








Language Effects in Early Development of Number Writing and Reading

Anna F. Steiner^{1,2} , Sabrina Finke¹ , Francina J. Clayton³ , Chiara Banfi^{1,4} , Ferenc Kemény¹ ,
Silke M. Göbel^{3,5} , Karin Landerl^{1,6,7} 

[1] Institute of Psychology, University of Graz, Graz, Austria. [2] FH JOANNEUM, University of Applied Sciences, Graz, Austria. [3] Department of Psychology, University of York, York, United Kingdom. [4] Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria. [5] Department of Special Needs Education, University of Oslo, Oslo, Norway. [6] BioTechMed-Graz, Graz, Austria. [7] Department of Cognitive Science, Macquarie University, Sydney, NSW, Australia.

Journal of Numerical Cognition, 2021, Vol. 7(3), 368–387, <https://doi.org/10.5964/jnc.6929>

Received: 2019-10-28 • Accepted: 2021-02-17 • Published (VoR): 2021-11-30

Handling Editors: Mojtaba Soltanlou, University of Surrey, Guildford, UK; Krzysztof Cipora, Loughborough University, Loughborough, UK

Corresponding Author: Anna F. Steiner, FH JOANNEUM, Studiengang Logopädie, Eggenberger Allee 13, 8020 Graz, Austria. Tel: +43 316 5453 6642. E-mail: anna.steiner@fh-joanneum.at

Related: This article is part of the JNC Special Issue “Direct and Conceptual Replication in Numerical Cognition”, Guest Editors: Mojtaba Soltanlou & Krzysztof Cipora, Journal of Numerical Cognition, 7(3), <https://doi.org/10.5964/jnc.v7i3>

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]



Abstract

Reading and writing multidigit numbers requires accurate switching between Arabic numbers and spoken number words. This is particularly challenging in languages with number-word inversion such as German (24 is pronounced as four-and-twenty), as reported by Zuber, Pixner, Moeller, and Nuerk (2009, <https://doi.org/10.1016/j.jecp.2008.04.003>). The current study aimed to replicate the qualitative error analysis by Zuber et al. and further extended their study: 1) A cross-linguistic (German, English) analysis enabled us to differentiate between language-dependent and more general transcoding challenges. 2) We investigated whether specific number structures influence accuracy rates. 3) To consider both transcoding directions (from Arabic numbers to number words and vice versa), we assessed performance for number reading in addition to number writing. 4) Our longitudinal design allowed us to investigate transcoding development between Grades 1 and 2. We assessed 170 German- and 264 English-speaking children. Children wrote and read the same set of 44 one-, two- and three-digit numbers, including the same number structures as Zuber et al. For German, we confirmed that a high amount of errors in number writing was inversion-related. For English, the percentage of inversion-related errors was very low. Accuracy rates were strongly related to number syntax. The impact of number structures was independent of transcoding direction or grade level and revealed cross-linguistic challenges of transcoding multidigit numbers. For instance, transcoding of three-digit numbers containing syntactic zeros (e.g., 109) was significantly more accurate than transcoding of items with lexical zeros (e.g., 190). Based on our findings, we suggest adaptations of current transcoding models.

Keywords

early transcoding, cross-linguistic, number word inversion, number writing, number reading, transcoding models

Number writing and reading are basic cultural skills acquired early in formal education. After their first experience with single-digit numerals, children develop their own ideas on how to read and write more complex numbers, for example, *five hundred and sixty* might be written as *50060*, or the Arabic number 203 might be read as *twenty, three* (Power & Dal Martello, 1990, 1997). In the present study, we specifically addressed language-independent and language-specific challenges in early number writing and reading.



Switching between the verbal code (number word) and the Arabic code (digits) is known as transcoding (Dehaene, 1992). The transcoding process is highly language-dependent and relies on number-word structures. The Arabic number 24 is pronounced as *twenty-four* in English, so the structure of Arabic numbers is consistent with the structure of number words. German, on the other hand, has an inconsistent correspondence between number formats, as units are pronounced before decades, for example, 24 is “*vierundzwanzig*”—literally translated as *four-and-twenty*. Up to now, theoretical models (e.g., Barrouillet, Camos, Perruchet, & Seron, 2004; Power & Dal Martello, 1990; Seron & Noël, 1995) are mostly language-specific (e.g., Moeller, Zuber, Olsen, Nuerk, & Willmes, 2015; Pixner et al., 2011; Steiner et al., 2021; Zuber et al., 2009) and model transcoding performance in languages with consistent number word structure, but are not always adequate for languages with deviant number word structures. The aim of this study was to establish a sound empirical basis to improve current transcoding models.

The Impact of Language on Transcoding

A decade ago, Zuber et al. (2009) were among the first to provide empirical evidence on language-specific transcoding problems based on a sample of 128 first graders, speaking a language with inversion: An impressive 50% of errors in writing Arabic numbers were due to incorrect digit order (categorized as inversion errors). In this widely recognized study, German-speaking children at the end of Grade 1 were asked to write 64 one-, two-, and three-digit numbers from dictation. A fine-grained error analysis for two- and three-digit numbers showed that problems with digit order were more marked in three- than in two-digit numbers, which the authors attributed to higher load on working memory. While this is a likely explanation for this finding, it should also be noted that these children may not have had much experience with three-digit numbers, which are only introduced later in the curriculum. Zuber et al.’s main conclusion was that the remarkable amount of problems with digit order was a result of the inversion principle in German. Their assumption is supported by earlier studies in languages without inversion, which did not report decade-unit order as an obstacle in number writing (Power & Dal Martello, 1990, 1997; Seron & Fayol, 1994). Subsequent cross-linguistic studies confirmed that children learning to read and write numbers in languages with decade-unit inversion have language-specific problems with digit order (Imbo, Vanden Bulcke, De Brauwere, & Fias, 2014; Moeller, Zuber, et al., 2015; Poncin, Van Rinsveld, & Schiltz, 2019; Steiner et al., 2021).

Inversion seems to have a particularly strong impact on transcoding performance in early development. Only small effects were reported on transcoding in older children and adults, both in languages with left-to-right writing systems (Steiner et al., 2021; van der Ven, Klaiber, & van der Maas, 2017) and right-to-left writing systems (Ganayim, Ganayim, Dowker, & Olkun, 2020; Hayek, Karni, & Eviatar, 2020). Decade-unit inversion was also reported to influence transcoding in a second (non-inverted) language. Native Arabic-speaking adults made (some) inversion errors in their second (non-inverted) language of Hebrew, but hardly any in their first (inverted) language (Ganayim et al., 2020).

Theoretical Models of Early Number Writing

Semantic and asemantic models were developed to explain early number writing. Verguts and Fias (2005) pointed out that both accounts are based on the prediction that rules need to be applied to successfully transcode spoken number words.

The distinctive feature of semantic transcoding models is the mandatory involvement of number semantics (McCloskey, 1992; Power & Dal Martello, 1990), which is strongly debated (Barrouillet et al., 2004; Power & Dal Martello, 1997; van Loosbroek, Dirx, Hulstijn, & Janssen, 2009; Verguts & Fias, 2005; Zuber et al., 2009). The central assumption is that spoken number words such as *three hundred and twenty-four* are translated into internal semantic representations of three hundreds (<C100 * C3>), two tens (<C10 * C2>), and four units (<C4>). These semantic representations are then again transformed to Arabic numbers based on concatenation and overwriting rules (Power & Dal Martello, 1990).

Asemantic models do not assume an intermediate semantic step but fully rely on direct transcoding procedures. The ADAPT-model (A Developmental Asemantic Procedural Transcoding model) by Barrouillet et al. (2004) was developed for the French number system. Spoken number words are parsed into lexical primitives where each element triggers application of corresponding production rules. For instance, for *three hundred and twenty-four* lexical units 3, 2, and 4 are

retrieved from long-term memory. The separator *hundred* defines number length. Procedures will fill empty slots with retrieved digits and terminate the transcoding process.

Zuber and colleagues (2009) pointed out that current theoretical models fail to account for the extra step involved in transcoding in languages with inversion such as in German. According to both semantic and asemantic transcoding models, transcoding procedures differ between number structures. The high relevance of number structure was demonstrated by a positive correlation between the percentage of transcoding errors and the number of procedures required according to the ADAPT-model (Barrouillet et al., 2004) in French second graders (Camos, 2008). Error rates for different number structures are expected to vary within a language due to structure-dependent transcoding procedures. Surprisingly, only a few studies took number structures into account (Lopes-Silva et al., 2016; Moura et al., 2013, 2015; Van Rinsveld & Schiltz, 2016). Most of the (cross-linguistic) studies explicitly addressing differences in number word systems did not differentiate number structures (Imbo et al., 2014; Moeller, Pixner, Zuber, Kaufmann, & Nuerk, 2011; Moeller, Zuber, et al., 2015; Zuber et al., 2009). Accuracy rates of number structures without inversion, such as X0 (e.g., 60), X0X (e.g., 206), or XX0 (e.g., 260) should not differ between languages with and without inversion.

Interestingly, Zuber et al. (2009) did report accuracy rates separately for specific number structures in their Table 1, but all statistical analyses were based on error rates across number structures. This may be relevant as their item set included inversion-demanding numbers such as XX- (e.g., 24, X represents a number between 1 and 9) or XXX-numbers (e.g., 624), and also single digits, double-digit numbers (X0: e.g., 40) and three-digit numbers (X00: e.g., 600, X0X: e.g., 206 or XX0: e.g., 260) that do not require inversion. Obviously, error types (and rates) in mixed item sets depend on the number structures included.

Challenges in Three-Digit Number Writing

Given the focus on inversion, Zuber et al. (2009) did not discuss unexpected outcomes in accuracy for certain number structures. Of particular interest is that error rates were higher for XX0- (e.g., 260) than for X0X-numbers (e.g., 206). Exactly the opposite pattern is predicted by the ADAPT-model: Barrouillet et al. (2004) suggested that two-digit numbers are lexicalized within the first two school years. Once a number is in the mental lexicon, it can be directly and quickly retrieved, which reduces the number of required procedures. For example, *two hundred and sixty* would be retrieved as lexical units 2 and 60 that would fill all empty slots. In contrast, *two hundred and six* would require retrieval of the lexical units 2 and 6 to fill the hundred and unit slots, plus an additional procedure to fill in the syntactic zero in the decade position. Note that the proposed early lexicalization of two-digit numbers would make language-specific rules superfluous since consistent as well as inverted numbers would be equally retrieved from long-term memory.

The pattern observed by Zuber et al. (2009) is more in line with the predictions of the semantic model by Power and Dal Martello (1990). This model proposes eight rules for XX0-numbers and only six for X0X-numbers. In detail, XX0-numbers require two concatenation rules (e.g., 2&00 → 200, 6&0 → 60) and one overwriting rule (200 # 60 → 260). X0X-numbers require only one concatenation rule (2&00 → 200) and one overwriting rule (200 # 6 → 206). If transcoding difficulty is indeed driven by individual number structures (as suggested by Camos [2008]) and not related to number size, mandatory semantic involvement is unlikely. A major aim of the current study was to further investigate Zuber et al.'s intriguing finding as it would imply that current transcoding models make incorrect predictions and need to be updated.

The Influence of Transcoding Direction on Transcoding Performance

While transcoding is a two-way process, theoretical models typically addressed either number writing (Barrouillet et al., 2004; Power & Dal Martello, 1990) or number reading (Dotan & Friedmann, 2018; McCloskey, 1992; Power & Dal Martello, 1997) or specified distinct sets of transcoding rules for each direction (Deloche & Seron, 1987). Obviously, opposite transcoding directions rely on distinct processes and therefore induce different error types (see Moura et al., 2013). However, so far, it has been overlooked that number structures may have a similar influence on both number writing and reading. We expect to see a similar influence of number structures on accuracy patterns in both transcoding directions. For instance, lack of place-value understanding should lead to low accuracy in number writing as well as

reading, even though errors might not be the same: *five hundred and sixty* may be written as 50060 whereas 560 might be read as *fifty-six, zero*.

The Current Study

Four specific research questions (Q) were addressed:

(Q1) Are digit-order errors related to decade-unit inversion?

The aim of the current study was to replicate and extend findings reported by Zuber et al. (2009), which had an important impact on later work on transcoding development. We thus decided to run a replication study based on a similar but larger sample of German-speaking children. Our cross-linguistic, longitudinal study design also allowed us to extend the seminal findings of Zuber et al. in a number of critical ways: Importantly, we investigated similarities and differences between a language with and a language without decade-unit inversion. German and English are ideal for such a comparison as they are closely related Germanic languages with highly similar lexical number primitives (e.g., *sechs*—*six*; *fünfzehn*—*fifteen*; *neunzig*—*ninety*; *hundert*—*hundred*), whereas only German applies decade-unit inversion (e.g., *vierundzwanzig*—*twenty-four*). If digit-order errors are related to decade-unit inversion, they should be frequent in German, as demonstrated by Zuber et al. but exceptional in English.

(Q2) Can accuracy rates be explained by current theoretical models?

An important extension of the Zuber et al. (2009) study is a detailed accuracy analysis separately for each number structure (e.g., X0-, and XX-numbers within two-digit and X00, X0X-, XX0-, and XXX-numbers within three-digit numbers). This fine-grained perspective will enable us to distinguish between language-dependent and more general challenges in transcoding. For German, we expected differences in accuracy rates between non-inverted numbers (e.g., X0-numbers as 60) and inverted numbers (e.g., XX-numbers as 67). No such differences should appear for non-inverted English. Thus, we predicted cross-linguistic differences in accuracy rates for number structures that are inverted but not for number structures that are not inverted in German.

We were particularly intrigued by Zuber et al.'s (2009) exploratory finding that X0X-numbers including a syntactic zero may be less error-prone than XX0-numbers including a lexical zero. A systematic replication of this finding would suggest that the procedures proposed by the ADAPT-model should be reconsidered. Table 1 presents a direct comparison of the number of transcoding procedures predicted by the two transcoding models.

Table 1

Predicted Number of Transcoding Procedures by the ADAPT-Model (Barrouillet et al., 2004) and the Model by Power and Dal Martello (1990)

Transcoding Model	Number of Transcoding Procedures					
	Two-Digit Numbers		Three-Digit Numbers			
	X0	XX	X00	X0X	XX0	XXX
ADAPT-model (Barrouillet et al., 2004)	2	2 (3) ^a	3	5	4	4 (5) ^a
Model by Power and Dal Martello (1990)	3	5	3	6	8	9

^aShown in parentheses are the numbers of rules if XX-numbers are algorithmically transcoded and not lexicalized.

(Q3) Are accuracy patterns similar for number writing and reading?

A further extension of Zuber et al. (2009) was that we asked children to write and read aloud the very same numbers. Previous studies comparing number writing and reading typically used different items for the two conditions (Habermann, Donlan, Göbel, & Hulme, 2020; Lopes-Silva et al., 2016; Moura et al., 2013). For some studies, authors did not provide a detailed description of items: Dowker & Roberts, 2015; Moeller, Zuber, et al., 2015). Our approach allowed us to investigate commonalities in the two transcoding directions for the presented number structures. Using the very same set of numbers ensured that any differences between number writing and reading could be attributed to transcoding strategies and not to item-specific characteristics. We assumed that children who had acquired certain procedures

for specific number structures in one transcoding direction should profit from this knowledge when transcoding in the opposite direction. Thus, we hypothesized that number structure would affect accuracy rates comparably in both transcoding directions.

(Q4) Are effects of number word structure consistent?

Most studies on early transcoding had only one assessment point at the end of Grade 1 (Moeller, Zuber, et al., 2015; Pixner et al., 2011; Zuber et al., 2009) or beginning of Grade 2 (Imbo et al., 2014). The longitudinal design of our study enabled us to investigate developmental trajectories of transcoding at two time points: towards the end of Grade 1 and again 1 year later, in Grade 2. Note that in Grade 1 two-digit numbers are explicitly taught, while three-digit numbers are formally introduced only in Grade 3. Thus, both assessments mostly investigated the conceptions of three-digit numbers that children developed on their own. This is important information as such independent conceptions may have a relevant impact on how more complex numbers should best be taught.

Method

Participants

Participants were recruited from five schools in Graz (Austria) and 11 schools across Yorkshire (United Kingdom) at the end of first grade. Overall, 170 native German-speaking children (47.1% female; T1: $M = 7;2$ years, $SD = 0;3$ years; T2: $M = 8;2$ years, $SD = 0;3$ years) and 264 native English-speaking children (47.7% female; T1: $M = 6;4$ years, $SD = 0;4$ years; T2: $M = 7;4$ years, $SD = 0;4$ years) took part in Grade 1 and 2. In Grade 2, 7 German- and 57 English-speaking children were not available due to relocation and school changes or temporary school absence. A power-analysis was conducted with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). Sample size was computed for an analysis of variance (ANOVA) with three within-subjects factors. We entered the effect size of .47 reported by Zuber et al. (2009) in the ANOVA on error categories. Power was set to .80, corresponding to conventions suggested by Cohen (1988). The probability of an alpha-error was fixed to .05. As results yielded a total sample size of 5 participants, both the original study and the present study easily meet sample size requirements.

Language groups were matched on duration of formal education, which resulted in an age difference of about 11 months between Austrian and English children as the UK school-system starts one year earlier. Two-digit numbers are explicitly taught in Grade 1 in both countries. The Austrian national curriculum (Bundesministerium für Bildung, Wissenschaft und Forschung, 2012) does not specify separate learning goals for Grades 1 and 2, but it requires that children should be able to count, read and write numbers up to 100 at the end of Grade 2. Most mathematical textbooks used in Year 1 include numbers up to 100. According to the UK national curriculum (Department for Education, 2013), children are expected to be able to count, read and write numbers to 100 by the end of Grade 1. According to both curricula, three-digit numbers are not explicitly taught before Grade 3.

In Grade 1, children completed selected items of the Numerical Operations subtest from the Wechsler Individual Achievement Test 2nd Edition (WIAT-II UK; Wechsler, 2005). Items were adapted for group use and assimilated to language-dependent notation of arithmetic operations. First, children had to master six items that involved identifying and writing Arabic digits to dictation and counting dots. Afterwards, they worked on nine standard arithmetic calculations (addition, subtraction, multiplication, and division with one- to three-digit numbers) with increasing difficulty for 15 minutes. German-speaking children ($M = 11.29$, $SD = 1.47$) performed significantly better than English-speaking children ($M = 9.86$, $SD = 2.15$), $t(459) = 8.18$, $p < .001$, $d = 0.78$.

We also individually administered language-specific word reading tests in which children had to read aloud a list of words as quickly as possible for 1 minute (in German: SLRT-II, Moll & Landerl, 2010) or 45 s (in English: TOWRE-2, Torgesen, Wagner, & Rashotte, 2012). In Grade 1, both groups showed average percentiles compared to test norms (German: between 81st and 83rd percentile, English: 77th percentile).

The current study was performed in accordance with the latest version of the Declaration of Helsinki and in compliance with national legislation. The University of Graz and the University of York Psychology Department Ethics

Committees approved the study and written informed consent was obtained from participants' legal guardians and school head teachers.

Materials and Procedure

Children were given a large battery of tasks to assess numerical cognition at three time points from Grade 1 to Grade 3. Here, we focus on number writing and reading, which were assessed in Grades 1 and 2.

Number Writing

Children were asked to write Arabic numbers to dictation.

Stimuli — The stimulus set contained one-, two-, and three-digit numbers based on the same number structures as in Zuber et al. (2009): X, teens, X0, XX, X00, X0X, XX0, XXX. In line with Zuber et al., we were interested in inversion-related errors. For that reason, we included more numbers that feature inversion (1X, XX, XXX) than numbers that do not feature inversion in German (X0, X00, X00, X0X, XX0). In first grade, the task comprised 52 items, including 4 one-digit numbers, 24 two-digit numbers, 16 three-digit numbers, and 8 four-digit numbers. In Grade 2, a slightly adapted task with 74 items was conducted. One-digit numbers were not re-administered due to ceiling effects observed in Grade 1. At the same time, the number of three- and four-digit items was increased. To ensure performance comparability, only those 40 items which were assessed at both grade levels were included in our analysis (for a complete list of the stimuli used see Table A.1 in the Appendix).

Procedure — The number-writing task was presented in four balanced blocks and was conducted as classroom activity where children had to fill in dictated items on separate lines. Each line was indicated by a picture to ensure that children were filling in the correct slot.

Analysis — Performance was first scored in terms of accuracy (correct = 1, incorrect = 0), and percentage correct for each number structure was computed. In order to replicate findings by Zuber et al. (2009), we conducted exactly the same error analysis for number writing in Grade 1: *lexical errors* were coded when incorrect lexical primitives were used (sixty-four → 65) or teens and decades were substituted with each other (seventy → 17). *Syntactic errors* were coded when either the principle of *additive composition* (three hundred and forty → 30040) or *multiplicative composition* (three hundred and forty → 3140) were ignored as well as when digit order was not correct. The third subcategory was *other syntactic errors* (sixty-four → 4). Incorrect digit order was further differentiated in *inversion errors* when transposed digits concerned decades and units or in *inversion wrongly applied* when the inversion rule was overgeneralized and applied to three-digit numbers (700 → 107). Errors that could not be explained as simple syntactic or simple lexical errors but were better explained by combination of two or three errors were coded as *combination errors* (three hundred and forty → 20040). We followed the procedures used by Zuber et al. (2009), who differentiated between six combination error types by combining lexical and syntactic errors or by combining two syntactic errors. To investigate the frequencies of all items which had *inversion-related errors*, we again followed the procedures described by Zuber et al. and pooled pure inversion, pure inversion wrongly applied, and combination errors including one of these two error categories. For further details, we refer to the original study by Zuber et al. To guarantee comparability, mean error rates were arcsine transformed following Zuber et al. Based on 26 German-speaking children, Cohen's kappa (κ) was computed to determine the interrater reliability for accuracy rating, which was .860. Two independent raters conducted the entire error analysis (interrater reliability for German: κ -values between .633 and 1.000, all $ps < .001$; interrater reliability for English: κ -values between .501 and .970, all $ps < .001$). To minimize ambivalences in error coding, deviations were discussed, and the adjusted data-set was used for statistical analysis.

Number Reading

Children were asked to read Arabic numbers aloud.

Stimuli — The item set was the same as for number-writing (see Table A.1 in the Appendix).

Procedure — In an individual testing session, children were presented with a list of numbers with increasing difficulty. Arabic numbers were presented as one item per line (font Calibri, size 20 point).

Analysis — Performance was scored in terms of accuracy (correct = 1, incorrect = 0), and percentage correct for each number structure was computed. Focusing on inversion, we performed an error analysis and coded incorrect digit order either as *inversion error*, when decades and units were transposed or as *inversion wrongly applied*, when hundred digits were named after units and decades. Inter-rater reliability was high for accuracy, $\kappa = .883$, $p < .001$, in the German-speaking sample as well as for inversion-related errors in German ($\kappa = .948$, $p < .001$) and English ($\kappa = .822$, $p < .001$).

Results

The data supporting the findings of this study were collected as part of a larger longitudinal project. The data-set of the present study is available on the Open Science Framework and can be accessed at <https://osf.io/nzevr>. Accuracy rates for individual number structures are shown in Table 2 separately for number writing and reading in Grade 1 and 2. Table 3 presents inversion-related error rates for individual number structures in both transcoding directions.

Table 2

Means and SDs of Accuracy Rates (ratio: 1.00 = 100% Correct) by Number Structure for Number Writing and Number Reading at Grade 1 and 2

Number Structure	(Example)	Grade 1								Grade 2							
		Number Writing				Number Reading				Number Writing				Number Reading			
		German		English		German		English		German		English		German		English	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
One- and Two-Digit Numbers																	
X	(4)	1.00	0.00	1.00	0.03	1.00	0.00	1.00	0.02	–	–	–	–	–	–	–	–
Teen	(13)	0.96	0.14	0.90	0.20	0.92	0.19	0.90	0.21	1.00	0.03	0.97	0.12	0.99	0.06	0.98	0.10
X0	(70)	0.86	0.27	0.83	0.28	0.89	0.23	0.93	0.20	0.99	0.04	0.96	0.12	0.97	0.12	0.99	0.08
XX	(56)	0.82	0.28	0.95	0.13	0.85	0.26	0.94	0.16	0.99	0.06	0.99	0.04	0.99	0.05	0.99	0.06
Three-Digit Numbers																	
X00	(700)	0.78	0.33	0.81	0.31	0.84	0.30	0.79	0.35	0.92	0.20	0.97	0.11	0.95	0.17	0.97	0.16
X0X	(304)	0.45	0.42	0.44	0.42	0.62	0.39	0.58	0.41	0.78	0.35	0.79	0.34	0.90	0.23	0.88	0.27
XX0	(340)	0.32	0.42	0.25	0.40	0.47	0.46	0.53	0.45	0.72	0.41	0.64	0.43	0.87	0.31	0.88	0.28
XXX	(349)	0.28	0.41	0.25	0.40	0.48	0.47	0.57	0.45	0.69	0.42	0.64	0.44	0.90	0.26	0.91	0.24

Note. Mean parameter values are shown for German- (number writing: $N = 170$ besides number reading in T1: $N = 169$) and English-speaking children (number writing and reading: $N = 264$).

Error Analysis of Number Transcoding in First Graders

Inversion-Related Errors

In order to replicate Zuber et al. (2009), we first examined whether the inversion property is a challenge in early transcoding. Zuber et al. reported that about 50% of all errors were inversion-related, and this was not significantly different from the percentage of non-inversion errors. The comparison of frequencies of inversion-related versus non-inversion-related errors in our sample of German-speaking first graders revealed that about 40% of all number writing errors were inversion-related ($M = 40.23\%$, $SD = 37.99\%$; non-inversion-related errors: $M = 60.48\%$, $SD = 48.62\%$). This confirms Zuber et al.'s finding that inversion is a major challenge at this age, but in our study, non-inversion-related errors were overall more frequent, $t(169) = -3.54$, $p = .001$, $d = 0.27$.

We extended these findings in two ways: First, to confirm that inversion problems were language-specific, we compared German- and English-speaking first graders. Secondly, we compared inversion errors in both transcoding directions.¹ Figure 1 shows significantly higher inversion rates (based on incorrect items) for German- than for

Table 3

Children's Transcoding Inversion Error Rates (Ratio: 1.00 = 100% Incorrect) as a Function of Types of Numbers

Number Structure	(Example)	Number Writing				Number Reading			
		German		English		German		English	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Two-Digit Numbers									
Teen	(13)	0.01	0.06	0.04	0.11	0.03	0.11	0.02	0.08
X0	(70)	0.08	0.20	0.05	0.14	0.01	0.10	0.00	0.00
XX	(56)	0.16	0.26	0.01	0.04	0.11	0.21	0.02	0.05
Three-Digit Numbers									
X00	(700)	0.14	0.28	0.02	0.09	0.11	0.24	0.04	0.15
X0X	(304)	0.09	0.23	0.02	0.09	0.21	0.31	0.09	0.20
XX0	(340)	0.16	0.27	0.11	0.19	0.20	0.29	0.03	0.12
XXX	(349)	0.37	0.39	0.01	0.05	0.18	0.29	0.01	0.04

Note. Mean parameter values are shown for German- (number writing: $N = 170$, number reading: $N = 169$) and English-speaking children (number writing and reading: $N = 264$).

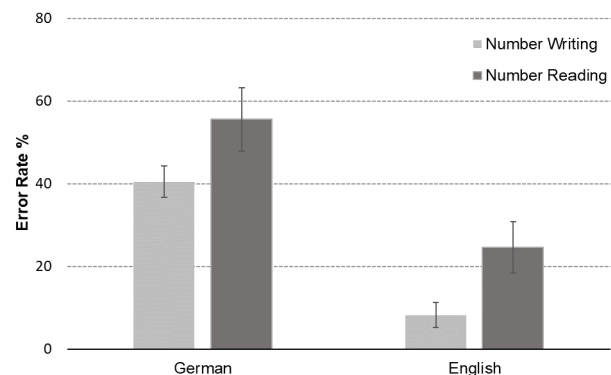
English-speaking children, $F(1, 431) = 120.630$, $p < .001$, $\eta_p^2 = .22$. In both languages, inversion errors were even more prominent in number reading than in number writing, $F(1, 431) = 34.85$, $p < .001$, $\eta_p^2 = .08$. There was no significant interaction between transcoding direction and language, $F(1, 431) = 0.06$, $p = .808$.

Zuber et al. (2009) reported significantly more inversion-related errors for number writing of three- ($M = 25.20\%$, $SD = 7.32$) than two-digit numbers ($M = 24.06\%$, $SD = 4.19$), $t(127) = 3.18$, $p < .01$, $d = 0.23$. We replicated this finding for the German-speaking sample and found a marked length effect (error rate based on all presented items for two-digit numbers: $M = 12.62\%$, $SD = 20.69\%$, for three-digit numbers: $M = 20.67\%$, $SD = 26.49\%$), $t(169) = -3.45$, $p = .001$, $d = 0.26$. In the current sample, variation for inversion-related errors was quite high and clearly higher than reported for the original study by Zuber et al. Most children did not experience problems with inversion as 82 participants made no inversion-related errors in two-digit numbers and 53 children made no inversion-related errors in three-digit numbers. However, some children still struggled with digit order and made inversion-related error rates in more than 50% of the items (10 children in two-digit numbers and 12 children in three-digit numbers).

A length effect was also observed for number reading in German (error rate based on all presented items for two-digit numbers: $M = 8.28\%$, $SD = 15.46\%$, for three-digit numbers: $M = 18.13\%$, $SD = 22.01\%$), $t(168) = -5.05$, $p < .001$, $d = 0.40$. English-speaking children did not struggle with digit order and showed a floor effect for inversion-related errors in two digit-numbers (number writing: $M = 2.00\%$, $SD = 4.43\%$, number reading: $M = 1.36\%$, $SD = 3.96\%$), and an only slightly higher error rate for three-digit numbers (number writing: $M = 3.95\%$, $SD = 6.42\%$, number reading: $M =$

Figure 1

Mean Inversion-Related Error Rates Based on All Errors Separately for Number Writing and Reading in German- and English-Speaking Children



Note. Error bars depict 95% CI.

1) Note that one German-speaking child did not complete number reading in Grade 1 and was therefore excluded from analysis. The Language x Transcoding direction ANOVA is based on 169 German-speaking participants.

4.10%, $SD = 7.77\%$). Nevertheless the length effect was also significant for English-speaking children for both, number writing, $t(263) = -4.27$, $p < .001$, $d = 0.26$, and reading, $t(263) = -5.02$, $p < .001$, $d = 0.31$.

Fine-Grained Analysis of Error Types

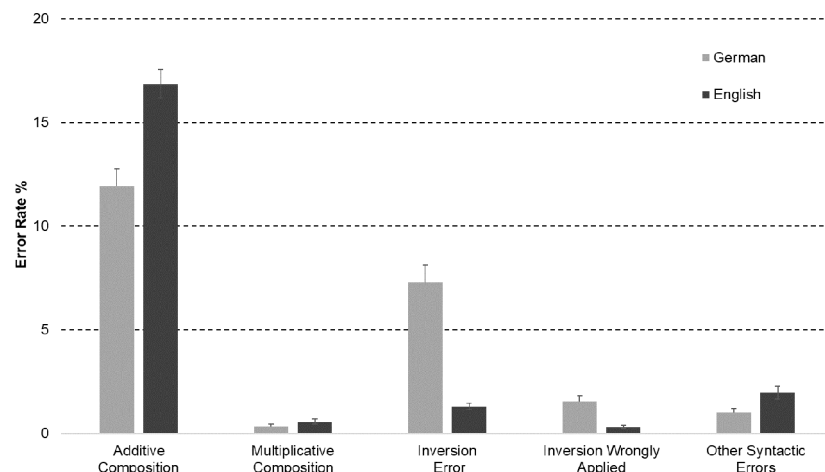
Following procedures by Zuber et al. (2009), we looked more closely at different error categories in number writing. In a repeated measures ANOVA² for German-speaking first graders, we replicated the findings of the original study: Lexical errors rarely occurred ($M = 2.02\%$, $SD = 3.15\%$). The predominant category were syntactic errors ($M = 23.92\%$, $SD = 20.00\%$) followed by combination errors ($M = 7.76\%$, $SD = 8.96\%$), with significant differences between all three error types, $F(1, 215) = 179.45$, $p < .001$, $\eta_p^2 = .52$, all follow-up t -tests $ps < .001$. When we ran the very same analysis for English-speaking first graders, the main effect of error type was again significant, $F(1, 352) = 455.08$, $p < .001$, $\eta_p^2 = .63$. However, it turned out that rates of combination errors ($M = 3.62\%$, $SD = 6.17\%$) were low and did not significantly differ from lexical error rates ($M = 3.37\%$, $SD = 5.36\%$, $p = .499$). Syntactic errors ($M = 21.22\%$, $SD = 13.32\%$) were again clearly more prevalent than the two other error categories (both $ps < .001$).

Next, following Zuber et al. (2009), we split up the prominent syntactic error category into additive composition errors, multiplicative errors, inversion errors, inversion wrongly applied errors, and other syntactic errors. For German-speaking children, we replicated Zuber et al.'s distribution (see Figure 1B in Zuber et al., 2009) and confirmed significant differences between frequencies, $F(2, 386) = 91.82$, $p < .001$, $\eta_p^2 = .35$. As displayed in Figure 2, inversion errors represented a large portion of syntactic errors in our sample of German-speaking first graders. Pairwise comparisons (Bonferroni corrected α -level is .010) replicated the original finding that additive composition errors were the most frequent error type (all $ps < .001$), followed by inversion errors (all $ps < .001$). For English-speaking first graders, we also found differences in frequency of syntactic error types, $F(1, 381) = 421.67$, $p < .001$, $\eta_p^2 = .62$. However, as evident from Figure 2, inversion errors hardly occurred in English, which is in stark contrast to German. A similarity to the German pattern was that additive composition errors were the most frequent syntactic error type (all $ps < .001$) in English as well. Other error types were rare.

In the study by Zuber et al. (2009), most combination errors were co-occurrences of inversion errors and other syntactic error types (see Figure 1C in Zuber et al.). As displayed in Figure 3, we replicated this finding in our sample of German-speaking first graders. A repeated measures ANOVA showed differences in the frequency of error types, $F(2, 280) = 69.49$, $p < .001$, $\eta_p^2 = .29$, and pairwise comparisons (Bonferroni corrected α -level is .008) confirmed the combination of inversion errors and any other syntactic error type as the most frequent type (all $ps < .001$). As error rates for combination errors in English-speaking first graders were very low, we refrained from further analyses for the English-speaking sample.

Figure 2

Mean Syntactic Error Rates in Each Error Category for German- and English-Speaking First Graders

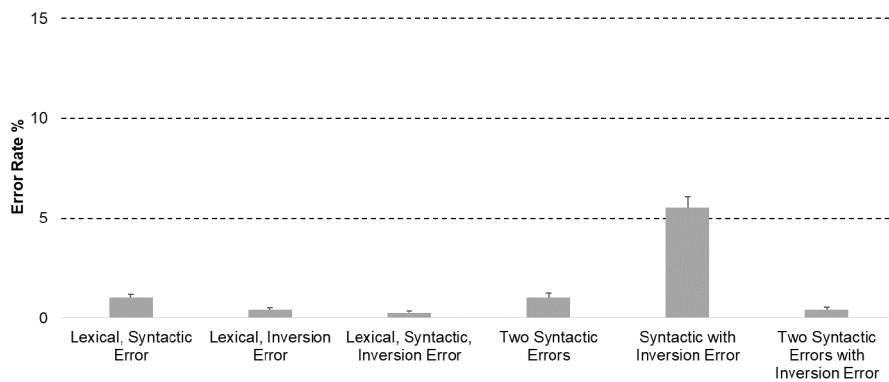


Note. Error bars depict 1 SEM.

2) In case Mauchly's test indicated a violation of the sphericity assumption, results for all conducted ANOVAs were Greenhouse-Geisser corrected.

Figure 3

Mean Combination Error Rates in Each Error Category for German-Speaking Children



Note. In line with Zuber et al. (2009), inversion errors were separated from other syntactic errors. In this figure, “syntactic” refers to “syntactic except inversion” and highlights the role of inversion. In English-speaking children, combination errors were not further analyzed due to low frequency (<4%). Error bars depict 1 SEM.

Summary

We replicated previous findings by Zuber et al. (2009) for early number writing. German-speaking first graders struggled with inconsistent digit order in number words. The cross-linguistic comparison with English-speaking children elucidated that digit order problems were language-specific and clearly driven by decade-unit inversion in German number words. This was evident in number writing and reading. The major source of errors in number writing, however, was the additive composition principle. In both languages, children struggled with overwriting zeros. In German, inversion errors and additive composition errors were large subcategories of syntactic errors and also co-occurred often. In English, additive composition errors were clearly the largest error category, while inversion errors were rare, which also explains why combination errors were less frequent than in German. Zuber et al. neglected additive composition errors and focused on inversion with the aim to constrain theoretical transcoding models with empirical data. The fact that we replicated the high prevalence of additive composition errors, however, highlights the importance of investigating this error type in more detail: Problems with overwriting unit digits (e.g., two hundred six → 206) and/or decade digits (e.g., two hundred and sixty → 260) were subsumed under additive composition errors but might have distinct implications for current transcoding models. In the following, we analyze individual number structures to disentangle distinct sources of additive composition errors.

Number Structures: Accuracy Analysis for Number Writing and Reading

We analyzed the impact of number structures on transcoding performance based on accuracy rates. This allowed us to perform cross-linguistic analyses and to compare both transcoding directions longitudinally (Grades 1 and 2). We tested whether accuracy rates can be explained by the number of required procedures as proposed in the ADAPT-model. The analysis of accuracy rates for German- and English-speaking samples for number writing and reading is based on the mean scores displayed in Table 2. For one-digit numbers and teens, accuracy was very high across language groups and assessment points. Thus, we focused our analyses on multidigit numbers. Two-digit numbers are taught in Grade 1 (and accuracy was at ceiling in Grade 2), whereas according to the curricula in Austria and the UK, there is no focus on three-digit numbers before Grade 3. Due to those differences in educational experience, analyses were run separately for two- and three-digit numbers.

Two-Digit Numbers

To test whether the German inversion rule causes a language-specific pattern that is different from non-inverted numbers in English, we compared performance on X0- and XX-numbers in the two language groups. Note that the

inversion rule only affects XX-, but not X0-numbers. We focused on the additional load of inversion and excluded teen numbers. Teens are particulars, as they include decade-unit inversion in German and English (*thirteen–dreizehn*). For a detailed cross-linguistic analysis on teens, see Clayton et al. (2020). In Grade 2, performance on two-digit numbers was at ceiling, therefore we only analyzed Grade 1 data. We conducted a $2 \times 2 \times 2$ ANOVA with number structure (X0, XX) and transcoding direction (number writing, number reading) as within-subjects factors and language as between-subjects factor. Results are presented in Table 4. Crucially, the three-way interaction Number structure \times Language \times Transcoding direction was significant.

Table 4

Results of a 2 (Transcoding Direction) \times 2 (Number Structure) \times 2 (Language) ANOVA for Two-Digit Numbers

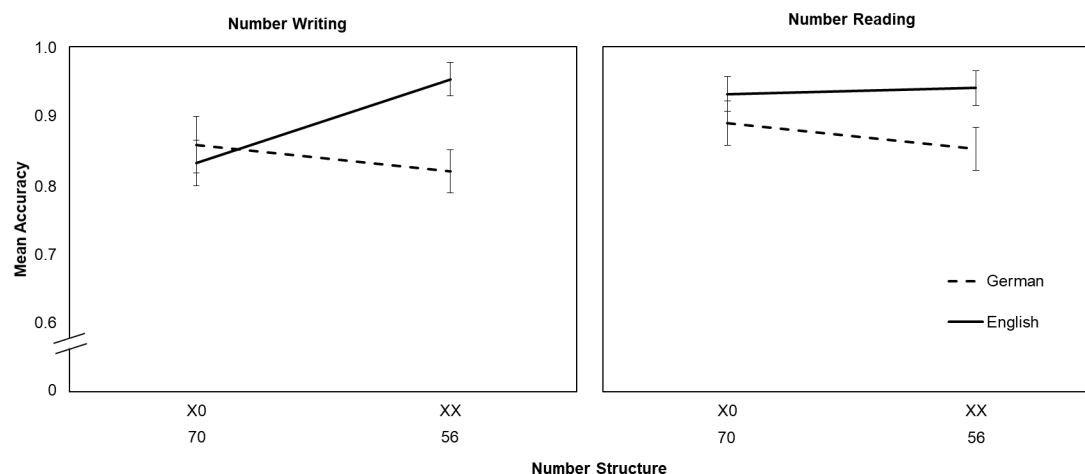
Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η_p^2
Between subjects						
Language	1	1.45	1.45	12.96	< .001	.03
Error (Language)	431	48.06	0.11			
Within subjects						
Transcoding Direction	1	0.58	0.58	20.05	< .001	.04
Transcoding Direction \times Language	1	0.01	0.01	0.46	.500	.00
Error (Transcoding Direction)	431	12.42	0.03			
Number Structure	1	0.08	0.08	2.10	.148	.01
Number Structure \times Language	1	1.08	1.08	28.93	< .001	.06
Error (Number Structure)	431	16.04	0.04			
Number Structure \times Transcoding Direction	1	0.31	0.31	13.32	< .001	.03
Number Structure \times Language \times Transcoding Direction	1	0.33	0.33	14.05	< .001	.03
Error (Number Structure \times Transcoding Direction)	431	9.96	0.02			

Note. *SS* = sum of squares; *MS* = mean square. *N* = 433 (German-speaking second graders: *N* = 169, English-speaking second graders: *N* = 264). Greenhouse-Geisser correction was applied and degrees of freedom were commercially rounded. Significant results are printed in bold.

Accuracy was generally high for two-digit numbers in both language groups. Nevertheless, number writing performance on XX-numbers was higher among English- than German-speaking children (see Figure 4).

Figure 4

Mean Accuracy Rates (Ratio: 1.00 = 100% Correct) in Two-Digit Numbers Showing Number Structure \times Language Interaction for Different Transcoding Directions in Grade 1



Note. On the left: number writing, on the right: number reading. Both number structures X0 and XX are displayed for German-speaking (dashed line) and English-speaking (solid line) children. The presented two-digit numbers below the number structure labels exemplify respective items. Error bars represent 95% CI.

Number writing accuracy for X0-numbers was comparable across languages. For number reading, accuracy was lower in German- than in English-speaking children for both number structures with a more pronounced language difference in XX- than X0-numbers.

Inspection of accuracy rates within languages revealed similar accuracy rates for both number structures in number writing as well as reading among German-speaking children. Among English-speaking children, accuracy rates for both number structures were again comparable for number reading. Number writing performance on XX-numbers, however, exceeded performance on X0-numbers. Note that based on the ADAPT-model (Barrouillet et al., 2004), we would have expected similar or even lower accuracy rates for XX- than X0-numbers.

Three-Digit Numbers

Due to the overall higher difficulty of three- compared to two-digit numbers, no ceiling effects were evident in Grade 2. This allowed us to look at longitudinal effects in transcoding development. Two aspects were of particular interest: We investigated the language-specificity of transcoding procedures by focusing on the cross-linguistic comparison of XXX-numbers versus numbers with language-independent number structures (e.g., X00-, X0X- or XX0-numbers) as only XXX-numbers require decade-unit inversion in German. Within the three number structures that are similar in German and English (X00, X0X- and XX0), we were particularly interested whether we could replicate Zuber et al.'s unexpected observation that numbers with a syntactic zero (X0X) yielded lower error rates than numbers with a lexical zero (XX0).

A 4 (Number structure) \times 2 (Grade level) \times 2 (Transcoding direction) \times 2 (Language) ANOVA revealed two significant three-way interactions (Number structure \times Grade level \times Language and Number structure \times Transcoding direction \times Language). In Table 5, all outcomes of the ANOVA are presented.

Table 5

Results of a 2 (Grade Level) \times 2 (Transcoding Direction) \times 2 (Number Structure) \times 2 (Language) ANOVA for Three-Digit-Numbers

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η_p^2
Between Subjects						
Language	1	0.00	0.00	0.02	.887	< .01
Error (Language)	431	25.82	0.06			
Within subjects						
Grade Level	1	157.73	157.73	643.20	< .001	.60
Grade Level \times Language	1	0.01	0.01	0.02	.876	.00
Error (Grade Level)	431	105.69	0.25			
Transcoding Direction	1	37.75	37.75	297.75	< .001	.41
Transcoding Direction \times Language	1	0.30	0.30	2.39	.123	.01
Error (Transcoding Direction)	431	54.65	0.13			
Number Structure	2	92.99	49.66	451.90	< .001	.51
Number Structure \times Language	2	0.37	0.20	1.82	.165	< .01
Error (Number Structure)	807	88.69	0.11			
Grade Level \times Transcoding Direction	1	0.31	0.31	2.67	.103	.01
Grade Level \times Transcoding Direction \times Language	1	0.03	0.03	0.27	.607	< .01
Error (Grade Level \times Transcoding Direction)	431	5.45	0.12			
Number Structure \times Grade Level	2	16.24	7.83	106.57	< .001	.20
Number Structure \times Grade Level \times Language	2	0.55	0.26	3.59	.026	.01
Error (Number Structure \times Grade Level)	894	65.68	0.07			
Number Structure \times Transcoding Direction	2	13.10	5.32	110.05	< .001	.20
Number Structure \times Transcoding Direction \times Language	2	2.10	0.85	17.66	< .001	.04
Error (Number Structure \times Transcoding Direction)	1,060	51.29	0.05			
Number Structure \times Grade Level \times Transcoding Direction	2	0.13	0.06	1.04	.361	< .01
Number Structure \times Grade Level \times Transcoding Direction \times Language	2	0.12	0.05	0.99	.380	< .01
Error (Number Structure \times Grade Level \times Transcoding Direction)	999	53.36	0.05			

Note. *SS* = sum of squares; *MS* = mean square. *N* = 433 (German-speaking second graders: *N* = 169, English-speaking second graders: *N* = 264). Greenhouse-Geisser correction was applied and degrees of freedom were commercially rounded. Significant results are printed in bold letters.

We expected differences in accuracy rates between adjacent number structures as presented by Zuber et al. (2009) (Table 1). Thus, we ran repeated planned contrasts³ comparing X00- with X0X-numbers, X0X- with XX0-numbers and XX0- with XXX-numbers. Results revealed higher accuracy for X00- than X0X-numbers, $F(1, 431) = 360.11, p < .001, \eta_p^2 = .46$ which in turn had higher accuracy than XX0-numbers, $F(1, 431) = 174.44, p < .001, \eta_p^2 = .29$. Thus, we replicated the accuracy pattern reported by Zuber et al. (2009), which is in contrast to predictions of the ADAPT-model. Number structures XX0 and XXX did not differ significantly from each other, $F(1, 431) = 3.63, p = .057, \eta_p^2 = .01$. As expected, results revealed significantly better transcoding performance in Grade 2 than in Grade 1, $F(1, 431) = 643.20, p < .001, \eta_p^2 = .60$. Accuracy in number reading was significantly higher than in number writing, $F(1, 431) = 297.75, p < .001, \eta_p^2 = .41$. No main effect of language was found, $F(1, 431) = 0.02, p = .887$.

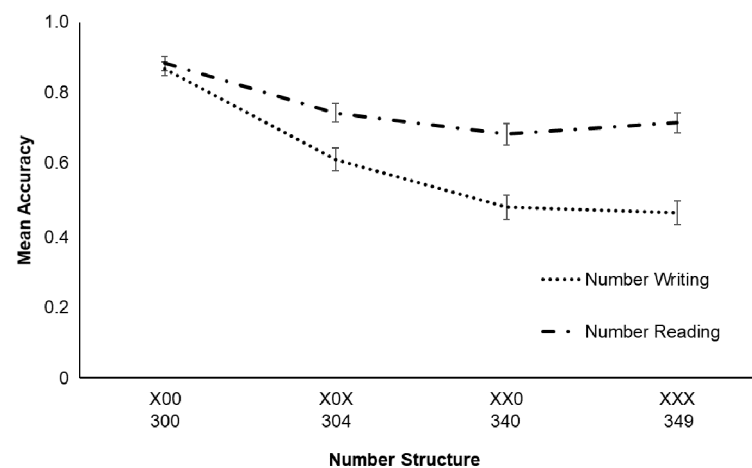
Influence of Grade Level — There was a significant interaction between number structure (X00, X0X, XX0, XXX) and grade level (Grade 1, Grade 2). Planned contrasts showed that the decrease in accuracy from X00-numbers to X0X-numbers was smaller in Grade 2 than in Grade 1, $F(1, 431) = 108.19, p < .001, \eta_p^2 = .20$. This was also observed for X0X- compared to XX0-numbers, $F(1, 431) = 21.16, p < .001, \eta_p^2 = .05$. These findings indicated overall better performance in transcoding three-digit numbers in Grade 2 compared to Grade 1. The difference in accuracy between XX0-numbers and XXX-numbers was not affected by grade level, $F(1, 431) = 1.03, p = .311$.

Influence of Transcoding Direction — Figure 5 presents average accuracy rates for three-digit numbers separately for number writing and reading. There was a significant interaction between number structure and transcoding direction. While performance in X00-numbers was close to ceiling in number writing as well as reading, the decrease in accuracy between X00 and X0X was larger in number writing than reading, $F(1, 431) = 64.51, p < .001, \eta_p^2 = .13$. A similar difference between number writing and reading was observed for the contrast between X0X and XX0, $F(1, 431) = 25.76, p < .001, \eta_p^2 = .06$. Number writing accuracy did not differ between XX0- and XXX-numbers whereas number reading accuracy was higher for XXX- than XX0-numbers; $F(1, 431) = 27.743, p < .001, \eta_p^2 = .06$.

Figure 4. Mean accuracy rates (ratio: 1.00 = 100% correct) in two-digit numbers showing Number structure \times Language interaction for different transcoding directions (on the left: number writing, on the right: number reading) in Grade 1. Both number structures X0 and XX are displayed for German-speaking (dashed line) and English-speaking (solid line) children. The presented two-digit numbers below the number structure labels exemplify respective items. Error bars represent 95% CI.

Figure 5

Mean Accuracy Rates (Ratio: 1.00 = 100% Correct) Showing Transcoding Direction \times Number Structure Interaction not Differentiating Between Grade Level and Language



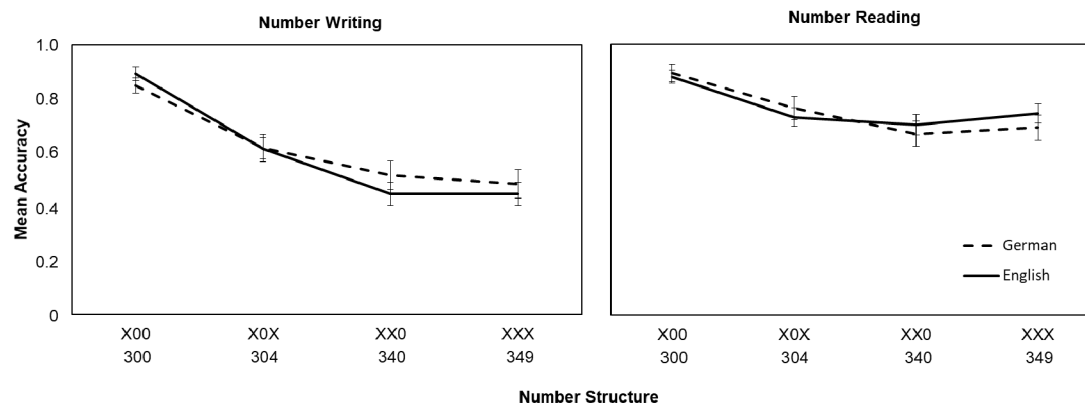
Note. Results for four number structures (X00, X0X, XX0, XXX) are displayed for number writing (dotted line) and number reading (dashed line). The presented three-digit numbers below the number structure labels exemplify respective items. Error bars represent 95% CI.

Influence of Language — As reported above, the interactions Number structure \times Grade level and Number structure \times Transcoding direction were modulated by language. The effect size of the Number structure (X00, X0X, XX0, XXX) \times Grade level (Grade 1, Grade 2) \times Language (English, German)

3) Planned contrasts define in advance a set of comparisons between variable levels which are based on hypothesis (Field, 2019).

Figure 6

Mean Accuracy Rates (Ratio: 1.00 = 100% Correct) in Three-Digit Numbers Showing the Number Structure \times Language Interaction for Different Transcoding Directions Including Both Grades



Note. On the left: number writing, on the right: number reading. Number structures (X00, X0X, XX0, and XXX) are displayed for German-speaking (dashed line) and English-speaking (solid line) children. The presented three-digit numbers below the number structure labels exemplify respective items. Error bars represent 95% CI.

interaction was small but significant. However, the conducted contrasts did not reveal any language-dependent accuracy difference, $0.74 < F_s(1, 431) < 2.49$, all $p_s > .1$.

The interaction Number structure \times Transcoding direction \times Language, is depicted in Figure 6. Contrasts revealed that the difference in accuracy between X0X- and XX0-numbers in number writing compared to reading was less marked in German- than English-speaking children, $F(1, 431) = 21.38$, $p < .001$, $\eta_p^2 = .05$. No other contrast was significant, X00 to X0X: $F(1, 431) = 0.73$, $p = .393$, XX0 to XXX: $F(1, 431) = 0.64$, $p = .425$. Note that we did not see the expected language difference for XXX-numbers which required decade-unit inversion in German, but not in English.

Summary

Accuracy for two-digit numbers was already at ceiling in Grade 2, whereas three-digit numbers displayed higher intersubject variation and showed a stable accuracy pattern over time and across transcoding directions. The analysis of accuracy rates for individual number structures revealed language-related differences in two-digit numbers: XX-numbers were transcribed less accurately in German (with inversion) than in English (without inversion). On the contrary, negligible language-related differences were found for X0-numbers, without inversion across languages. Interestingly, English-speaking children achieved lower accuracy rates in writing X0- than XX-numbers. If two-digit numbers are lexicalized early, no such difference should appear. In three-digit numbers, accuracy patterns were comparable across languages. Contrary to the predictions by the ADAPT-model, XX0-numbers yielded lower accuracy than X0X-numbers, although more procedures are required to transcode X0X- than XX0-numbers.

Discussion

We investigated language-specific and language-independent challenges in transcoding by comparing languages with and without decade-unit inversion (German vs. English). Importantly, we fully replicated findings by Zuber et al. (2009), showing that in German decade-unit inversion has a strong impact on early number writing. Language-specific characteristics, however, are largely lacking from current transcoding models. Our study specifically benefits from the comparison of the two Germanic languages, which have highly similar number word structures but differ in the critical issue of decade-unit inversion. The cross-linguistic comparison clearly confirmed that problems with digit order are related to decade-unit inversion.

We intended to establish a sound empirical basis to improve current transcoding models and specifically accounted for individual number structures as modulators of transcoding performance (Camos, 2008; Moura et al., 2013). Indeed, cross-linguistic accuracy comparisons revealed language-specific and language-independent obstacles in early transcoding. Accuracy rates allowed us to compare both transcoding directions (number writing and reading) and highlighted the strong impact of number structures on transcoding in both directions. In our longitudinal study, we tracked transcoding development from Grade 1 to Grade 2 and showed enduring challenges in transcoding.

Language-Specific Challenges in Transcoding

Our error analyses of number writing in Grade 1 confirmed results by Zuber et al. (2009) that decade-unit inversion in German is a major obstacle in early transcoding: About 40% of incorrect responses were inversion-related. We also replicated the finding that three-digit numbers were more (inversion) error-prone than two-digit numbers in German. Zuber et al. attributed the difference in accuracy between two- and three-digit numbers to working memory effects as proposed in the ADAPT-model (Barrouillet et al., 2004). Indeed, we found that for both a language with and a language without inversion, increased number length led to more digit-order problems. However, it should be noted that first graders have not yet received formal instruction on how to write three-digit numbers. The fact that in German three-digit numbers are only taught much later than two-digit numbers may have an adverse influence on conceptions of inversion: We observed that German-speaking children frequently overgeneralized the decade-unit inversion rule and wrote the first number of the number word (the hundred) at the end of their number writing attempt (e.g., 700 → 107). Even though this error type was occasionally observed in English, too, it was far more frequent in German. Thus, the existence of this remarkable type of inversion-related error strongly suggests that the higher error rate for longer numbers may result from conceptual rather than working memory problems, at least in German.

Further evidence for the assumption that inversion errors result from conceptual rather than memory problems comes from the fact that they occurred in both transcoding directions. Indeed, in German-speaking children, number reading revealed an even bigger proportion of inversion-related errors than writing (even though the absolute error rate was lower). Note that in number reading, digit order is permanently visible to the child, so working memory demands are clearly lower during number reading than during number writing. Nevertheless, digit order was still a fundamental challenge for German-speaking children also during number reading, at least in the early stages of development. Large standard deviations in inversion-related errors indicate that children do not solve the task in the same way. Differences in (prior) knowledge on number writing and reading lead to a marked variance in task performance.

The special challenges of inversion were also reflected in our analysis of accuracy rates for individual number structures. In both number writing and reading, transcoding (inverted) XX-numbers in German was more difficult than (non-inverted) XX-numbers in English. Thus, we found strong evidence that decade-unit inversion specifically affected inversion-related number structures. In three-digit numbers, however, XXX-numbers did not reveal language-specific accuracy rates. Untrained three-digit numbers increased error-proneness in German and English. Language-independent challenges and a lack of experience may have masked the impact of decade-unit inversion. In German, inversion errors often co-occurred with other syntactic errors, as shown by the high frequency of combination errors.

In summary, our cross-linguistic comparison provides support for Zuber et al.'s conclusions (2009) that digit-order problems are inversion-related. Problems with decade-unit inversion were clearly less prominent in English-speaking children. Findings confirmed earlier cross-linguistic studies on somewhat more distant languages like Dutch and French (Imbo et al., 2014) or German and Japanese (Moeller, Zuber, et al., 2015). Moreover, they are in line with recent findings showing long-lasting inversion effects on transcoding: German-speaking fourth graders needed more time to select Arabic two-digit numbers matching a spoken number word than French-speaking fourth graders who do not have to deal with inversion (Poncin et al., 2019). An enduring impact of inversion was recently also demonstrated in a verbal-visual matching task where German-speaking elementary school children and even adults with ample experience in number writing and reading needed longer to reject inverted Arabic numbers than English-speaking samples (Steiner et al., 2021). Decade-unit inversion impacts early number development and leads to permanent adverse effects on transcoding. We also replicated Zuber et al.'s (2009) finding that lexical errors rarely occurred in German, and this was also true for English. In both languages, corresponding units and decades (such as six and sixty) are phonologically similar, and this

could facilitate lexical retrieval. Findings seem to be different in French with its vigesimal number word system (Imbo et al., 2014; Van Rinsveld & Schiltz, 2016). French number words above 60 involve inconsistencies in decade numbers (e.g., 80 is literally pronounced as four-twenty) which was shown to have enduring effects on two-digit number transcoding. For example, French-speaking 10-year-olds showed longer reaction times in a verbal-visual matching task than their English-speaking peers for numbers above 60 but not for lower numbers (Van Rinsveld & Schiltz, 2016).

There is emerging evidence that the linguistic impact on number processing may even go beyond lexical features such as number word peculiarities. For example, orthographic characteristics (e.g., reading direction) were shown to influence number processing (Moeller, Shaki, Göbel, & Nuerk, 2015) and the influence of further linguistic factors such as the markedness of words (conceptual linguistic feature) are currently discussed (Bahnmüller, Cipora, Goebel, Nuerk, & Soltanlou, this issue).

Language-Independent Challenges in Transcoding

Across languages, the biggest challenge of transcoding development appears to be the understanding of additive composition rules. The frequent occurrence of additive composition errors in both languages, characterized by supernumerary zeros, corroborates previous research (Imbo et al., 2014; Moeller, Zuber, et al., 2015; Moura et al., 2013). Note that this error type encompasses problems with overwriting units (304 → 3004) and/or decades (e.g., 340 → 30040). Thus, overwriting errors are typically combined across different complexity levels depending on number structure. Such combined scores do not reflect that children may acquire decade-unit additive composition for two-digit numbers (e.g., twenty-four → 24) but may still struggle with hundred-decade composition (e.g., three hundred and twenty-four → 30024). The current analysis of individual number structures revealed interesting (language-independent) challenges in transcoding: In both languages, XX0-numbers were more difficult to transcode than X0X-numbers. Current transcoding models assume that number structures activate specific procedures (Barrouillet et al., 2004; Dotan & Friedmann, 2018; Power & Dal Martello, 1990). Importantly, the accuracy patterns observed in our analysis were not in line with the ADAPT-model (Barrouillet et al., 2004), which proposes more procedures and therefore lower accuracy rates for syntactic zeros as in X0X- than for lexical zeros as in XX0-structures. Interestingly, the observed pattern fits with the assumptions of the semantic model by Power and Dal Martello (1990). This model proposes an additional overwriting rule for XX0-numbers compared to X0X-numbers. The current study was, of course, not aimed at contributing to the important discussion whether or not semantic access is mandatory during transcoding (development). That said, we would like to point out that mandatory semantic involvement seems rather unlikely as we observed that errors are driven by number structures rather than numerical value.

It might be promising to consider the type of required procedures in addition to the number of required procedures. The procedures required for lexical retrieval might be less demanding than syntactic procedures. Furthermore, as mentioned above, novel procedures such as overwriting decades might be more challenging than already trained procedures such as overwriting units. Our results strongly suggest that children establish their transcoding concepts for three-digit numbers based on earlier experience with two-digit numbers (Byrge, Smith, & Mix, 2014). This would not only explain the overgeneralization of the inversion rule for hundreds (107 for seven hundred) but also the larger difficulties in overwriting decades than units.

Even though accuracy increased from Grade 1 to Grade 2, accuracy patterns did not change. Similar accuracy patterns for number writing and reading confirmed the relevance of number structures for transcoding performance. In line with previous results (Moura et al., 2013), accuracy rates were higher for number reading than writing. This may have two reasons: First, working memory load is higher for number writing than reading. Even though number words were repeated twice, children had to rely on their active mental representations of those number words. During number reading, the Arabic number was always present, allowing continuous monitoring of the transcoding procedure. So far, the impact of working memory on transcoding has only been investigated for number writing (Imbo et al., 2014; Simmons, Willis, & Adams, 2012; Zuber et al., 2009). Future studies should investigate the association of individual differences in working memory with number reading skills. Secondly, procedures in number writing are more demanding than in number reading. Moeller, Zuber, et al. (2015) remarked that overwriting errors such as 20,060 for *two hundred and sixty* are specific to number writing. Hardly any children would read 260 as *twenty thousand and sixty*.

Thus, we agree that complexity levels were different between the two transcoding directions: Encoding Arabic numbers entailing overwriting zeros seemed to be more difficult than decoding multidigit numbers in digital-Arabic notation.

Place-Value Understanding in Early Transcoding

When children are learning to write and read numbers, they have to surmount language-specific and language-independent challenges. Both inversion and additive composition are syntactic principles and are required by the use of a place-value system. To understand that the value of a digit is defined by its position within a sequence of digits is not only a key factor for complex numerical skills but was also found to be related to arithmetic performance (Moeller et al., 2011). The current findings expand earlier evidence (Byrge et al., 2014; Zuber et al., 2009), showing that children experience difficulties in applying the place-value concept when trying to transcode complex numbers that have not yet been formally taught. Thus, an important educational implication of this evidence is that children should get explicit information on multidigit numbers and place value as early as possible in order to prevent misconceptions.

Conclusion

Number word structure has a significant impact on early transcoding. Decade-unit inversion induced more problems with number order in German than English, a language without inversion. Cross-linguistic comparisons revealed language-independent challenges in transcoding with more problems with lexical zeros (e.g., 260) than syntactic zeros (e.g., 206). Importantly, results of our accuracy analysis provide a sound empirical basis to reconsider current transcoding models. First, adaptations to linguistic peculiarities such as decade-unit inversion are needed. For algorithmic transcoding, the ADAPT-model (Barrouillet et al., 2004) suggests that number words are sequentially parsed and constituent digits are retrieved from long-term memory. In German number writing, an additional procedure would be necessary to reorder digits stored in working memory. Second, it is important for research and teaching practices to consider number structures, as they can explain commonalities in accuracy patterns between number writing and reading and can help to reveal specific error sources.

Funding: This work is supported by the Austrian Science Fund [I 2778-G16] and the Economic and Social Research Council [ES/N014677/1].

Acknowledgments: We would like to thank Clare Copper and Viktoria Wirth for help with data collection.

Competing Interests: The authors have declared that no competing interests exist.

Data Availability: For this article, a dataset is freely available (Steiner et al., 2021)

Supplementary Materials

The data supporting the findings of this study were collected as part of a larger longitudinal project. The data-set of the present study is available on the Open Science Framework (for access see [Index of Supplementary Materials](#) below).

Index of Supplementary Materials

Steiner, A. F., Finke, S., Clayton, F. J., Banfi, C., Kemény, F., Göbel, S. M., & Landerl, K. (2021). *Supplementary materials to "Language effects in early development of number writing and reading"* [Research data and code]. OSF. <https://osf.io/nzevr>

Journal of Numerical Cognition. (Ed.). (2021). *Supplementary materials to "Language effects in early development of number writing and reading"* [Open peer-review]. PsychOpen GOLD. <https://doi.org/10.23668/psycharchives.5223>

References

- Bahnmueller, J., Cipora, K., Göbel, S. M., Nuerk, H.-C., & Soltanlou, M. (2021). Pick the smaller number: No influence of linguistic markedness on three-digit number processing. *Journal of Numerical Cognition*, 7(3), 295-307. <https://doi.org/10.5964/jnc.6057>
- 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(2), 368-394. <https://doi.org/10.1037/0033-295X.111.2.368>
- Bundesministerium für Bildung, Wissenschaft und Forschung. (2012). *Lehrplan der Volksschule*. Retrieved from 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>
- 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. <https://doi.org/10.1016/j.jecp.2007.06.006>
- Clayton, F. J., Copper, C., Steiner, A. F., Banfi, C., Finke, S., Landerl, K., & Göbel, S. M. (2020). Two-digit number writing and arithmetic in Year 1 children: Does number word inversion matter? *Cognitive Development*, 56, Article 100967. <https://doi.org/10.1016/j.cogdev.2020.100967>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44(1-2), 1-42. [https://doi.org/10.1016/0010-0277\(92\)90049-N](https://doi.org/10.1016/0010-0277(92)90049-N)
- Deloche, G., & Seron, X. (1987). Numerical transcoding: A general production model. In G. Deloche & X. Seron (Eds.), *Mathematical disabilities: A cognitive neuropsychological perspective* (pp. 137-170). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
- Department for Education. (2013). *The national curriculum in England: Key stages 1 and 2 framework document*. Retrieved from <https://www.gov.uk/government/publications/national-curriculum-in-england-primarycurriculum>
- Dotan, D., & Friedmann, N. (2018). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex*, 101, 249-281. <https://doi.org/10.1016/j.cortex.2017.10.025>
- Dowker, A., & Roberts, M. (2015). Does the transparency of the counting system affect children's numerical abilities? *Frontiers in Psychology*, 6, Article 945. <https://doi.org/10.3389/fpsyg.2015.00945>
- 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>
- Ganayim, D., Ganayim, S., Dowker, A., & Olkun, S. (2020). Linguistic effects on the processing of two-digit numbers. *Open Journal of Modern Linguistics*, 10(1), 49-69. <https://doi.org/10.4236/ojml.2020.101004>
- Habermann, S., Donlan, C., Göbel, S. M., & Hulme, C. (2020). The critical role of Arabic numeral knowledge as a longitudinal predictor of arithmetic development. *Journal of Experimental Child Psychology*, 193(5), Article 104794. <https://doi.org/10.1016/j.jecp.2019.104794>
- Hayek, M., Karni, A., & Eviatar, Z. (2020). Transcoding number words by bilingual speakers of Arabic: Writing multi-digit numbers in a units-decades inverting language. *Writing Systems Research*, 11(2), 1-15. <https://doi.org/10.1080/17586801.2020.1787298>
- Imbo, I., Vanden Bulcke, C., De Brauwier, 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, Article 313. <https://doi.org/10.3389/fpsyg.2014.00313>
- Lopes-Silva, J. B., Moura, R., Júlio-Costa, A., Wood, G., Salles, J. F., & Haase, V. G. (2016). What is specific and what is shared between numbers and words? *Frontiers in Psychology*, 7, Article 22. <https://doi.org/10.3389/fpsyg.2016.00022>
- McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, 44(1-2), 107-157. [https://doi.org/10.1016/0010-0277\(92\)90052-J](https://doi.org/10.1016/0010-0277(92)90052-J)
- 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. <https://doi.org/10.1016/j.ridd.2011.03.012>
- Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H.-C. (2015). Language influences number processing—A quadrilingual study. *Cognition*, 136, 150-155. <https://doi.org/10.1016/j.cognition.2014.11.003>
- 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, Article 740. <https://doi.org/10.3389/fpsyg.2015.00740>
- Moll, K., & Landerl, K. (2010). *Salzburger Lese- und Rechtschreibtest* [Reading and Spelling Test]—Second edition (SLRT-II). Bern, Switzerland: Hans Huber.

- Moura, R., Lopes-Silva, J. B., Vieira, L. R., Paiva, G. M., Prado, A. C. D. A., Wood, G., & Haase, V. G. (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., & Haase, V. G. (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. <https://doi.org/10.1016/j.jecp.2013.07.008>
- Pixner, S., Zuber, J., Heřmanová, 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. <https://doi.org/10.1016/j.ridd.2011.06.004>
- Poncin, A., Van Rinsveld, A., & Schiltz, C. (2019). Units-first or tens-first: Does language matter when processing visually presented two-digit numbers? *Quarterly Journal of Experimental Psychology*, *73*(5), 726-738. <https://doi.org/10.1177/1747021819892165>
- Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, *5*(3), 237-254. <https://doi.org/10.1080/01690969008402106>
- Power, R. J. D., & Dal Martello, M. F. (1997). From 834 to eighty thirty four: The reading of Arabic numerals by seven-year-old children. *Mathematical Cognition*, *3*(1), 63-85. <https://doi.org/10.1080/135467997387489>
- Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*(3), 281-300. <https://doi.org/10.1111/j.2044-835x.1994.tb00635.x>
- Seron, X., & Noël, M.-P. (1995). Transcoding numerals from the Arabic code to the verbal one and vice versa: How many routes? *Mathematical Cognition*, *1*(2), 215-243.
- 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. <https://doi.org/10.1016/j.jecp.2011.08.011>
- Steiner, A. F., Banfi, C., Finke, S., Kemény, F., Clayton, F. J., Göbel, S. M., & Landerl, K. (2021). Twenty-four or four-and-twenty: Language modulates cross-modal matching for multi-digit numbers in children and adults. *Journal of Experimental Child Psychology*, *202*, Article 104970. <https://doi.org/10.1016/j.jecp.2020.104970>
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency—Second edition (TOWRE-2)*. Austin, TX, USA: PRO-ED, Inc.
- van der Ven, S. H. G., Klaiber, J. D., & van der Maas, H. L. J. (2017). 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. <https://doi.org/10.1080/01443410.2016.1150421>
- 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. *British Journal of Developmental Psychology*, *34*(3), 461-468. <https://doi.org/10.1111/bjdp.12151>
- Verguts, T., & Fias, W. (2005). Interacting neighbors: A connectionist model of retrieval in single-digit multiplication. *Memory and Cognition*, *33*(1), 1-16. <https://doi.org/10.3758/BF03195293>
- Wechsler, D. (2005). *Wechsler Individual Achievement Test—Second UK Edition (WIAT-II UK)*. London, England: Harcourt Assessment.
- 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. <https://doi.org/10.1016/j.jecp.2008.04.003>

Appendix

Table A.1

One-, Two- and Three-Digit Numbers of Both Transcoding Tasks, Which Were Written and Read in Grades 1 and 2

Number Structure	Number of Items	Items
One-digit numbers		
X	4	1, 7, 4, 6
Two-digit numbers		
Teens (1X)	4	11,15, 16, 13
X0	4	80, 40, 70, 30
XX	16	73, 42, 34, 81, 32, 89, 53, 64, 25, 68, 56, 91, 48, 27, 79, 47
Three-digit numbers		
X00	4	700, 300, 200, 600
X0X	4	203, 107, 304, 109
XX0	4	560, 340,190, 220,
XXX	4	242, 349, 123, 643

Note. One-digit numbers were only presented at T1.



Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).



leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.