Cognition 214 (2021) 104789

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Short Communication

Interference between naïve and scientific theories occurs in mathematics and is related to mathematical achievement

Johannes Stricker [a,*], Stephan E. Vogel [b], Silvia Schöneburg-Lehnert [c], Thomas Krohn [c], Susanne Dögnitz [c], Nina Jud [b], Michele Spirk [b], Marie-Christin Windhaber [b], Michael Schneider [a], Roland H. Grabner [b]

[a] University of Trier, Trier, Germany
[b] University of Graz, Graz, Austria
[c] University of Leipzig, Leipzig, Germany

ARTICLE INFO

Keywords:
Conceptual change
Naïve theory
Mathematics
Education

ABSTRACT

When students learn a scientific theory that conflicts with their earlier naïve theories, the newer and more correct knowledge does not always replace the older and more incorrect knowledge. Both may coexist in a learner's long-term memory. Using a new speeded reasoning task, Shtulman and Valcarcel (2012) showed that naïve theories interfere with retrieving scientific theories. Although mathematics learning is a central aim of schooling and a vital prerequisite for success in life, no study has tested whether Shtulman and Valcarcel's (2012) findings generalize to mathematical subdomains such as algebra, geometry, and probability. Additionally, it is unclear how the interference strength relates to domain-specific and domain-general competencies. We investigated these questions using the speeded reasoning task with new mathematical items in a sample of 62 university students. Solution rates and reaction times indicated interference between naïve and scientific mathematical theories. Additionally, interference strength was inversely related to mathematical achievement and unrelated to general inhibitory control. After controlling for general inhibitory control, mathematical achievement was still substantially related to interference strength. These findings indicate that interference strength reflects domain-specific achievement rather than domain-general inhibitory control.

1. Introduction

Conceptual change has long been viewed as a process in which an initial naïve theory is discarded, replaced, modified, or restructured (Potvin, 2013). However, to date, a growing body of research shows that naïve and scientific theories *coexist* in a learner (e.g., Goldberg & Thompson-Schill, 2009) and that naïve theories *interfere* with operating scientific theories (e.g., Shtulman & Valcarcel, 2012). This interference between naïve and scientific theories has been demonstrated for different scientific domains such as astronomy, genetics, or thermodynamics (Shtulman & Valcarcel, 2012). Concerning mathematics, several studies indicate the presence of hindering misconceptions, e.g., in geometry (Babai, Zilber, Stavy, & Tirosh, 2010; Stavy & Babai, 2010), fractions (DeWolf & Vosniadou, 2015), decimal proportions (Varma & Karl, 2013), and arithmetical operations (Vamvakoussi, Van Dooren, & Verschaffel, 2013). However, to date, no study has systematically assessed interference between naïve and mathematical theories across a broad spectrum of mathematical subdomains. Additionally, it is unclear whether individual differences in interference strength reflect domain-specific or domain-general competencies. Regarding domain-specific competencies, previous research shows that persons with high domain-specific proficiency show lower discrepancies in judging the correctness of intuition-consistent vs. intuition-inconsistent statements (Shtulman & Harrington, 2016). Thus, interference strength may be positively linked to domain-specific achievement. Regarding domain-general competencies, inhibitory control (i.e., the ability to suppress irrelevant information and inhibit incorrect responses) may be recruited in tasks that require rejection of naïve scientific concepts (e.g., Vosniadou et al., 2018, for opposing findings, see Kelemen, Rottman, & Seston, 2013) and is thought to play a central role in conceptual learning more broadly (Mason & Zaccoletti, 2020).

In this study, we aimed to test whether naïve theories interfere with retrieving mathematical theories across a broad range of mathematical

* Corresponding author at: University of Trier, Psychology Department, Universitaetsring 15, 54286 Trier, Germany.
E-mail address: stricker@uni-trier.de (J. Stricker).

https://doi.org/10.1016/j.cognition.2021.104789
Received 27 October 2020; Received in revised form 19 May 2021; Accepted 20 May 2021
Available online 29 May 2021

subdomains. Additionally, we investigated whether the interference strength is related to mathematical achievement and general inhibitory control.

1.1. Interference of naïve and scientific theories in learners

Before formal instruction, learners acquire naïve theories through observation and interaction with their environment in everyday life (Vosniadou, 1994). Although naïve theories are useful for making preliminary sense of one's environment, they are usually at odds with the theories accepted by the scientific community. Interestingly, even after many years of formal education, naïve theories seem to be resilient to change in that they coexist with scientific theories in a learner's long-term memory (Shtulman & Legare, 2020).

In a novel paradigm, Shtulman and Valcarcel (2012) showed that college students are less accurate and slower in judging statements that are true according to naïve theories and false according to scientific theories (or vice versa; *inconsistent statements*; e.g., “Air is composed of matter.“) compared to statements that are true or false according to naïve and scientific theories (*consistent statements*; e.g., “Rocks are composed of matter.“). This finding suggests that a cognitive conflict occurs when learners retrieve scientific information that contradicts naïve theories.

Several studies have identified hindering misconceptions in mathematics (e.g., the whole number bias; DeWolf & Vosniadou, 2015). However, to date, research using Shtulman and Valcarcel's (2012) paradigm on interference strength in mathematics has been limited to fractions. Thus, the degree of interference between naïve—and mathematical theories across a wide range of mathematical subdomains is largely unclear.

1.2. Conceptual change and the resilience of naïve theories in mathematics

Why may conceptual change in mathematics be similar to—or different from conceptual change in other scientific domains? On the one hand, conceptual change in mathematics may be similar to other scientific domains because learners hold naïve theories that conflict with operating mathematical theories (e.g., Vamvakoussi et al., 2013). On the other hand, conceptual change in mathematical domains could differ from other scientific domains because naïve theories in mathematical domains may be less strongly reinforced through learners' perceptual experiences. For example, the sun seemingly moves across the sky, and a strong breeze feels subjectively cold (Shtulman & Harrington, 2016; Shtulman & Legare, 2020). In contrast to these examples from astronomy and physics, naïve theories in mathematical subdomains are more formal and abstract. Consequently, naïve mathematical theories may be less resilient to change compared to naïve theories in other domains.

1.3. Correlates of the interference strength

There are substantial individual differences in how strongly naïve theories interfere with scientific theories (e.g., Shtulman & Harrington, 2016). Only few studies have investigated variables that account for these individual differences in interference strength. Regarding domain-specific achievement, groups with different achievement levels have been compared. For example, one study showed that science professors display a lower discrepancy in judging the correctness of intuition-consistent and intuition-inconsistent scientific statements than humanities professors and non-professors (Shtulman & Harrington, 2016).

Regarding domain-general competencies, general inhibitory control has been investigated in relation to the endorsement of naïve theories (Kelemen & Rosset, 2009). For example, a recent fMRI study in domain-specific experts (i.e., professors) showed increased activation in brain areas related to inhibitory control in inconsistent compared to consistent trials in the Shtulman and Valcarcel (2012) task (Potvin, Malenfant-Robichaud, Cormier, & Masson, 2020). Additionally, in a study with school children, a behavioral general inhibitory control measure was related to performance in a task that required using science/mathematics concepts inconsistent with naïve concepts (Vosniadou et al., 2018). However, in a different study, performance in a behavioral general inhibitory control task was unrelated to the endorsement of naïve teleological statements (Kelemen et al., 2013). Most previous studies used Stroop-like tasks to capture general inhibitory control. In Stroop-like tasks, participants need to ignore irrelevant *external* stimuli (e.g., color-words when naming colors). This process may only partly capture inhibition related to competing *internal* conceptual representations. Additionally, previous studies often tested explicit endorsement of naïve theories. Thus far, no study has tested whether interference assessed with the Shtulman and Valcarcel (2012) task is related to behavioral measures of general inhibitory control. Due to the substantial mathematical achievement - inhibitory control link (e.g., Gilmore et al., 2013), it is also unclear whether interference strength is related to mathematical achievement after controlling for inhibitory control.

1.4. The present study

The present study aimed to test three hypotheses formulated based on the literature reviewed above. First, we hypothesized that naïve and scientific mathematical theories interfere in a broad range of mathematical subdomains (Hypothesis 1). Second, we expect that mathematical achievement (Hypothesis 2a) and inhibitory control (Hypothesis 2b) are related to the interference strength. Third, we hypothesized that mathematical achievement is uniquely related to interference strength beyond the contribution of inhibitory control (Hypothesis 3).

2. Material and methods

2.1. Participants

Sixty-two healthy university students (39 female) between 18 and 33 years ($M = 23.50$, $SD = 3.53$) participated in this study. Most participants were students at the faculty of natural sciences enrolled in psychology (68%) who received course credit for their participation. The remaining participants were not compensated. The study was approved by the local ethics committee. We set a predefined time period for data collection and did not a priori specify a sample size. Post-hoc power analyses ($\alpha = 0.05$) showed that statistical power was high for large correlations (99% for $r = 0.50$), moderate for medium correlations (68% for $r = 0.30$), and small for small correlations (12% for $r = 0.10$).

2.2. Procedure

The study was conducted in a computer laboratory with a maximum of six participants. The interference task and the inhibitory control task were administered using *PsychoPy* (PY3–1.90.2). Participants completed the measures in the order presented below.

2.3. Materials and measures

2.3.1. Interference of naïve and mathematical theories

We adopted Shtulman and Valcarcel's (2012) task to assess conceptual interference in mathematics (approx. duration: 25 min). Participants decided via button press whether a presented statement was mathematically correct or incorrect. Each statement was presented until button press (or a maximum of 15 s), with a 1 s fixation point before item onset and a 1 s inter-trial interval. This task contained 196 newly developed statements from the mathematical subdomains fractions, algebra, units and geometry, probability, and basic concepts (see Tables S1 to S5 for all statements and the corresponding naïve or mathematical theories). The statements were developed by experts in mathematics education based on frequent misconceptions. We included mathematical subdomains and concepts that are central to mathematical education and in which misconceptions frequently occur (e.g., Padberg, 2005; Padberg & Wartha, 2017; Welder, 2012). Within each subdomain, 10 concepts were assessed through 4 items each,[1] with one statement being mathematically and naïvely true (e.g., 1/4 + 2/4 = 3/4), one being mathematically and naively false (e.g., 1/4 + 1/4 = 1/4), one being mathematically true and naively false (e.g., 1/10 + 1/10 = 1/5), and one being mathematically false and naively true (1/3 + 1/4 = 2/7). The first two conditions are *consistent*—the latter two *inconsistent* across naïve and mathematical theories. The subdomain's presentation order, concepts within subdomains, and statements within concepts were randomized. We computed two indices of the interference strength:

$$\text{interference}_{\text{accuracy}} = \text{accuracy}_{\text{consistent trials}} - \text{accuracy}_{\text{inconsistent trials}} \quad (1)$$

$$\text{interference}_{\text{reaction time (RT)}} = \text{RT}_{\text{inconsistent trials}} - \text{RT}_{\text{consistent trials}} \quad (2)$$

$\text{Interference}_{\text{RT}}$ was calculated based on correct responses only.

2.3.2. General inhibitory control

We assessed general inhibitory control with the picture-word task (approx. duration: 10 min; Heidekum, Grabner, De Smedt, De Visscher, & Vogel, 2019). This task was used because it assesses semantic prepotent response inhibition, which has been hypothesized to play a role in suppressing naïve theories (Mason & Zaccoletti, 2020). Participants were presented picture-word pairs and had to decide via button press whether the word meaning matches the concept displayed in the picture. Three conditions were administered: (a) correct: matching picture-word pair, (b) related lure: picture and incorrect word from the same semantic category (e.g., dog – cat), (c) unrelated lure: picture and incorrect word from different semantic categories (e.g., cup – cat). The test comprised 32 concepts from six semantic categories (animals, insects, plants, fruits, tools, and clothing), which were presented four times in different combinations (50% correct) divided into 4 blocks with 32 items each. After a fixation point (0.5 s), the picture-word pair was shown for 1.5 s, followed by an inter-trial-interval between 1 s and 4 s. There is a stronger conceptual overlap between the presented picture and the incorrect word in trials with related lures compared to trials with unrelated lures. Thus, stronger inhibitory control is required in these trials. We computed two indices of inhibitory control in which higher scores indicate higher inhibitory control:

$$\text{inhibitory control}_{\text{accuracy}} = \text{accuracy}_{\text{related lures}} - \text{accuracy}_{\text{unrelated lures}} \quad (3)$$

$$\text{inhibitory control}_{\text{RT}} = \text{RT}_{\text{unrelated lures}} - \text{RT}_{\text{related lures}} \quad (4)$$

2.3.3. Mathematical achievement

We included three indicators of mathematical achievement.

2.3.3.1. Mathematical competence. Participants completed a short version of the German Mathematics Test for Personnel Selection (time limit: 15 min; Jasper & Wagener, 2011). This paper-pencil test covers a wide range of mathematical problems, including fractions, conversion of units, exponentiation, division with decimals, algebra, geometry, roots, and logarithm.

2.3.3.2. Arithmetic fluency. Participants completed a brief time-based paper-pencil test (time limit: 90s, Vogel et al., 2017) in which they solved simple (single-digit) additions, subtractions, and multiplications.

2.3.3.3. Mathematics grade. Participants self-reported the mathematics grade of their university entrance diploma (5 grade levels ranging from 1 "very good"to 5 "not sufficient"). Lower grade scores indicate higher mathematical achievement.

[1] In the fractions subdomain, one statement was erroneous so that only 9 concepts with 4 statements each were included in the analysis.

3. Results

Table S6 displays the descriptive statistics. Table S7 displays the bivariate correlation matrix. Tables S8 and S9 display the intercorrelations of interference strength across mathematical subdomains (all $r \leq 0.33$). Hypotheses 1 states that naïve and scientific theories interfere in mathematical subdomains. To test this assumption, we conducted repeated-measures analyses of variance (ANOVAs) with mathematical domains (fractions, algebra, geometry, probability, basic concepts) and consistency (consistent, inconsistent) as independent variables separately for accuracy and RT as dependent variables. For accuracy, we found a significant main effect of consistency indicating that consistent statements were solved more accurately than inconsistent statements, $F(1,61) = 187.89$, $p < .001$, $\eta_p^2 = 0.76$. We also found a main effect of domain indicating general performance differences between domains, $F(2.86,174.42) = 38.86$, $p < .001$, $\eta_p^2 = 0.39$. The interaction of domain and consistency was significant, indicating that the interference strength differed between domains, $F(3.35,204.58) = 22.99$, $p < .001$, $\eta_p^2 = 0.27$, see Fig. 1a. Follow-up *t*-tests showed that participants were more accurate in consistent compared to inconsistent statements in all subdomains ($t(61) \geq 3.82$, $p < .001$ for all comparisons) with Cohen's *d* ranging from 0.49 (for probability) to 1.36 (for units and geometry; see Table S10).

A similar picture emerged for RT. Consistent statements were solved faster than inconsistent statements, $F(1,61) = 75.85$, $p < .001$, $\eta_p^2 = 0.55$. There were also significant performance differences between domains, $F(4,244) = 25.82$, $p < .001$, $\eta_p^2 = 0.30$. A significant interaction between consistency and domain indicated that the interference strength differs between domains, $F(4,244) = 12.08$, $p < .001$, $\eta_p^2 = 0.16$, see Fig. 1b. Follow-up *t*-tests showed that participants were faster in consistent compared to inconsistent statements in four mathematical subdomains ($t(61) \geq |3.00|$, $p \leq .004$ for all comparisons) with Cohen's *d* ranging from −0.38 (for basic concepts) to −0.88 (for fractions; see Table S10). Only in probability, participants were equally fast in inconsistent and consistent trials, $t(61) = 0.14$, $p = .887$, $d = 0.02$. In sum, the results largely support Hypothesis 1.

Hypothesis 2 states that mathematical achievement (Hypothesis 2a) and inhibitory control (Hypothesis 2b) predict interference strength in mathematics. As presented in Table 1, bivariate regression analyses revealed that all three mathematical achievement measures predicted $\text{Interference}_{\text{accuracy}}$, but not $\text{interference}_{\text{RT}}$. The regression coefficient for mathematical competence ($\beta = -0.69$) was significantly larger than those for arithmetic fluency ($\beta = -0.29$; $z = -3.37$, $p < .001$) and math grade ($\beta = 0.40$; $z = -2.46$ $p = .014$). Inhibitory control did not predict interference strength ($ps \geq 0.053$). These findings partially support Hypothesis 2a and do not support Hypothesis 2b. A supplementary multiple regression showed that only mathematical competence predicted $\text{interference}_{\text{accuracy}}$ ($\beta = -0.64$, $p < .001$) after controlling for the other mathematical achievement measures (see Table S13).

Hypothesis 3 posits that mathematical achievement and interference strength are related even after controlling for general inhibitory control. We tested this expectation in hierarchical regression analyses. The three mathematical achievement measures were uniquely related to $\text{interference}_{\text{accuracy}}$ ($\beta = -0.69$, $p < .001$, for mathematical competence; $\beta = -0.28$, $p = .027$, for arithmetic fluency, $\beta = 0.35$, $p = .009$, for the mathematics grade), but not to $\text{interference}_{\text{RT}}$ ($|\beta| = 0.16$ to 0.18, $ps \geq 0.206$) after controlling for inhibitory control (see Tables S11 and S12). Thus, the analyses partly support Hypothesis 3.

4. Discussion

This study showed that naïve and scientific theories interfere in a broad range of mathematical subdomains. In line with Hypothesis 1, we observed a higher accuracy or faster responses for consistent compared to inconsistent statements in all mathematical subdomains under investigation. Our study added to the conceptual clarification of interference observed in the Shtulman and Valcarcel (2012) task. In previous

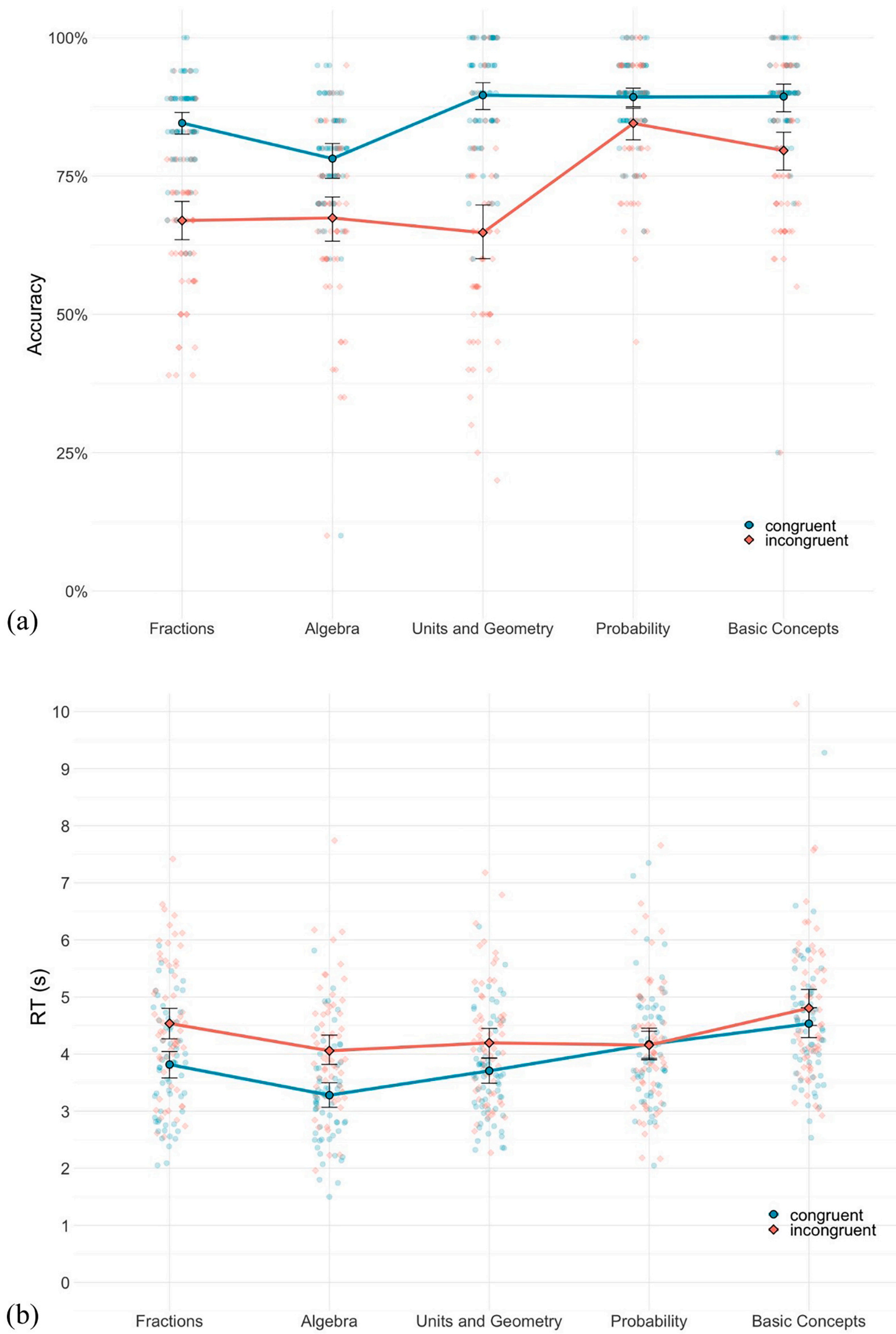

Fig. 1. Accuracy (a) and reaction times (RT) (b) across mathematical subdomains and trial types. Error bars indicate 95% confidence intervals. Light dots (circles and squares) represent the individual values in both conditions.

studies applying this task, a large proportion of the presented statements referred to everyday perceptual experiences. As pointed out by Potvin, Masson, Lafortune, and Cyr (2015), it is not entirely clear whether interference for such statements reflects a conflict of naïve *theories* or *everyday experiences* with scientific theories: Statements inconsistent with everyday perceptual experiences may produce interference simply because these statements are less familiar than statements consistent with everyday perceptual experiences. The mathematical statements administered in this study, in contrast, were more formal and abstract. Thus, the observed effects demonstrate conflict between naïve and scientific theories independent of perceptual experience.

This study also revealed characteristics that account for individual differences in interference strength. Our results, partially supporting Hypothesis 2a, showed that mathematical achievement indicators are associated with interference in accuracy. This finding suggests that individuals with higher mathematical achievement can better inhibit

Table 1
Standardized regression coefficients for mathematical achievement and inhibitory control as predictors of interference strength.

Interference in mathematics	Mathematical achievement			Inhibitory control	
	Mathematical competence	Arithmetic fluency	Math grade	Accuracy	RT
Accuracy	−0.69***	−0.29*	0.40**	−0.25	−0.04
RT	−0.12	−0.15	0.15	0.06	0.02

Note. RT = reaction time. $N = 62$ for all variables except math grade ($N = 58$). A lower math grade indicates higher mathematical achievement.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

naïve theories in mathematics, which is, generally, in line with the results of previous comparisons between achievement groups (Goldberg & Thompson-Schill, 2009; Shtulman & Harrington, 2016). The observation that 51.41% of the inter-individual differences in interference strength in accuracy can be accounted for by mathematical achievement measures (see Table S13) highlights the relevance of conceptual interference for mathematical achievement. The low magnitude of correlations between interference scores across mathematical subdomains further supports the domain-specificity of conceptual interference.

In contrast to mathematical achievement, general inhibitory control did not predict interference strength (Hypothesis 2b not supported). Thus, conceptual interference in mathematics appears relatively unrelated to the general ability to suppress irrelevant semantic information (also see Kelemen et al., 2013), which is in line with the domain-specificity of knowledge and learning more broadly. Yet, in previous work, general inhibitory control has been related to performance in tasks that require rejecting naïve theories (Vosniadou et al., 2018). Thus, one may speculate whether fundamentally different components of inhibitory control were assessed in this study. An alternative explanation could be that inhibitory control is less relevant for rejecting naïve mathematical theories that are usually not reinforced by perceptual experiences. This study also showed that mathematical achievement predicts lower interference in accuracy but not interference in reaction times beyond the contribution of inhibitory control (partially supporting Hypothesis 3). Thus, domain-specific knowledge-based reasoning processes, but not general inhibitory control play a significant role in the interference between naïve and scientific theories.

This study has some limitations. Due to our cross-sectional design, it is unclear whether stronger interference leads to lower mathematical achievement or vice versa. Future longitudinal studies are needed to reveal temporal ordering. Additionally, we assessed semantic prepotent response inhibition using a relatively novel task. Although semantic prepotent response inhibition appears central to the Shtulman and Valcarcel (2012) task, further research is needed to test relations with other inhibitory control tasks and processes. Also, future research using larger samples is needed to detect smaller correlations of interference strength with relevant outcomes. Finally, differences in interference strength between mathematical subdomains may be due to differences in the stimulus material (e.g., statement length). For example, statements in the probability subdomain were substantially longer than statements in other subdomains, which may explain why no interference in reaction times occured.

Regarding practical implications, this study showed that formal instruction rarely succeeds in entirely eradicating naïve theories in mathematics. Thus, teachers should carefully check whether naïve theories reoccur and place particular emphasis on counterintuitive mathematical operations, even long after the respective scientifically correct mathematical theories have been learned.

Declarations of interest

None.

Acknowledgements

We thank Anke Geißler as well as Christian Dies for their support in developing the items used in this study and Dennis Wambacher as well as Clemens Brunner for their technical support in the study preparation and data analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2021.104789.

References

Babai, R., Zilber, H., Stavy, R., & Tirosh, D. (2010). The effect of intervention on accuracy of students' responses and reaction times to geometry problems. *International Journal of Science and Mathematics Education, 8*, 185–201. https://doi.org/10.1007/s10763-009-9169-8.

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction, 37*, 39–49. https://doi.org/10.1016/j.learninstruc.2014.07.002.

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., … Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One, 8*, Article e67374. https://doi.org/10.1371/journal.pone.0067374.

Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental "roots" in mature biological knowledge. *Psychological Science, 20*, 480–487. https://doi.org/10.1111/j.1467-9280.2009.02320.x.

Heidekum, A. E., Grabner, R. H., De Smedt, B., De Visscher, A., & Vogel, S. E. (2019). Interference during the retrieval of arithmetic and lexico-semantic knowledge modulates similar brain regions: Evidence from functional magnetic resonance imaging (fMRI). *Cortex, 120*, 375–393. https://doi.org/10.1016/j.cortex.2019.06.007.

Jasper, F., & Wagener, D. (2011). *Mathematiktest für die Personalauswahl [Mathematics test for personnel selection]*. Hogrefe.

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition, 111*, 138–143. https://doi.org/10.1016/j.cognition.2009.01.001.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology. General, 142*, 1074–1083. https://doi.org/10.1037/a0030399.

Mason, L., & Zaccoletti, S. (2020). Inhibition and conceptual learning in science: A review of studies. *Educational Psychology Review, 33*, 181–212. https://doi.org/10.1007/s10648-020-09529-x.

Padberg, F. (2005). *Didaktik der Arithmetik [Didactics of arithmetic]* (3rd ed.) (Spektrum).

Padberg, F., & Wartha, S. (2017). *Didaktik der Bruchrechnung [Didactics of fractions]*. Springer.

Potvin, P. (2013). Proposition for improving the classical models of conceptual change based on neuroeducational evidence: Conceptual prevalence. *Neuroeducation, 1*, 16–43. https://doi.org/10.24046/neuroed.20130201.16.

Potvin, P., Malenfant-Robichaud, G., Cormier, C., & Masson, S. (2020). Coexistence of misconceptions and scientific conceptions in chemistry professors: A mental chronometry and fMRI study. *Frontiers in Education*. https://doi.org/10.3389/feduc.2020.542458.

Potvin, P., Masson, S., Lafortune, S., & Cyr, G. (2015). Persistence of the intuitive conception that heavier objects sink more: A reaction time study with different levels of interference. *International Journal of Science and Mathematics Education, 13*, 21–43. https://doi.org/10.1007/s10763-014-9520-6.

Shtulman, A., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science, 8*, 118–137. https://doi.org/10.1111/tops.12174.

Shtulman, A., & Legare, C. H. (2020). Competing explanations of competing explanations: Accounting for conflict between scientific and folk explanations. *Topics in Cognitive Science, 12*, 1337–1362. https://doi.org/10.1111/tops.12483.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124*, 209–215. https://doi.org/10.1016/j.cognition.2012.04.005.

Stavy, R., & Babai, R. (2010). Overcoming intuitive interference in mathematics: Insights from behavioral, brain imaging and intervention studies. *ZDM – Mathematics Education, 42*, 621–633. https://doi.org/10.1007/s11858-010-0251-z.

Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2013). *Brief report. Educated adults are still affected by intuitions about the effect of arithmetical operations: Evidence from a reaction-time study. Educational studies in mathematics. 82* pp. 323–330). https://doi.org/10.1007/s10649-012-9432-8.

Varma, S., & Karl, S. R. (2013). Understanding decimal proportions: Discrete representations, parallel access, and privileged processing of zero. *Cognitive Psychology, 66*, 283–301. https://doi.org/10.1016/j.cogpsych.2013.01.002.
Vogel, S. E., Haigh, T., Sommerauer, G., Spindler, M., Brunner, C., Lyons, I. M., & Grabner, R. H. (2017). Processing the order of symbolic numbers: A reliable and unique predictor of arithmetic fluency. *Journal of Numerical Cognition, 3*(2), 288–308. https://doi.org/10.5964/jnc.v3i2.55.
Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45–69. https://doi.org/10.1016/0959-4752(94)90018-3.
Vosniadou, S., Pnevmatikos, D., Makris, N., Lepenioti, D., Eikospentaki, K., Chountala, A., & Kyrianakis, G. (2018). The recruitment of shifting and inhibition in on-line science and mathematics tasks. *Cognitive Science, 42*, 1860–1886. https://doi.org/10.1111/cogs.12624.
Welder, R. M. (2012). Improving algebra preparation: Implications from research on student misconceptions and difficulties: Improving algebra preparation: Improving algebra preparation. *School Science and Mathematics, 11*, 255–264. https://doi.org/10.1111/j.1949-8594.2012.00136.x.