Language and the Algebraic Mind: Learning Unfamiliar Rules

Qi Zhou (Qi.Zhou23@student.xjtlu.edu.cn)
Academy of Future Education, Xi'an Jiaotong-Liverpool University
Suzhou, 215123, China

Jike Qin (Jike.Qin@xjtlu.edu.cn)
Academy of Future Education, Xi'an Jiaotong-Liverpool University
Suzhou, 215123, China

Daniel Roberts (D.J.Roberts@liverpool.ac.uk)
Department of Psychology, University of Liverpool
Liverpool, L69 7ZX, UK

Abstract

Natural language is often considered fundamental to mathematical thinking, a view supported by research on language-of-training effects in bilinguals. However, these effects have been primarily examined in arithmetic. While previous research has explored language effects in algebra, it has largely focused on well-established rules. This study extends prior work by investigating whether similar effects apply to learning of algebra with unfamiliar rules. Thirty-nine Chinese-English bilingual undergraduates were trained to solve arithmetic and algebraic problems in Chinese or English and were later tested on both old and novel problems in both languages. Consistent with previous findings, results revealed a clear dissociation between arithmetic and algebra. Bilinguals responded faster in the trained versus the untrained language for arithmetic problems and solved old arithmetic problems faster than novel ones. However, these effects were absent in algebra, suggesting that algebraic learning does not necessarily depend on natural language encoding.

Keywords: language; algebra; arithmetic; bilingual training

Introduction

Mathematics and language are two unique human cognitive abilities, and their relationship has been a topic of debate in cognitive science. A central question is whether language is a prerequisite for mathematical thinking (Carey, 2001; Dehaene et al., 1999; Gordon, 2004; Pica et al., 2004; Spelke, 2003).

One perspective suggests that natural language refines an innate, approximate sense of numbers, enabling exact mathematical thinking about large numbers (Dehaene et al., 1999; Spelke & Tsivkin, 2001). In a seminal study, Spelke and Tsivkin (2001) trained eight Russian-English bilinguals with multiple sets of exact (e.g., "What is the sum of fifty-four and forty-eight?") and approximate (e.g., "Estimate the approximate cube root of twenty-nine.") mathematical statements presented in one language and then tested them in both languages. Their findings revealed a language switch effect in exact arithmetic: participants solved problems faster in the trained than the untrained language and were faster for trained than novel problems. In contrast, approximate arithmetic was unaffected by language or problem novelty. Subsequent studies have replicated these findings across different bilingual populations, including Chinese-English solving single-digit arithmetic problems (Chen et al., 2016) and German-English performing auditory arithmetic tasks (Hahn et al., 2019).

One possible interpretation of these findings is that natural language serves as a medium for mathematical thinking that extends beyond the limitations of the approximate number system. As Spelke (2003) concludes, "Human knowledge of number appears to be quintessentially abstract. The concept of 'seven' appears to transcend any of the particular sets of seven entities that a person enumerates, the particular situations in which she enumerates them, and (one would think) the particular language in which she expresses this enumeration. However, our findings suggest that 'seven' is a language-dependent concept, distinct from the Russian *sem* or the French *sept*" (Spelke & Tsivkin, 2001, p. 81).

A previous study (Qin & Opfer, 2018) challenged this view by extending Spelke and Tsivkin (2001)'s paradigm to algebraic problems, which emphasize patterns, relationships, and structures—core features of algebra (Lang, 1993/2002) . The authors argued that if algebraic thinking showed no language-of-training effects, it would constitute strong evidence against the language-dependency view of mathematical cognition. In their study, Chinese-English bilingual adults were trained to solve arithmetic problems (multiplication, e.g., twenty-one · five = one hundred five) and algebraic problems (distributive law transformations, e.g., twelve · twenty-six = twelve · twenty + twelve · six) in either English or Chinese. Participants were then tested on both problems in both languages. Results revealed the language-of-training and problem-novelty effects emerged in arithmetic but were absent in algebra. Based on these findings, the authors concluded that when mathematics is learned through rule-based reasoning that emphasizes relational structures, language plays no role.

However, a key limitation of this study may affect the interpretation of its findings. Specifically, the algebra tasks involved an overlearned rule (i.e., the distributive law), creating a substantial disparity in difficulty between arithmetic and algebra tasks. Notably, participants took

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

nearly twice as long to solve arithmetic problems compared to algebraic problems. Although task difficulty was controlled, differences in cognitive load may have influenced strategy selection. For instance, familiarity with the distributive law might have facilitated automatic processing in the algebra task, reducing reliance on language.

The Current Study

To address this limitation, the present study investigates whether algebraic learning depends on natural language by introducing a novel algebraic rule to bilingual adults. We conducted a near replication of Spelke and Tsivkin (2001) and Qin and Opfer (2018), asking Chinese-English adults to solve arithmetic or algebra problems. Unlike Qin and Opfer (2018), who used the well-known distributive law, our study introduced an artificial rule (i.e., A#B=A×B+A−B), requiring participants to *learn* it through training. This rule reflects a particularly important algebraic topic—function—which is used to express a relationship between variables (Mielicki et al., 2017). If only arithmetic knowledge is encoded through language, we would expect language-of-training and problem novelty effects for arithmetic problems but not for algebra problems, which can be solved by applying algebraic rules through visual pattern recognition.

Furthermore, we investigated how learners prioritize different aspects of a rule when acquiring a novel algebraic operation. Algebra involves two key components: operators (e.g., multiplication, addition, and their operational sequence) and operands (how numerical values are applied across operations). To determine whether learners focus more on operators or operands, we designed incorrect answer choices that violated either the sequence of operators or the sequence of operands. For the problem "A#B=", the correct answer follows the rule "A×B+A−B". If participants prioritize the relational structure of operators (×+−) over the specific arrangement of operands (ABAB), they should reject *operator-violating* foils (e.g., A+B+A−B) faster than *operand-violating* foils (e.g., A×B+B−A).

Methods

Participants

A total of 39 undergraduates (M_{age} = 20.58 years; 23 females) from an English-medium university in Suzhou, China, participated in this study. Their language profiles were assessed using the Chinese version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). All participants were unbalanced Chinese-English bilinguals, having acquired Chinese earlier and with greater immersion in it than in English. Despite these differences, they reported medium-to-high proficiency in both languages on a 0 to 10 scale: speaking (M_{CH} = 8.67, M_{EN} = 6.08), understanding (M_{CH} = 8.85, M_{EN} = 6.59), and reading (M_{CH} = 8.85, M_{EN} = 6.79). For every participant in the study, the medium of instruction for mathematics courses changed from Chinese to English after their university admission.

Design

Following previous studies (Qin & Opfer, 2018; Spelke & Tsivkin, 2001), the experiment comprised a two-day training and one-day test session. During training, participants were randomly assigned to one of two groups. In one group, they performed the arithmetic task in Chinese and the algebra task in English, while in the other group, this was reversed. In the test session, all participants performed both tasks in both languages. The order of languages was counterbalanced across participants in all sessions.

Tasks and Materials

Arithmetic Task

Participants were asked to complete multiplication problems (see Table 1 for examples) with half involving one-digit multiplication and half one-digit and two-digit multiplication. Answers were two-digit numbers ranging 14 to 98. Incorrect answers were generated by adding or subtracting 2 for one-digit multiplication problems and 10 for one-digit and two-digit multiplication problems. To establish a language processing environment, problems were presented using Chinese and English words (e.g., "一" or "one").

Algebra Task

Participants were asked to learn an artificial algebraic rule, "A#B=A×B+A−B," through trial-and-error. This was chosen to introduce a novel learning environment for adults. Incorrect answers were categorized as either operator-violating or operand-violating foils to analyze algebra learning patterns. In the operator-violating condition, one of the three operators was altered while the operands remained unchanged (e.g., A+B+A−B). In the operand-violating condition, the positions of the latter A and B were swapped while the operator remained unchanged (e.g., A×B+B−A). As with the arithmetic problems, algebraic problems were presented using Chinese or English number words.

Table 1: Examples of Problems.

	English	Chinese
Arithmetic	Seven × Twelve = Eighty-four Seventy-four	七 × 十二 = 八十四 七十四
Algebra	Six # Eight = Six × Eight + Six − Eight Six × Eight + Eight − Six	六 # 八 = 六 × 八 + 六 − 八 六 × 八 + 八 − 六

Procedure

The experiment was programmed in PsychoPy (Peirce et al., 2019) and conducted on a MacBook. To ensure participants were fully immersed in the target language before each task, the experimenter briefly conversed with them and asked them to browse news websites for three minutes at the beginning of each training and test session.

For training sessions, each task consisted of 12 problems, repeated 6 times, totaling 72 trials. In each trial, the problem was displayed at the upper center of the screen, with two candidate answers vertically aligned at the lower center.

Participants were asked to select the correct answer using the "up" or "down" arrow keys; correct answers appeared equally in the top and bottom positions across trials. The problem sequence was fixed but arranged in a pseudo-random pattern to prevent consecutive repetition of the same problem. Each task began with instructions and example problems in the trained language. After pressing the space bar, each trial started and remained on the screen for 10 seconds or until a response was made. Immediately after responding (or after timeout), feedback was displayed for 600ms, indicating "Correct", "Incorrect" or "Time out". Participants were instructed to respond as quickly and accurately as possible.

The test session followed the same procedure as the training sessions, except that 6 trained problems and 6 novel problems were included for both tasks. Each problem was presented twice, resulting in 24 trials in total.

Results

1. Dissociation Between Arithmetic and Algebra Tasks Over Three Days.

Response times (RTs) for correctly answered trials were analyzed using a generalized linear mixed-effect model (GLMM), implemented with the *lme4* package (Bates et al., 2015) in R (Version 4.4.2) (analyses of accuracy data revealed a similar pattern). The model included the following fixed effects: task (arithmetic/algebra), language (Chinese/English), 3-day period (day 1/2/3), language switch (no switch/switch), problem novelty (old/new), and all the possible interactions. A maximal random-effects structure was applied following the recommendations of Barr et al. (2013) and Judd et al. (2017) under the premise of model converging, with random intercepts and slopes of task, language, and language switch on participants, and random intercepts and slopes of language on stimuli. The variable "day" was coded using a backward difference coding scheme, distinguishing between "day 2 vs. day 1" and "day 3 vs. day 2". The other variables were coded as follows—task: arithmetic = 0.5, algebra = -0.5; language: English = 0.5, Chinese = -0.5; language switch: switch = 0.5, no switch = -0.5; problem novelty: new = 0.5, old = -0.5.

Figure 1 displays overall task performance across the three days, and Table 2 presents the detailed model results. A non-significant main effect of task indicated that participants responded at similar speeds to both arithmetic and algebraic problems. Significant main effects were found for language, day, and language switch, indicating longer RTs for English compared to Chinese (2761 vs. 2007ms, $p<.001$), on day 1 compared to day 2 (2698 vs. 2124ms, $p<.001$), on day 3 compared to day 2 (2357 vs. 2124ms, $p<.05$), and during language switch compared to no switch (2497 vs. 2219ms, $p<.001$).

Importantly, (marginally) significant two-way interactions were observed between task and each of the other four variables, suggesting dissociations between arithmetic and algebra tasks over the 3-day period. Three-way interactions were present among task, language, and day, and task, language, and language switch, reflecting the fact that the task × language interaction differed by day, and that the task × language switch interaction on day 3 varied across languages.

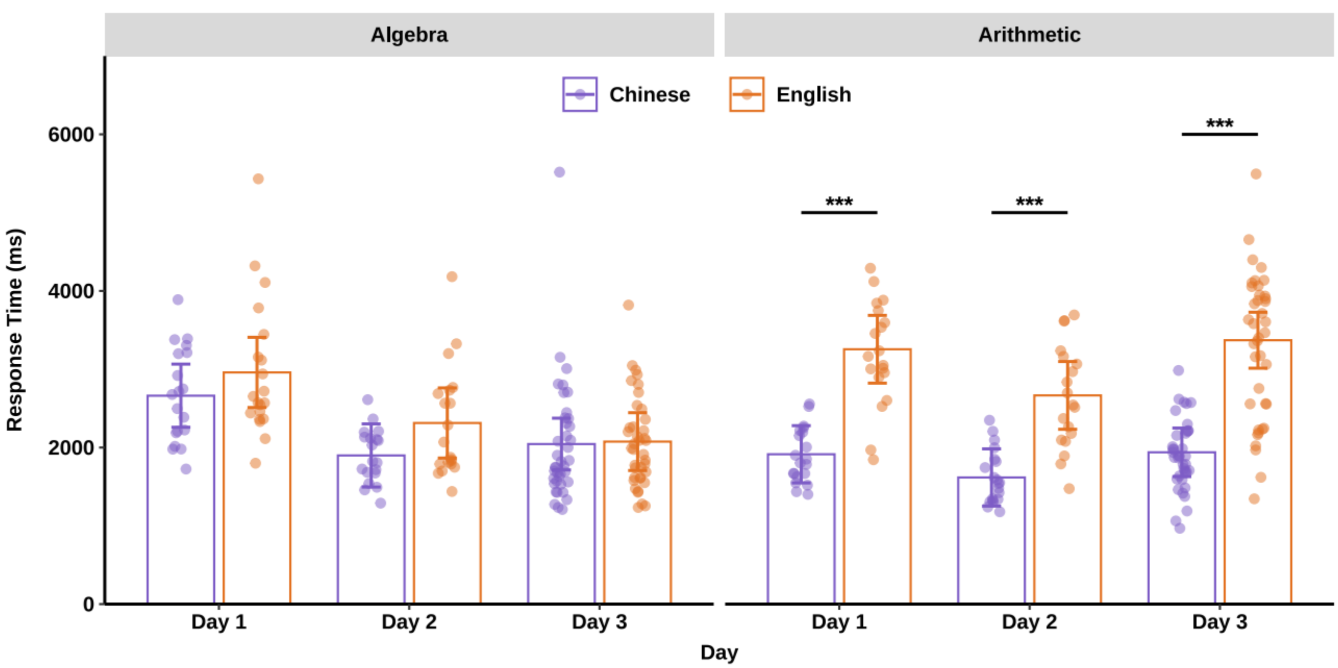

Figure 1: RTs Across Days for Algebra and Arithmetic Tasks by Language Group. The dots represent the raw individual mean, while the bar chart and the error bar represent the estimated group mean and 95% CI based on the model results. *** $p<.001$.

Table 2: GLMM Analyses Predicting RTs from Task, Day, Language, Language Switch, and Problem Novelty.

Predictor	B	SE_B	β	t	p
(intercept)	2.39	0.13	.04	18.89	<.001***
Task	**0.14**	**0.22**	**.10**	**0.63**	**.53**
Language	**0.76**	**0.08**	**.55**	**9.39**	**<.001*** **
Day 2 − Day 1	**−0.57**	**0.02**	**−.42**	**−28.82**	**<.001*** **
Day 3 − Day 2	**0.23**	**0.10**	**.17**	**2.30**	**<.05***
Language Switch	**0.28**	**0.06**	**.03**	**4.76**	**<.001*** **
Problem Novelty	0.29	0.20	.03	1.44	.16
Task × Language	**1.03**	**0.23**	**.75**	**4.52**	**<.001*** **
Task × (Day 2 − Day 1)	**0.26**	**0.04**	**.19**	**6.58**	**<.001*** **
Task × (Day 3 − Day 2)	**0.56**	**0.21**	**.41**	**2.71**	**<.01*** **
Task × Language Switch	**0.18**	**0.10**	**.02**	**1.76**	**.08#**
Task × Problem Novelty	**0.80**	**0.40**	**.08**	**1.97**	**.06#**
Language × (Day 2 − Day 1)	−0.09	0.04	−.06	−2.21	<.05*
Language × (Day 3 − Day 2)	0.00	0.10	−.00	0.00	.99
Language × Language Switch	0.13	0.17	.01	0.77	.44
Language × Problem Novelty	0.09	0.12	.01	0.74	.46
Language Switch × Problem Novelty	−0.18	0.07	−.00	−2.65	<.01**
Task × Language × (Day 2 − Day 1)	**−0.41**	**0.08**	**−.30**	**−5.15**	**<.001*** **
Task × Language × (Day 3 − Day 2)	**0.77**	**0.37**	**.56**	**2.09**	**<.05***
Task × Language × Language Switch	**1.48**	**0.64**	**.16**	**2.31**	**<.05***
Task × Language × Problem Novelty	−0.17	0.23	−.02	−0.73	.47
Task × Language Switch × Problem Novelty	−0.21	0.14	−.01	−1.53	.13
Language × Language Switch × Problem Novelty	−0.04	0.14	−.00	−0.29	.77
Task × Language × Language Switch × Problem Novelty	0.16	0.27	.00	0.57	.57

Note. Results discussed in the main text are indicated in bold. Coding scheme: Task: arithmetic = 0.5, algebra = –0.5; Language: English = 0.5, Chinese = –0.5; Language Switch: switch = 0.5, no switch = –0.5; Problem Novelty: new = 0.5, old = –0.5; Day was coded using a backward difference coding scheme, distinguishing between "day 2 vs. day 1" and "day 3 vs. day 2".
# p<.10. * p<.05. ** p<.01. *** p<.001.

To better understand these dissociations, we broke down the interactions using the *emmeans* package in R. The results showed that arithmetic problems were solved faster in Chinese than in English (1852 vs. 3166ms, p<.001), whereas algebraic problems were solved at similar speeds in both languages (2162 vs. 2356ms, p=.38). Also, for arithmetic problems, RTs decreased from day 1 to day 2 (2585 vs. 2142ms, p<.001), and then increased from day 2 to day 3 (2142 vs. 2655ms, p<.01), whereas for algebraic problems, RTs dropped more dramatically from day 1 to day 2 (2811 vs. 2106ms, p<.001), but remained similar from day 2 to day 3 (2106 vs. 2060ms, p=.99). The task × language × (day 2 – day 1) and task × language × (day 3 – day 2) interactions further revealed that for arithmetic problems, the decrease of RTs from day 1 to day 2 (Chinese vs. English: 296 vs. 589ms), and the increase of RTs from day 2 to day 3 (Chinese vs. English: 321 vs. 705ms) were smaller in Chinese than English, whereas for algebraic problems, the decrease from day 1 to day 2 was similar in Chinese and English (763 vs. 646ms), and no significant difference was found between day 2 and day 3 in both languages (Chinese: difference = 145ms, p=.99; English: difference = 237ms, p=.99) (see Figure 1).

Most notably, on day 3, participants trained in English solved English arithmetic problems significantly faster than those trained in Chinese (2971 vs. 3773ms, p<.01). However, no language switch effects were observed for Chinese arithmetic problems (1975 vs. 1906ms, p=.99) or for algebraic problems in either language (Chinese: 1798 vs. 2290ms, p=.21; English: 2132 vs. 2020ms, p=.99; Figure 2A). Additionally, participants tended to solve old arithmetic problems faster than novel ones (2312 vs. 3000ms, p=.08), but solved old and new algebraic problems equally fast (2114 vs. 2006ms, p=.98; Figure 2B). These findings suggest that encoding specificity effects were considerably stronger for arithmetic than for algebra.

2. Operator Priority Under Specific Conditions.

To investigate whether learners prioritize operators or operands when acquiring a novel algebraic rule, RTs were analyzed using GLMM (analyses of accuracy data revealed a similar pattern), with foil type (operand-violating/operator-violating), language (Chinese/English), 3-day period (day 1/2/3), problem novelty (old/new), and all the potential interactions as fixed effects. The maximal random-effects structure included random intercepts and slopes of foil type and language on participants, random intercepts and slopes of language on stimuli. The coding for day, language, and problem novelty followed the same scheme as in the previous models. For foil type, operand-violating was coded as -0.5 and operator-violating was coded as 0.5.

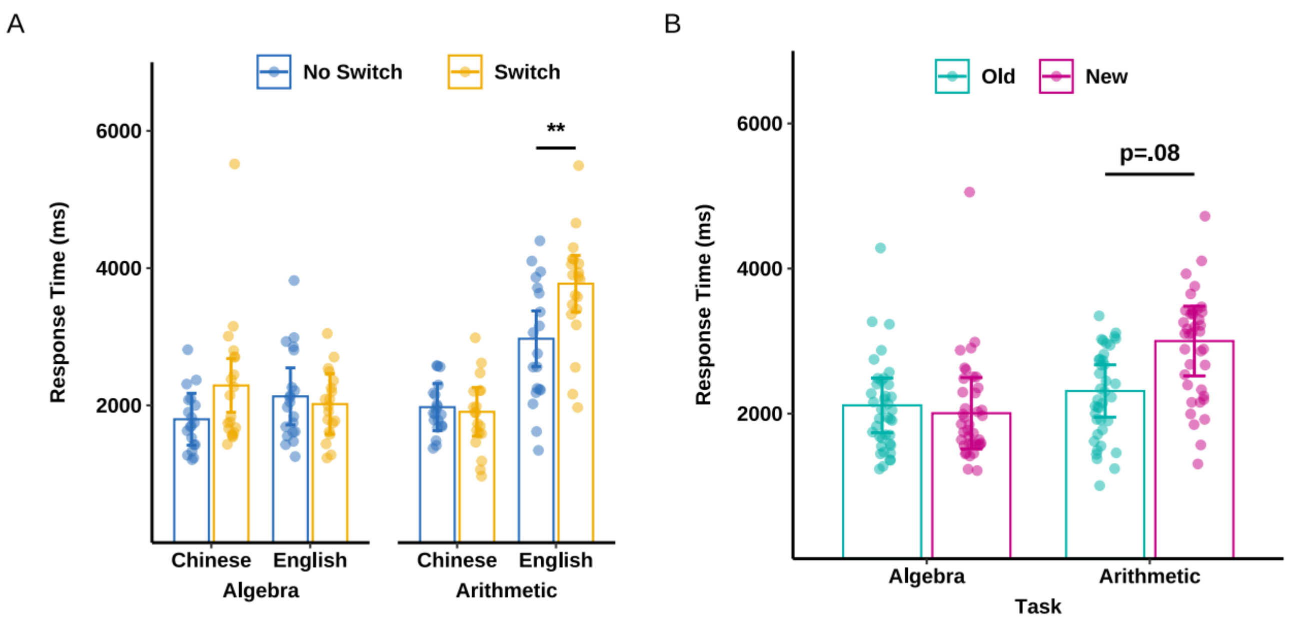

Figure 2: Interactions for Language Switch Effect (A) and Problem Novelty Effect (B) on RTs on Day 3. The dots represent the raw individual mean, while the bar chart and the error bar represent the estimated group mean and 95% CI based on the model results. Chinese and English in the panel A refers to the test language. ** p<.01.

Table 3 shows the detailed model results. A significant two-way interaction was found between foil type and language, indicating that the operator-violating problems were solved faster than operand-violating problems in English (2163 vs. 2476ms, p<.05), but not in Chinese (2308 vs. 2266ms, p=.43). This interaction was further moderated by the comparison between day 2 and day 1, as well as by problem novelty. Specifically, the difference between English operand-violating and operator-violating on day 2 was larger than that on day 1 (335 vs. 90ms), but neither difference reached significant level. This difference was also observed in old English problems on day 3, but not in novel problems (old: 645ms, p<.001; novel: 183ms, p=.96). In contrast, no such differences were found for Chinese algebra problems across days (differences of 32ms on day 1, 116ms on day 2, 3ms for novel problems on day 3, and 79ms for old problems on day 3, all ps=.99). These results suggest that, for only English problems, familiarity with the task could influence how participants focus on and process the rule: with more exposure, participants come to prioritize the operator structure over the operand.

Table 3: GLMM Analyses Predicting RTs from Foil Type, Day, Language, and Problem Novelty

Predictor	B	SE_B	β	t	p
(intercept)	2.38	0.12	.05	20.40	<.001***
Foil Type	−0.12	0.10	−.10	−1.21	.23
Language	0.04	0.08	.03	0.46	.65
Day 2 − Day 1	−0.70	0.03	−.57	−25.67	<.001***
Day 3 – Day 2	−0.10	0.05	−.08	−2.26	<.05*
Problem Novelty	−0.10	0.08	−.01	−1.34	.19
Foil Type × Language	**−0.32**	**0.11**	**−.26**	**−2.80**	**<.05***
Foil Type × (Day 2 – Day 1)	−0.05	0.05	−.04	−0.89	.37
Foil Type × (Day 3 – Day 2)	−0.08	0.09	−.06	−0.90	.37
Foil Type × Problem Novelty	0.19	0.15	.02	1.26	.22
Language × (Day 2 − Day 1)	0.12	0.05	.10	2.26	<.05*
Language × (Day 3 – Day 2)	−0.09	0.09	−.07	−0.99	.32
Language × Problem Novelty	0.17	0.13	.02	1.30	.20
Foil Type × Language × (Day 2 – Day 1)	**−0.39**	**0.11**	**−.32**	**−3.59**	**<.001***
Foil Type × Language × (Day 3 – Day 2)	−0.00	0.17	−.00	−0.03	.98
Foil Type × Language × Problem Novelty	**0.54**	**0.27**	**.07**	**2.02**	**<.05***

Note. Results discussed in the main text are indicated in bold. Coding scheme: Foil Type: operator-violating = 0.5, operand-violating = –0.5; Language: English = 0.5, Chinese = –0.5; Problem Novelty: new = 0.5, old = –0.5; Day was coded using a backward difference coding scheme, distinguishing between “day 2 vs. day 1” and “day 3 vs. day 2”.
* p<.05. ** p<.01. *** p<.001.

Discussion

This study adds to the growing body of research on the role of language in mathematical thinking by examining arithmetic and algebraic learning in a bilingual context. The findings highlighted a clear dissociation between the two, with arithmetic showing a reliance on language, as evidenced by the language switch and problem novelty effects, and algebra demonstrating independency from linguistic encoding. Furthermore, our analysis of algebraic learning revealed that participants prioritize the operators when solving the trained English algebraic problems.

Aligning with previous studies (Chen et al., 2016; Hahn et al., 2019; Qin & Opfer, 2018; Spelke & Tsivkin, 2001), our results demonstrate that arithmetic relies on language and tends to be rote-based. However, the directionality of the language switch effect we observed diverges from previous findings. Specifically, while earlier research reported a bidirectional switch cost (Spelke & Tsivkin, 2001), our results showed a significant switch cost only when participants transitioned from Chinese to English. This asymmetry, consistent with findings from other studies involving unbalanced bilinguals (Chen et al., 2016; Saalbach et al., 2013), may suggest that participants relied on their dominant language (Chinese) during arithmetic training in non-dominant language (English). They may have translated English problems into Chinese for computation or fact retrieval and then recoded the answers into English for output. This translation step was unnecessary when switching from English to Chinese, potentially explaining why participants trained in English performed better when tested in Chinese.

In contrast, algebraic learning showed no evidence of a language switch effect, supporting the hypothesis that algebraic thinking is independent of linguistic processing. This finding is consistent with brain injury and neuroimaging evidence indicating separate neural mechanisms for algebra and language (Klessinger et al., 2007; Monti et al., 2012). For example, Monti et al. (2012) identified a clear neural dissociation using fMRI—language syntax activated classical left-hemisphere language regions, while algebraic operations engaged bilateral parietal areas responsible for magnitude processing. Furthermore, the absence of language switch and problem novelty effects in algebra suggests that algebraic reasoning does not rely on verbal encoding of operands or numerical values. Instead, prior research suggested that algebraic thinking is grounded in perceptual and visuo-spatial encoding (Harrison et al., 2020; Landy & Goldstone, 2007; Marghetis et al., 2016). Participants may have relied on visual representations to process the structure of algebraic rules, bypassing the need for verbal mediation.

The introduction of a novel algebraic rule in this study provided valuable insights into the learning process, revealing that learners prioritize operator processing under specific conditions. Faster RTs were observed for operator-violating foils in English problems, but not Chinese, and only for trained not untrained problems. In English, number words are processed linearly from left to right, mirroring the direction of reading, which likely reinforces attention to the order of operators. In contrast, Chinese number words are compact and visually distinct, not following a linear phonetic structure, which may increase the salience of operands and reduce the focus on operators. Additionally, familiarity with the problems may have directed attention toward the operator structure. For new problems, however, the rule may still be inadequately internalized, leading to less differentiations between operator-violating and operand-violating foils.

Collectively, our study advances theoretical understanding by challenging the view that language is essential for all forms of mathematical cognition. Specifically, we demonstrate that learning novel algebraic rules can proceed independently of language, providing critical behavioral evidence that addresses key limitations in prior research. These findings also have implications for mathematics education. The strong link between language and arithmetic suggests that instructional strategies emphasizing verbal explanations and linguistic reinforcement could enhance arithmetic learning. Conversely, algebra instruction might benefit from approaches that foster abstract reasoning and visuospatial problem-solving skills, potentially incorporating visual aids or manipulative tools to reinforce relational thinking.

In conclusion, this study highlights the distinct roles of language in arithmetic and algebra. While arithmetic is intertwined with linguistic processing, algebra engages distinct nonlinguistic cognitive processes. Recognizing these differences could inform curriculum design and teaching strategies, ultimately supporting more effective learning across mathematical domains.

Acknowledgments

This study was funded by XJTLU Research Development Funding RDF-21-02-063.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48.

Carey, S. (2001). Whorf versus continuity theorists: Bringing data to bear on the debate. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge University Press.

Chen, Y., Yanke, J., & Campbell, J. I. D. (2016). Language-specific memory for everyday arithmetic facts in Chinese-English bilinguals. *Psychonomic Bulletin & Review*, *23*(2), 526–532.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974.

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2019). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57.

Harrison, A., Smith, H., Hulse, T., & Ottmar, E. R. (2020). Spacing out! Manipulating spatial features in mathematical expressions affects performance. *Journal of Numerical Cognition*, *6*(2), 186–203.

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with More Than One Random Factor: Designs, Analytic Models, and Statistical Power. *Annual Review of Psychology*, *68*, 601–625.

Klessinger, N., Szczerbinski, M., & Varley, R. (2007). Algebra in a man with severe aphasia. *Neuropsychologia*, *45*(8), 1642–1648.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 720–733.

Lang, S. (2002). *Algebra*. Springer. (Original work published in 1993).

Marghetis, T., Landy, D., & Goldstone, R. L. (2016). Mastering algebra retrains the visual system to perceive hierarchical structure in equations. *Cognitive Research: Principles and Implications*, *1*(1), 25.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

Mielicki, M. K., Kacinik, N. A., & Wiley, J. (2017). Bilingualism and symbolic abstraction: Implications for algebra learning. *Learning and Instruction*, *49*, 242–250.

Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought Beyond Language: Neural Dissociation of Algebra and Natural Language. *Psychological Science*, *23*(8), 914–922.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and Approximate Arithmetic in an Amazonian Indigene Group. *Science*, *306*(5695), 499–503.

Qin, J., & Opfer, J. E. (2018). Rote versus Rule: Revisiting the Role of Language in Mathematical Thinking. *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 918–923). Madison, WI: Cognitive Science Society.

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction*, *26*, 36–44.

Spelke, E. S. (2003). What Makes Us Smart? Core Knowledge and Natural Language. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind: Advances in the Study of Language and Thought*. The MIT Press.

Spelke, E. S., & Tsivkin, S. (2001). Language and number: A bilingual training study. *Cognition*, *78*(1), 45–88.