follow-up studies failed to fully replicate this finding (Lin & Göbel, 2019; Sasanguie, Lyons, De Smedt, & Reynvoet, 2017) and a cross-sectional study of children in Years 1–6 did not find a significant association between matching efficiency and arithmetic performance (Lyons et al., 2014).

Alternatively, place-value understanding might be more important for accurate number transcoding. For multi-digit number transcoding children need to understand the principle of place-value: that the meaning of each single Arabic digit depends on its location in a digit string. Evidence suggests that individual differences in early place-value understanding are apparent even prior to any formal schooling on multi-digit numbers (Mix, Prather, Smith, & Stockton, 2014; Yuan et al., 2019) and understanding of this principle in first graders predicts later performance in arithmetic (Moeller et al., 2011).

Non-inversion errors committed by children in our study included both lexical and syntactic errors for both language-groups. This indicates that even at the end of Year 1 children are still – at least occasionally – struggling with both the lexical and syntactic aspects of two-digit number writing. An important goal of future research will be to tease apart components of transcoding using longitudinal studies to understand the mechanism driving this relationship with mathematical ability.

## 5. Conclusions

In summary, while the occurrence of inversion errors was driven by the number word structure at the item-level and thus varied across our two language groups, overall number writing performance for two-digit numbers was a significant concurrent predictor of arithmetic performance over and above domain-general predictors for both languages, with and without number word inversion.

### CRedit authorship contribution statement

**Francina J. Clayton:** Data curation, Formal analysis, Investigation, Project administration, Resources, Visualization, Writing - original draft. **Clare Copper:** Data curation, Investigation, Resources, Writing - review & editing. **Anna F. Steiner:** Data curation, Resources, Writing - review & editing. **Chiara Banfi:** Data curation, Investigation, Project administration, Resources, Writing - review & editing. **Sabrina Finke:** Data curation, Investigation, Resources, Writing - review & editing. **Karin Landerl:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing - review & editing. **Silke M. Göbel:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

### Declaration of Competing Interest

The authors report no declarations of interest.

### Acknowledgements

This work was supported by the ESRC, UK [grant number ES/N014677/1] and the FWF, Austria [grant number I 2778-G16]. We would like to thank Erin Dysart, Christina Roberts, Rebecca Reed, Philippa Gibbons and Viktoria Wirth for help with data collection.

### Appendix A. Overview of studies investigating the relation between number writing/number knowledge and arithmetic

Study	Task	Items
Göbel, Watson et al. (2014)	Target number was presented verbally, children had to identify corresponding Arabic numeral out of a choice of 4 or 5	Target numbers: 6, 14, 28, 52, 76, 163, 235, 427
Habermann et al. (2020)	Number writing	Target numbers: 2, 9, 7, 4, 8, 10, 6, 1, 20, 3, 100, 5
Malone et al. (2019) Sowinski et al. (2015)	Number writing Number system knowledge: Numeration subtest of KeyMath test-Revised, digit naming, name next digit	Target numbers: 2, 7, 13, 28, 69, 145 Numbers used not stated in the method section, but Numeration subtest uses numbers 0–999, example given for digit naming is a three-digit number
Imbo et al. (2014)	Number writing	Five 1-digit, <b>twenty 2-digit</b> and forty 3-digit numbers, math grades only available for Dutch-speaking children – compared math grade for more-skilled versus less-skilled transcoders based on their overall errors in transcoding
van der Ven et al. (2016)	Number writing in Math Garden	All <b>two-digit items (10-99)</b> and 290 three-digit numbers, direct effect of transcoding ability (across all items) on addition
Moeller et al. (2011)	Number writing	Four 1-digit numbers, <b>20 2-digit numbers</b> , 40 three-digit numbers

**Appendix B. Simple and partial correlations (controlling for age, above diagonal) for the English-speaking group**

	1.	2.	3.	4.	5.	6.	7.
1. Age							
2. Number writing	0.16**						
3. Arithmetic	0.21***	0.45***					
4. Non-verbal reasoning	0.27***	0.26***	0.36***				
5. Verbal WM	0.02	0.30***	0.29***	0.23***			
6. Visuo-spatial WM	0.12*	0.37***	0.36***	0.35***	0.23***		
7. Central executive	0.19***	0.40***	0.53***	0.32***	0.30***	0.43***	

\*\*\* $p < .001$ , \*\* $p < .01$ , \* $p < .05$

**Appendix C. Simple and partial correlations (controlling for age, above diagonal) for the German-speaking group**

	1.	2.	3.	4.	5.	6.	7.
1. Age							
2. Number writing	0.04						
3. Arithmetic	0.15*	0.30***					
4. Non-verbal reasoning	0.18*	0.27***	0.37***				
5. Verbal WM	-0.16*	0.10	0.16*	0.22**			
6. Visuo-spatial WM	0.13	0.22**	0.20**	0.33***	0.17*		
7. Central executive	0.11	0.32***	0.36***	0.41***	0.27***	0.41***	

\*\*\* $p < .001$ , \*\* $p < .01$ , \* $p < .05$

**References**

- Steiner, A. F., Finke, S., Clayton, F. J., Banfi, C., Kemény, F., Göbel, S. M., & Landerl, K. (2020). Language effects in early development of number writing and reading. *Journal of Numerical Cognition*.
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7(2), 85.
- Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (vol. 8). New York: Academic Press.
- Bahnmueller, J., Nuerk, H. C., & Moeller, K. (2018). A taxonomy proposal for types of interactions of language and place-value processing in multi-digit numbers. *Frontiers in Psychology*, 9(1024), 1024.
- Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004). ADAPT: A developmental, asemantic, and procedural model for transcoding from verbal to Arabic numerals. *Psychological Review*, 111, 368–394. <https://doi.org/10.1037/0033-295X.111.2.368>
- Barrouillet, P., Thevenot, C., & Fayol, M. (2010). Evidence for knowledge of the syntax of large numbers in preschoolers. *Journal of Experimental Child Psychology*, 105(3), 264–271. <https://doi.org/10.1016/j.jecp.2009.10.005>
- Brame, R., Paternoster, R., Mazerolle, P., & Piquero, A. (1998). Testing for the equality of maximum-likelihood regression coefficients between two independent equations. *Journal of Quantitative Criminology*, 14(3), 245–261.
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, 33(3), 205–228. <https://doi.org/10.1080/87565640801982312>
- Bundesministerium Bildung, Wissenschaft und Forschung. (2012). *Lehrplan der Volksschule*. Available at [https://www.bmbwf.gv.at/Themen/schule/schulpraxis/lp/lp\\_vs.html](https://www.bmbwf.gv.at/Themen/schule/schulpraxis/lp/lp_vs.html).
- Byrge, L., Smith, L. B., & Mix, K. S. (2014). Beginnings of place value: How preschoolers write three-digit numbers. *Child Development*, 85(2), 437–443. <https://doi.org/10.1111/cdev.12162>
- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using stata*. College Station (TX): Stata Press.
- Camos, V. (2008). Low working memory capacity impedes both efficiency and learning of number transcoding in children. *Journal of Experimental Child Psychology*, 99(1), 37–57.
- Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical-methods for comparing regression-coefficients between models. *The American Journal of Sociology*, 100(5), 1261–1293.
- Comrie, B. (2005). *The world atlas of language structures*. Oxford: Oxford University Press.
- De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, 103(2), 186–201. <https://doi.org/10.1016/j.jecp.2009.01.004>
- Department for Communities and Local Government. (2015). *The English indices of deprivation 2015 statistical release*. Available at <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015>.
- Department for Education. (2013). *The national curriculum in England: Key stages 1 and 2 framework document*. Available at <https://www.gov.uk/government/publications/national-curriculum-in-england-primary-curriculum>.
- Dowker, A., & Nuerk, H. C. (2016). Editorial: Linguistic influences on mathematics. *Frontiers in Psychology*, 7(1035), 1035.
- Dowker, A., Bala, S., & Lloyd, D. (2008). Linguistic influences on mathematical development: How important is the transparency of the counting system? *Philosophical Psychology*, 21(4), 523–538. <https://doi.org/10.1080/09515080802285511>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology*, 77(3), 236–263. <https://doi.org/10.1006/jecp.2000.2561>
- Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology*, 74(3), 213–239.
- Göbel, S. M., Moeller, K., Kaufmann, L., Pixner, S., & Nuerk, H. C. (2014). Language affects symbolic arithmetic in children: The case of number word inversion. *Journal of Experimental Child Psychology*, 119, 17–25. <https://doi.org/10.1016/j.jecp.2013.10.001>

- Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science*, 25(3), 789–798.
- Habermann, S., Donlan, C., Göbel, S. M., & Hulme, C. (2020). The critical role of Arabic numeral knowledge as a longitudinal predictor of formal arithmetic development. *Journal of Experimental Child Psychology*, 193. <https://doi.org/10.1016/j.jecp.2019.104794>
- Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology*, 93(3), 615–626. <https://doi.org/10.1037/0022-0663.93.3.615>
- Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science*, 23(6), 572–577.
- Imbo, I., Vanden Bulcke, C., De Brauwer, J., & Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Frontiers in Psychology*, 5(313), 313.
- Johansson, B. S. (2005). Numerical writing skill and elementary arithmetic mental calculations. *Scandinavian Journal of Educational Research*, 49(1), 3–25. <https://doi.org/10.1080/0031383042000302119>
- Knudsen, B., Fischer, M., & Aschersleben, G. (2015). The development of Arabic digit knowledge in 4-to-7-year-old children. *Journal of Numerical Cognition*, 1(1), 21–37.
- Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H. C., & Willmes, K. (2011). Differential language effects on numerical skills in second grade. *Journal of Cross-Cultural Psychology*, 42(4), 614–629.
- LeFevre, J. A., Skwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(2), 55.
- Lewis, C. A., Bahmuel, J., Wesierska, M., Moeller, K., & Göbel, S. M. (2020). Inversion effects on mental arithmetic in English- and Polish-speaking adults. *The Quarterly Journal of Experimental Psychology*, 73(1), 91–103. <https://doi.org/10.1171/1747021819991983>
- Lin, C. J., & Göbel, S. M. (2019). Arabic digits and spoken number words: Timing modulates the cross-modal numerical distance effect. *The Quarterly Journal of Experimental Psychology*, 72(11), 2632–2646. <https://doi.org/10.1177/1747021819854444>
- Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends in Neuroscience and Education*, 4(1-2), 1–5.
- Lopes-Silva, J. B., Moura, R., Julio-Costa, A., Haase, V. G., & Wood, G. (2014). Phonemic awareness as a pathway to number transcoding. *Frontiers in Psychology*, 5, 13.
- Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, 17(5), 714–726.
- Malone, S. A., Burgoyne, K., & Hulme, C. (2019). Number knowledge and the approximate number system are two critical foundations for early arithmetic development. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000426>
- McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, 44(1-2), 107–157.
- McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, 4, 171–196. [https://doi.org/10.1016/0278-2626\(85\)90069-7](https://doi.org/10.1016/0278-2626(85)90069-7)
- Mix, K. S., Prather, R. W., Smith, L. B., & Stockton, J. D. (2014). Young children's interpretation of multidigit number names: From emerging competence to mastery. *Child Development*, 85(3), 1306–1319.
- Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H. C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities*, 32(5), 1837–1851.
- Moeller, K., Zuber, J., Olsen, N., Nuerk, H. C., & Willmes, K. (2015). Intransparent German number words complicate transcoding - a translingual comparison with Japanese. *Frontiers in Psychology*, 6(740), 740.
- Moura, R., Lopes-Silva, J. B., Vieira, L. R., Paiva, G. M., Prado, A. C., Wood, G., et al. (2015). From 'five' to 5 for 5 minutes: Arabic number transcoding as a short, specific, and sensitive screening tool for mathematics learning difficulties. *Archives of Clinical Neuropsychology*, 30(1), 88–98. <https://doi.org/10.1093/arclin/acu071>
- Moura, R., Wood, G., Pinheiro-Chagas, P., Lonnemann, J., Krinzinger, H., Willmes, K., et al. (2013). Transcoding abilities in typical and atypical mathematics achievers: The role of working memory and procedural and lexical competencies. *Journal of Experimental Child Psychology*, 116(3), 707–727.
- Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C)*. London, UK: Psychological Corporation.
- Pixner, S., Zuber, J., Hermanova, V., Kaufmann, L., Nuerk, H. C., & Moeller, K. (2011). One language, two number-word systems and many problems: Numerical cognition in the Czech language. *Research in Developmental Disabilities*, 32(6), 2683–2689.
- Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, 5(3), 237–254.
- Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology*, 105(2), 453–464.
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, 20(2), 110–122. <https://doi.org/10.1016/j.lindif.2009.10.005>
- Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology*, 91, 137–157.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven's standard progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.
- Sasanguie, D., & Reynvoet, B. (2014). Adults' arithmetic builds on fast and automatic processing of Arabic digits: Evidence from an audiovisual matching paradigm. *PLoS One*, 9(2), Article e87739.
- Sasanguie, D., Lyons, I. M., De Smedt, B., & Reynvoet, B. (2017). Unpacking symbolic number comparison and its relation with arithmetic in adults. *Cognition*, 165, 26–38.
- Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *The British Journal of Developmental Psychology*, 12(3), 281–300. <https://doi.org/10.1111/j.2044-835X.1994.tb00635.x>
- Siegler, R. S., & Mu, Y. (2008). Chinese children excel on novel mathematics problems even before elementary school. *Psychological Science*, 19(8), 759–763.
- Simmons, F. R., Willis, C., & Adams, A. M. (2012). Different components of working memory have different relationships with different mathematical skills. *Journal of Experimental Child Psychology*, 111(2), 139–155.
- Sowinski, C., LeFevre, J. A., Skwarchuk, S. L., Kamawar, D., Bisanz, J., & Smith-Chant, B. (2015). Refining the quantitative pathway of the Pathways to Mathematics model. *Journal of Experimental Child Psychology*, 131, 73–93.
- van der Ven, S. H. G., Klaiber, J. D., & van der Maas, H. L. J. (2016). Four and twenty blackbirds: How transcoding ability mediates the relationship between visuospatial working memory and math in a language with inversion. *Educational Psychology*, 37(4), 487–505.
- van Loosbroek, E., Dirx, G. S. M. A., Hulstijn, W., & Janssen, F. (2009). When the mental number line involves a delay: The writing of numbers by children of different arithmetical abilities. *Journal of Experimental Child Psychology*, 102(1), 26–39. <https://doi.org/10.1016/j.jecp.2008.07.003>
- Van Rinsveld, A., & Schiltz, C. (2016). Sixty-twelve = seventy-two? A cross-linguistic comparison of children's number transcoding. *The British Journal of Developmental Psychology*, 34, 461–468. <https://doi.org/10.1111/bjdp.12151/full>
- Verguts, T., & Fias, W. (2006). Lexical and syntactic structures in a connectionist model of reading multi-digit numbers. *Connection Science*, 18, 265–283. <https://doi.org/10.1080/09540090600639396>
- Wechsler, D. (2005). *Wechsler individual achievement test – Second UK edition (WIAT-IIUK)*. London, England: Harcourt Assessment.
- Yuan, L., Prather, R. W., Mix, K. S., & Smith, L. B. (2019). Preschoolers and multi-digit numbers: A path to mathematics through the symbols themselves. *Cognition*, 189, 89–104. <https://doi.org/10.1016/j.cognition.2019.03.013>
- Zuber, J., Pixner, S., Moeller, K., & Nuerk, H. C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, 102(1), 60–77.