Neural representations of transitive relations predict current and future math calculation skills in children

Flora Schwartz [a,*], Justine Epinat-Duclos [b], Jessica Léone [b], Alice Poisson [c], Jérôme Prado [b,**]

[a] *Department of Psychiatry and Behavioral Sciences, Stanford University, USA*
[b] *Lyon Neuroscience Research Center (CRNL), Experiential Neuroscience and Mental Training Team, INSERM U1028 - CNRS UMR5292, University of Lyon, France*
[c] *GénoPsy, Reference Center for Rare Diseases with Psychiatric Symptoms, Centre Hospitalier le Vinatier, France*

ARTICLE INFO

ABSTRACT

A large body of evidence suggests that math learning in children is built upon innate mechanisms for representing numerical quantities in the intraparietal sulcus (IPS). Learning math, however, is about more than processing quantitative information. It is also about understanding relations between quantities and making inferences based on these relations. Consistent with this idea, recent behavioral studies suggest that the ability to process transitive relations (A > B, B > C, therefore A > C) may contribute to math skills in children. Here we used fMRI coupled with a longitudinal design to determine whether the neural processing of transitive relations in children could predict their current and future math skills. At baseline (T_1), children (n = 31) processed transitive relations in an MRI scanner. Math skills were measured at T_1 and again 1.5 years later (T_2). Using a machine learning approach with cross-validation, we found that activity associated with the representation of transitive relations in the IPS predicted math calculation skills at both T_1 and T_2. Our study highlights the potential of neurobiological measures of transitive reasoning for forecasting math skills in children, providing additional evidence for a link between this type of reasoning and math learning.

1. Introduction

The ability to understand and manipulate logical relations has long been thought to be associated with the acquisition of math skills in children. For instance, the pioneering developmental psychologist Jean Piaget famously proposed that the development of math cognition relies on the emergence of logical skills from childhood to adolescence (Piaget, 1952). Educational policies have also often been influenced by the idea that learning math would contribute to the development of logical thinking in children, a claim whose premises can be found in the writings of philosophers such as Plato, Locke, and Bacon (Inglis and Attridge, 2017). To date, however, evidence that math learning is related to logical reasoning remains scarce, most likely because research has largely focused on investigating to what extent math development relies on evolutionarily old mechanisms for representing quantities (Feigenson et al., 2004).

Yet, mathematical development is about more than simply processing quantitative information. It is also about understanding relations between those quantities and making inferences based on these relations. For example, a type of logical relation that is prevalent in many math domains is that of transitivity. Transitivity is a property that arises from a set of items that can be ordered along a single continuum (Wright, 2001). A relation is said "transitive" when it allows reasoners to infer a relationship between two items (e.g., A > C) from two other overlapping pairs (e.g., A > B; B > C). The ability to recognize transitive relations and make associated inferences may contribute to the acquisition of many mathematical concepts. For example, transitive reasoning is fundamental to the acquisition of measurement skills in children (Inhelder and Piaget, 1958; Bryant and Kopytynska, 1976; Rabinowitz and Howe, 1994). It also allows for the processing of ordinal and categorical information, which in turn contributes to domains such as arithmetic, algebra and geometry (Bryant and Kopytynska, 1976; Rabinowitz and Howe, 1994; Wright, 2001).

Consistent with a role for transitive reasoning in math learning, a few studies have shown that the ability to understand transitive relations relate to math skills in children. For example, Morsanyi et al. (2013) found that 10-year-olds with math learning disability (i.e., dyscalculia) performed significantly worse than typically developing controls in a

---

transitive reasoning task. In contrast, children with high math ability performed significantly better than typically developing controls on that same task. This is in line with a prior report showing that math performance is generally related to transitive (as well as conditional) reasoning ability in elementary school children (Handley et al., 2004).

What neural mechanisms may underlie the relationship between transitive reasoning and math learning in children? Two lines of evidence suggest that this relationship may be mediated by brain mechanisms involved in relational processing in the posterior parietal cortex (Wendelken, 2015). First, neuroimaging studies indicate that the processing of transitive relations consistently activates the intraparietal sulcus (IPS) (for a review, see Prado et al., 2011a), a region that is also typically involved in tasks that require the manipulation of numbers (Nieder and Dehaene, 2009). Arguably, both transitive reasoning and numerical cognition require individuals to process ordinal information. Therefore, this common reliance on IPS mechanisms may stem from the fact that this region is central to ordinal processing more generally, in line with studies showing that the IPS is also activated when participants represent learned series such as letters, days or months (Fias et al., 2007).

However, it is also possible that transitive relations and numbers are encoded in similar regions of the IPS because they both rely on representations that are inherently spatial and for which the posterior parietal cortex is key (Prado et al., 2010a, 2010b). For example, numbers are often thought to be represented on a mental number line (Hubbard et al., 2005) in adults and children. This is notably suggested by the behavioral distance effect observed in number comparison tasks (i.e., reaction times decrease with the distance between numbers; Moyer and Landauer, 1967), indicating that numbers that are close (e.g., 4 vs 5) are more difficult to distinguish than numbers that are far (e.g., 2 vs 8). This phenomenon has also been reported in transitive reasoning (i.e., reaction times decrease with the distance between items in a transitive ordering task; Prado et al., 2010a, 2010b), suggesting that transitive items are also arranged on a mental representation of space (Alfred et al., 2018). Furthermore, activity in the same region of the IPS has been found to decrease with the distance between numbers (Pinel et al., 2001; Pinel et al., 2004; Mussolin et al., 2010) and transitive items (Prado et al., 2010a), suggesting that transitive relations and numbers engage common spatial representations in the IPS.

A second line of evidence suggesting that the relationship between transitive reasoning and math learning may be mediated by the posterior parietal cortex is provided by a recent neuroimaging study. In that study, we contrasted brain activity of typically developing children to brain activity of children with math learning difficulties during a transitive reasoning task. While transitive relations were associated with IPS activity in typically developing children, that was not the case in children with math learning difficulty (Schwartz et al., 2018). Children with math learning difficulty also showed significantly less activity in the IPS than typically-developing children during the transitive reasoning task. Thus, the extent to which IPS mechanisms process transitive relations may be related to levels of math competence in children.

It is important to note that our previous study only provides indirect correlational evidence that math abilities may be associated with reasoning-related activity in the IPS. However, it raises an intriguing possibility that may strengthen the claim that transitive reasoning plays an important role in math learning: current and future math skills in children might be predicted (at least to some extent) by the neural processing of transitive relations in the IPS. The goal of the present study was to formally test this hypothesis. FMRI activity of 31 typically-developing children from 9 to 13 was measured while they passively listened to transitive (e.g., A > B, B > C) and non-transitive (e.g., A > B, C > D) relations that were embedded in a coherent story that was designed to be as interactive (it took the form of a "choose your own adventure story") and as engaging as possible (Schwartz et al., 2018). This task allowed us to measure activity associated with transitive relation in a relatively ecological context (i.e., discourse

comprehension). Arithmetic and math problem-solving abilities were measured for each child at the time of the fMRI session (T_1), as well as 1.5 years later (T_2). The predictive power of the neural representations of transitive relations in the IPS on current and future math skills was assessed using a machine learning approach with cross-validation (Gabrieli et al., 2015). Specifically, we evaluated whether multivariate patterns of IPS activity associated with the processing of transitive relations (as compared to non-transitive relations) at T_1 could accurately predict math scores of children at both T_1 and T_2.

2. Material and methods

2.1. Sample size justification

To our knowledge, this is the first study to examine the link between individual differences in math skills and the neural processing of transitive relations in typically developing children. For this reason, and because decoding accuracy in multivariate analyses may not reflect effect sizes (Hebart and Baker, 2018), it is difficult to determine the optimal sample size for the multivariate analyses. However, we note that previous studies that have used multivariate analyses to predict math skills in children have used sample sizes ranging from n = 20 (Evans et al., 2015) to n = 28 (Qin et al., 2015). Power calculations (performed with G*Power, https://stats.idre.ucla.edu/other/gpower/) indicate that such sample sizes would provide 80% power to detect a univariate brain-behavior correlation in the r = 0.5 to r = 0.6 range (with α = 0.05). Such a range is typically the one observed in most studies measuring univariate brain-behavior correlations (providing that regions in which correlations are observed are not selected from circular analyses; see Fig. 5 in Vul et al., 2009). Therefore, we aimed to analyze a sample size of around 28 participants in the present study (similar to Qin et al. (2015)). Because we expected to discard about 20% of participants due to technical, performance, or movement issues, we aimed to recruit around 35 participants.

2.2. Participants

Thirty-eight right-handed children from 8 to 13 were recruited using advertisements in schools, newspapers and social media. Seven of these participants were excluded from the analyses because of excessive head motion in at least 3 of the 4 fMRI runs (n = 4), presence of a specific language impairment (n = 1), and lack of behavioral data (n = 2). Thus, 31 children were included in the analyses. Eighteen of these participants were already analyzed in Schwartz et al. (2018), in which only univariate analyses were used at T_1 to compare brain activity associated with transitive reasoning between children with and without math learning difficulty. The mean age of the final sample was 10.97 years ($SD = 1.37$) at T_1 and 12.57 years ($SD = 1.39$) at T_2. All children were native French speakers and had no diagnosis of mental retardation or high intellectual potential. They also had no hearing deficit, no MRI contraindications and no history of neurological and psychiatric disorder. Parents gave their written informed consent and children gave their assent to participate in the experiment, which was approved by the local ethics committee (CPP Lyon Sud-Est II). Families were paid 80 euros for their participation.

2.3. Testing

To ensure that participants had normal cognitive functioning and to provide a means of measuring math skill, children completed standardized tests at the time of scanning (T_1) and after a follow-up period of between 1.45 and 1.79 years (T_2) (mean = 1.60 years, $SD = 0.07$). First, the Nouvelle Echelle Metrique de l'Intelligence (NEMI-2) (Cognet, 2006) was used to obtain a measure of verbal intelligence (estimated using subtests of general knowledge, vocabulary and comparisons) and matrix reasoning at T_1. Second, math skills were measured at both T_1

and T_2 using a French version of the Woodcock-Johnson Test of achievement (WJ III; Woodcock et al., 2001). To provide a comprehensive measure of math abilities in each individual, 3 subtests were used: Calculation, Math fluency, and Applied problems. The Calculation subtest is an untimed test in which participants perform math computations of increasing difficulty. The test begins with single-digit addition, subtraction and multiplication problems and progressively moves to double-digit problems (including division) and algebra, as well as logarithmic and trigonometric operations. Test administration is stopped after six consecutive errors or when the last item is reached (raw scores range from 0 to 45). In the Math fluency subtest, participants have 3 min to solve as many single-digit addition, subtraction and multiplication problems (raw scores range from 0 to 160). The Applied problems subtest measures the ability to analyze and solve math problems. While early items require application of basic numerical concepts (e.g., counting, performing simple addition and subtraction, reading clocks and coin values), most items require children to understand and analyze word problems. The test is untimed and testing stops after 6 consecutive errors or when the last item is reached (raw scores range from 0 to 63). Standardized IQ scores at T_1 and raw math scores at T_1 and T_2 are shown in Table 1.

2.4. Task

Participants completed the task described in Schwartz et al. (2018). In the scanner, they listened to 4 stories that each included a series of 12 short scenarios (see Table 2). Each scenario ended up with a question that the child had to answer. Similar to a "choose your own story" structure, children were told to pay attention to every scenario because their responses to questions were critical for making progress in the story. Scenarios varied with respect to the type of relations involved. Twenty-four scenarios included linear-order relations such as "A is more than B", while 24 scenarios included set-inclusion relations such as "All As are Bs". In half (i.e., 12) of each type of scenario, the relations overlapped so that they could be integrated, and a conclusion could be inferred. For instance, the linear-order relations "White cows give more milk than black cows" and "Black cows give more milk than brown cows" in the bottom left cell of Table 2 should lead to the conclusion that "White cows give more milk than brown cows". Similarly, the set-inclusion relations "All old farms are made of stone" and "All farms that are made of stone are uphill" in the top left cell of Table 2 should lead to the conclusion that "All old farms are uphill". The question that followed scenarios with such transitive relations always tested whether children had inferred conclusions upon listening to the scenario (it was

Table 1
Demographic information and measures of verbal IQ, reasoning (matrices) abilities and math skills.

Variable	Mean (SD)
Demographic information	
Age at T_1 (in years)	10.97 (1.37)
Age at T_2 (in years)	12.57 (1.39)
Verbal IQ (NEMI-II)	
General knowledge[a]	105.85 (20.4)
Vocabulary[a]	106.75 (18.45)
Comparison[a]	112.15 (21.75)
Matrix reasoning (NEMI-II)	
Raven's matrices[a]	105.85 (20.85)
Math skills (WJ-III)	
Calculation (T_1)[b]	18.90 (3.82)
Math fluency at (T_1)[b]	51.16 (17.43)
Applied problems at (T_1)[b]	37.23 (5.33)
Calculation (T_2)[b]	22.03 (5.18)
Math fluency at (T_2)[b]	62.97 (16.36)
Applied problems at (T_2)[b]	40.03 (6.04)

Notes.
[a] , standardized scores (M = 100, SD = 15).
[b] , raw scores.

Table 2
Examples of each type of scenario (translated from French).

	Transitive relations	Non-transitive relations
Set-inclusion relations	**1.** S1. "You are going on vacation to the countryside." S2. "You are planning to stay in a farm for a few days." S3. "There are farms uphill and downhill." S4. "All old farms are made of stone." S5. "All farms that are made of stone are uphill." S6. "You have to find an old farm." *Q. Reasoning question: "Are you going uphill (response 1) or downhill (response 2)?"*	**2.** S1. "You are going uphill and you find the old farm. *(response 1)*"/"You are going downhill and the farmers pick you up" *(response 2)* S2. "The farmers invite you in." S3. "You need to bring your bag to your bedroom on the 2nd floor." S4. "All bedrooms with a red door are next to the chicken coop." S5. "All bedrooms with a green door are next to the barn." S6. "The farmers' house is very big." *Q. Memory question: "Are you taking your bag to the 3rd floor (response 1) or to the 2nd floor (response 2)?"*
Linear-order relations	**4.** S1. "You are taking the pastries out of the oven." *(response 1)*/"You let the pastries in the oven and they are overbaked" *(response 2)*. S2. "You would like some milk for your breakfast." S3. "You are going to milk cows with the farmer." S4. "White cows give more milk than black cows." S5. "Black cows give more milk than brown cows." S6. "You need to milk the cows giving the most milk." *Q. Reasoning question: "Are you milking the brown cows (response 1) or the white cows (response 2)?"*	**3.** S1. "You are going to the 2nd floor and bring your bag in." *(response 2)*/"You are going to the 3rd floor and the farmers tell you to go down to the 2nd floor." *(response 1)* S2. "The next morning, the farmer is baking pastries." S3. "The farmer is asking you to take them out of the oven now." S4. "The chocolate cake is baking faster than the apple pie." S5. "The strawberry pie is baking faster than the cheesecake." S6. "It is very hot in the kitchen." *Q. Memory question: "Are you taking the pastries out of the oven now (response 1) or later (response 2)?"*

Notes. Numbers 1 to 4 (bold) indicate the order of presentation within the experimental run. The labels S1 to S6 indicate sentence number for each scenario.

termed *Reasoning question*). The remaining scenarios included non-transitive relations, that is, relations that were not overlapping. Thus, no particular conclusion could be inferred upon listening to the scenario. For instance, the linear-order relations "The chocolate cake is baking faster than the apple pie" and "The strawberry pie is baking faster than the cheesecake" in the bottom right cell of Table 2 are non-transitive. Similarly, the set-inclusion relations "All bedrooms with a red door are on the side of the chicken coop" and "All bedrooms with a green door are on the side of the barn" in the top right cell of Table 2 are non-transitive. In those cases, the question that followed simply tested whether children remembered some information that was explicitly given in the scenario and was termed *Memory question* (see Table 2).

Overall word count was controlled across the 4 types of scenarios. On average, word count was 69.1 for scenarios with transitive set-inclusion relations, 69.7 for scenarios with non-transitive set-inclusion relations, 73.6 for scenarios with transitive linear-order relations and 72.4 for scenarios with non-transitive linear-order relations. Because word count was not normally distributed, non-parametric testing was used to assess differences as a function of type of scenarios. A Kruskal-Wallis ANOVA indicated that word count did not differ between types of scenarios ($H_{(3,48)} = 4.14$, $p = 0.25$).

The task was split into 4 runs that contained one story each. Each story contained 3 scenarios with transitive linear-order relations, 3

scenarios with non-transitive linear-order relations, 3 scenarios with transitive set-inclusion relations and 3 scenarios with non-transitive set-inclusion relations. Two wrap-up sentences concluded each story. Runs were randomized but scenarios were presented in a fixed order within a run. This was because scenarios were embedded in a coherent story. However, two scenarios of the same type were never following each other. Additionally, responses were counterbalanced across different variables. First, the order of the correct response was counterbalanced between the 4 types of stories across the whole experiment. Second, within a run, the order of the correct response was counterbalanced between scenarios with transitive and non-transitive relations and between scenarios involving linear-order relations and set-inclusion relations. Third, to prevent participants from developing expectations during the task and using heuristic strategies to respond to *Reasoning questions*, transitive relations were switched around in half of the problems. For example, linear-order relations could be presented in the order "A is faster than B and B is faster than C" in some scenarios and in the order "B is faster than C and A is faster than B" in other scenarios. Participants were instructed to press on a response button when they were ready to hear the next sentence. They were told to press on one of two other response buttons to choose between options in the question. Each story started with a red fixation cross at the center of the screen. The red cross turned orange after 6 s and green after 2 s. The green cross lasted 2 s and was immediately followed by the presentation of the first sentence of the first scenario. There was a 500 ms interval between each sentence in a scenario. A white fixation cross was displayed before and after each question for a random duration between 2 and 4 s. A fixation cross was also displayed for 20 s at the end of each story. See Fig. 1 for a timeline of one scenario.

2.5. Experimental procedure

After standardized tests were administered at T_1, children were familiarized with the fMRI environment in a mock scanner. They listened to a recording of the noises associated with all fMRI sequences. A motion tracker system (3D Guidance trak STAR, Ascension Technology Corporation) was used to measure head movements and provide on-line feedback to participants. Finally, children practiced the task in that mock scanner. Different stimuli were used in the practice and in the scanning sessions. Stimuli were generated using Presentation software (Neurobehavioral Systems, Albany, CA). Stories were spoken through

headphones sentence by sentence. During each scenario, a black-and-white picture illustrating the setting was displayed on a computer screen that was viewed by the participants through a mirror attached to the head coil. Behavioral responses were recorded using MR-compatible keypads placed in the left and right hands.

2.6. fMRI: data acquisition

Images were collected with a Siemens Prisma 3T MRI scanner (Siemens Healthcare, Erlangen, Germany) at the CERMEP Imagerie du vivant in Lyon, France. The BOLD signal was measured with a susceptibility weighted single-shot Echo-planar Imaging (EPI) sequence. Imaging parameters were as follows: TR = 2000 ms, TE = 24 ms, flip angle = 80°, matrix size = 128 × 120, field of view = 220 × 206 mm, voxel size = 1.72 × 1.72 mm, slice thickness = 3 mm (0.48 mm gap), number of slices = 32. A high-resolution T1-weighted whole-brain anatomical volume was also collected for each participant. Parameters were as follows: TR = 3500 ms, TE = 2.24 ms, flip angle = 8°, matrix size = 256 × 256, field of view = 224 × 224 mm, voxel size = 0.87 × 0.87 mm, slice thickness = 0.9 mm, number of slices = 192."

2.7. fMRI: data preprocessing

Images were analyzed with SPM12 (Welcome department of Cognitive Neurology, London, UK). The first 4 images of each run were discarded to allow for T1 equilibration effects. Functional images were corrected for slice acquisition delays and spatially realigned to the first image of the first run to correct for head movements. Realigned images were smoothed with a Gaussian filter (4 × 4 × 8 mm full-width at half maximum). As in our previous studies (Schwartz et al., 2017; Mathieu et al., 2018; Schwartz et al., 2018), volumes showing rapid scan-to-scan movement greater than 1.5 mm were identified using ArtRepair (Mazaika et al., 2009). Those were removed and replaced by the interpolation of the 2 nearest non-repaired volumes. Runs with more than 10% of repaired volumes were excluded from the analysis. Finally, functional images were normalized into the standard Montreal Neurological Institute (MNI) space.

2.8. fMRI: statistical modeling

Statistical modeling was performed in the context of the General
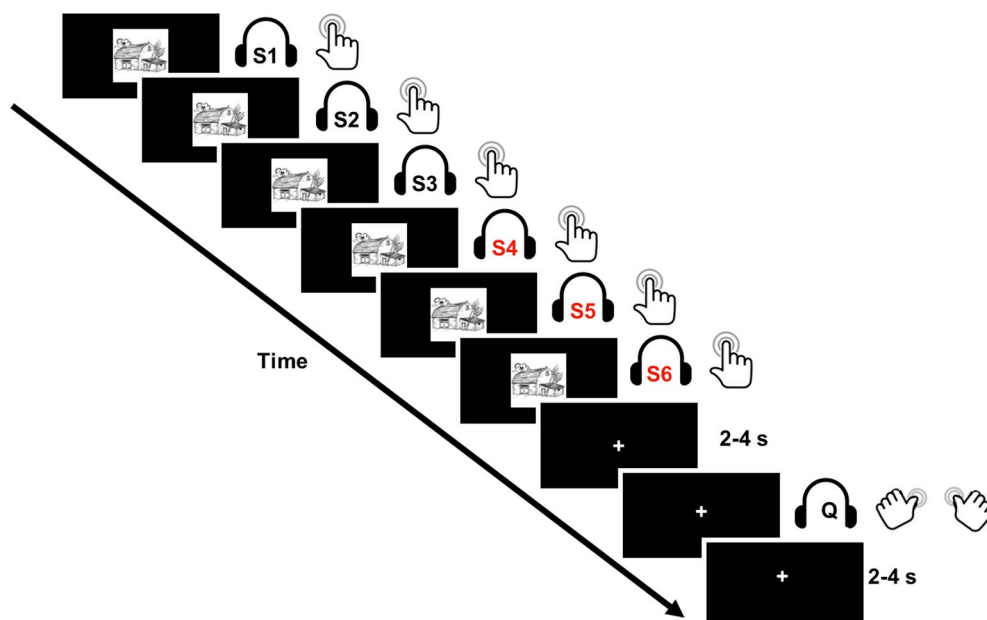
Fig. 1. Timeline for a sample scenario. Each of the 6 sentences (S) was spoken through headphones while a picture was displayed on the screen. The task was entirely self-paced. Participants pressed on a button to indicate that they were ready to listen to the next sentence, which was spoken after a 500 ms delay (not shown). The scenario ended with a question (Q), which was also spoken through headphones. This question was preceded and followed by a jittered interval ranging from 2 to 4 s. The sentences of interest considered in the analyses were sentences 4 to 6 (in red). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Linear Model, using the exact same procedure as in Schwartz et al. (2018). Sentences of interest were those containing transitive and non-transitive relations (i.e., sentences 4 and 5). These sentences were modeled as epochs with onsets time-locked to the presentation of sentence 4 and with offsets time-locked to the end of sentence 6 (i.e., the wrap-up sentence). Other sentences (as well as the question) were not explicitly modeled and were considered background noise. Linear-order and set-inclusion relations were modeled within the same regressors. All epochs were convolved with a canonical hemodynamic response function (HRF). The time series data were high-pass filtered (1/128 Hz), and serial correlations were corrected using an autoregressive AR(1) model.

2.9. fMRI: ROIs definition

Regions of interest (ROIs) were regions in which greater activity was observed for transitive than non-transitive relations. The definition of these regions involved the following 3 steps. First, for each subject, brain activity associated with non-transitive relations was subtracted from brain activity associated with transitive relations across the whole-brain. Second, these contrast images (one for each participant) were submitted to a second-level one-sample *t*-test. Third, the resulting t-map was thresholded using the non-parametric permutation-based Threshold-Free Cluster Enhancement (TFCE) method (Smith and Nichols, 2009), implemented in the TFCE Toolbox r164 (http://dbm.neuro.uni-jena.de/tfce/). As in Schwartz et al. (2018), we had an *a priori* hypothesis that transitive relations would be processed in the IPS. Therefore, clusters were considered significant at a Family-Wise Error- (FWE) corrected threshold of $p < 0.05$, either across the whole-brain or within an anatomical mask of the IPS (i.e., small volume correction) defined using the Anatomy Toolbox v2.2 (Eickhoff et al., 2005). Following Schwartz et al. (2018), the IPS mask consisted in voxels with at least 50% probability of belonging to one of the IPS subdivisions (hIP1, hIP2, and hIP3) as defined in the Anatomy Toolbox (Choi et al., 2006; Scheperjans et al., 2008a, 2008b).

2.10. fMRI: ROI analysis

Activity in the ROIs defined using the procedure above was analyzed using both univariate and multivariate methods. First, we calculated for each participant the average activity associated with the contrast of transitive versus non-transitive relations within an ROI by averaging the fMRI signal across all voxels within that ROI. This average activity was then correlated with Calculation, Math fluency, and Applied problems raw scores at T_1 and T_2 across participants.

Second, because such standard univariate regression analyses only show associations between two within-sample variables but do not have any predictive power (Gabrieli et al., 2015), we used a multivariate machine-learning approach with cross-validation to test whether brain activity associated with the contrast of transitive versus non-transitive relations could *predict* math scores of out-of-sample participants at both T1 and T2. This was done using the pipeline for pattern regression analyses provided by the Pattern Recognition for Neuroimaging Toolbox (PRoNTo) v2.0 implemented in Matlab (http://www.mlnl.cs.ucl.ac.uk/pronto/) (Schrouff et al., 2013). This pipeline involves a training phase and a test phase (Portugal et al., 2016). During the training phase, a pattern regression model is trained by associating labels with spatial patterns of activity from a "training set". During the test phase, the association between patterns and labels (learned in the training phase) is used to predict the label of novel (i.e., out-of-sample) patterns from a "test set" (Schrouff et al., 2013). In the present study, spatial patterns were voxel-by-voxel patterns of activity for the contrast of transitive versus non-transitive relations from each ROI. Labels were Calculation, Math fluency, and Applied problems raw scores at both T_1 and T_2. Following studies suggesting that Kernel Ridge Regression (KRR) (Shawe-Taylor and Cristianini, 2004) is computationally effective and may give better prediction accuracy than other linear methods (Chu et al.,

2011), we used KRR to predict math scores. Briefly, KRR estimates a relationship between samples by minimizing an error function that includes the sum of the squared differences between model predictions and regression targets as well as a model regularization term (Taylor et al., 2017). Partitioning between the training and the test sets was done using a leave-one-subject-out cross-validation method. This involves leaving one participant out for test and training the model on 30 (i.e., n - 1) participants. The procedure was then repeated 31 times, such that each participant was left out once.

Prediction accuracy of the model was measured using the Pearson's correlation coefficient (r) and the mean squared error (MSE) of the relationship between actual and predicted scores. Positive r values indicate that the model is able to predict the outcome, while negative r values (or r values close to 0) indicate that the model fails to predict the outcome. Statistical significance was calculated using permutation tests with 1000 iterations. That is, the value for each metric (i.e., r and MSE) was compared to the results of 1000 random permutations (representing the distribution of correlations corresponding to the null hypothesis). P values were calculated based on how many times the value obtained with random permutations was higher than the correlation value observed in the data (divided by 1000).

3. Results

3.1. ROIs definition

ROIs involved in transitive reasoning were identified by contrasting activity associated with transitive relations to activity associated with non-transitive relations. As shown in Table 3 and Fig. 2, this contrast revealed significant activity in two clusters of the left IPS and one cluster of the right IPS. These clusters defined the voxels that were used in the subsequent univariate and multivariate analyses (see **Methods**). Specifically, both left IPS clusters constituted the left IPS ROI (158 voxels) and the right IPS cluster constituted the right IPS ROI (154 voxels). No region outside of the IPS survived whole-brain correction for multiple comparisons.

3.2. Relationships between math scores, transitive reasoning performance, and univariate IPS activity

Transitive reasoning performance was measured using accuracy to reasoning questions during the fMRI task at T_1. Accuracy to *reasoning* questions ranged from 0.5 to 1 (average = 0.75, SD = 0.15). Accuracy to *memory* questions ranged from 0.61 to 1 (average = 0.82, SD = 0.12). Table 4 contains the correlations among math scores (at T1 and T2), accuracy to reasoning and memory questions, and univariate activity associated with the contrast of transitive versus non-transitive relations in the left and right IPS ROIs (all correlations partialled out differences in age and verbal IQ between participants). Although there was no relationship between Calculation scores and accuracy to Reasoning questions, there was a significant relationship between Calculation

Table 3
Clusters activated in the contrast of transitive versus non-transitive relations across all participants.

Anatomical location	Whole-brain P$_{FWE-corr}$	SVC P$_{FWE-corr}$	MNI coordinates			t-score	Cluster size (mm³)
			X	Y	Z		
Left hIP1/ hIP2	0.102	0.002	−46	−46	38	4.88	644
Left hIP1/ hIP3	0.226	0.007	−32	−56	41	3.57	1568
Right hIP1/ hIP2/hIP3	0.368	0.004	42	−48	41	4.03	2156

Notes. BA: Brodmann area; MNI: Montreal Neurological Institute; SVC: Small Volume Correction; FWE-corr: Family-wise error corrected.
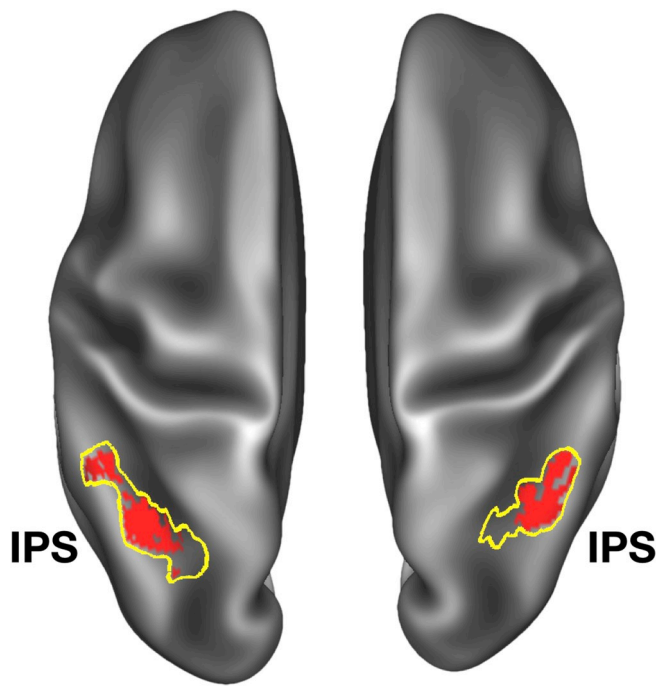
Fig. 2. Brain regions involved in transitive reasoning across all participants. Clusters in red are regions in which greater activity was observed for transitive than non-transitive relations. Yellow outlines delineate the IPS mask used for small volume correction (see Methods). Activations are overlaid on an inflated 3D rendering of the MNI-normalized anatomical brain (lateral views of the left and right hemispheres). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

scores and univariate activity in the left and right IPS, at both T1 and T2 (see Table 4). However, these brain-behavior correlations were specific to the Calculation subtest, as neither Math fluency nor Applied problem scores were correlated with univariate activity in the IPS (at T1 or T2). This was despite the fact that accuracy to reasoning questions did correlate with scores on these two subtests (but only at T_1). Thus, univariate activity associated with the contrast of transitive versus non-transitive relations was specifically related to math calculation skills at the time of scanning and 1.5 years later.

3.3. Predictive power of multivariate activity associated with transitive reasoning

The analyses above, however, only reflect in-sample associations between brain activity during the processing of transitive relations and math abilities. As such, they do not inform on whether brain activity during the processing of transitive relations may *predict* the outcome of

out-of-sample individuals (Gabrieli et al., 2015). To address this question, we used regression-based multivariate analyses with cross-validation. As in univariate analyses, age and measures of verbal IQ (general knowledge, comparison, and vocabulary) were included as confounding variables. These analyses tested whether patterns of IPS activity associated with the contrast of transitive versus non-transitive relations could accurately predict current math score in out-of-sample individuals, over and above differences in age and verbal IQ (see **Methods**). Accuracy of the machine learning algorithm was assessed by examining the relationship between actual and predicted math scores at T_1. As shown in Fig. 3 (left), there was no relationship between actual and predicted Calculation scores in the left IPS at T1 ($r = -0.07$, $p = 0.60$; MSE = 38.05, $p = 0.64$) or T2 ($r = -0.26$, $p = 0.87$; MSE = 92.77, $p = 0.89$). However, the relationship was reliable in the right IPS at both T1 ($r = 0.60$, $p = 0.002$; MSE = 11.87, $p = 0.006$) and T2 ($r = 0.39$, $p = 0.04$; MSE = 23.92, $p = 0.01$). Fig. 4 shows the consistency between the actual and the predicted Calculation score for each participant in the left and right IPS at both T_1 and T_2. In contrast to Calculation scores, there was no reliable relationship between actual and predicted Math fluency scores at either T_1 or T_2 (all $rs < 0.11$, all $ps > 0.32$; all MSEs > 466.17, all $ps > 0.30$). Similarly, there was no reliable relationship between actual and predicted Applied problem scores at either T_1 or T_2 (all $rs < 0.06$, all $ps > 0.39$; all MSEs > 54.31, all $ps > 0.36$). Consistencies between actual and predicted Math fluency and Applied problem scores are shown in Supplementary Fig. 1 and Supplementary Fig. 2. Therefore, patterns of right IPS activity associated with transitive relations predicted math calculation skills (but not arithmetic fluency or problem-solving skills) at the time of scanning and 1.5 years later.

4. Discussion

Over the past decades, developmental research on math learning has largely focused on exploring the role of mechanisms for representing numerical quantities (Ansari, 2008). Yet, acquiring math skills also involves understanding relations between quantities and drawing inferences based on these relations (Singley and Bunge, 2014). In keeping with this idea, recent behavioral studies suggest that understanding relations of transitivity may significantly contribute to math learning in children (Handley et al., 2004; Morsanyi et al., 2013; Morsanyi et al., 2017a, 2017b). Here we tested whether the brain representations of transitive relations may predict current and future math skills in children. Specifically, fMRI activity was measured while children listened to transitive and non-transitive relations embedded in a story context. Calculation, single-digit arithmetic, and problem-solving skills were also measured at the time of scanning (T_1) and 18 months later (T_2). A machine learning approach with cross validation was used to test whether activity associated with the contrast of transitive versus non-transitive relations could predict math skills at both T_1 and T_2.

First, we found that transitive relations were associated with greater activity than non-transitive relations in the IPS. This is consistent with

Table 4
Partial correlation matrix of math scores, performance on the transitive reasoning task, and univariate IPS activity, controlling for age and verbal IQ (General knowledge, Vocabulary, Comparison).

|	1		2		3		4		5		6		7		8		9		10
1- Calculation (T1)	–																		
2- Math fluency (T1)	0.50	**	–																
3- Applied problems (T1)	0.47	**	0.76	***	–														
4- Calculation (T2)	0.49	**	0.49	**	0.62	***	–												
5- Math fluency (T2)	0.31	†	0.74	***	0.72	***	0.68	***	–										
6- Applied problems (T2)	0.30		0.64	***	0.81	***	0.71	***	0.73	***	–								
7- Reasoning (accuracy)	0.27		0.38	*	0.42	*	0.12		0.17		0.14		–						
8- Memory (accuracy)	0.41	*	0.11		0.10		−0.01		0.02		−0.05		0.47	**	–				
9- Left IPS (beta)	0.41	*	−0.02		0.06		0.50	**	0.09		0.10		0.14		−0.01		–		
10- Right IPS (beta)	0.34	†	0.02		−0.09		0.41	*	0.04		−0.07		0.07		−0.12		0.84	***	–

Notes. †, p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.
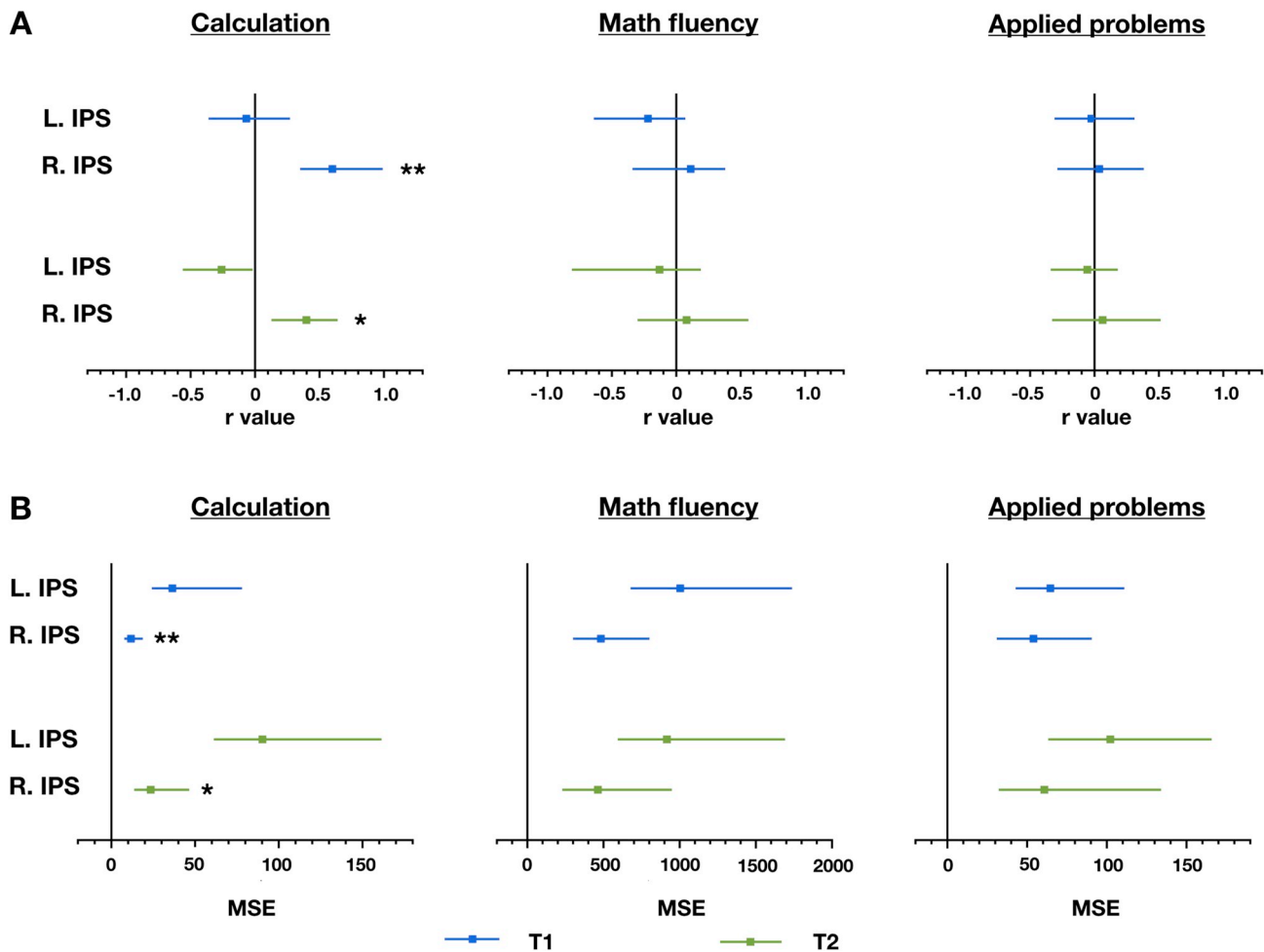
Fig. 3. Results of multivariate analyses. (A) Pearson's correlation coefficient (r) for the relationship between actual and predicted scores broken down by subtest (Calculation, Math fluency, and Applied problems) and time of testing (T_1, T_2). Error bars represent bootstrapped 95% confidence intervals. (B) Mean squared error (MSE) for the relationship between actual and predicted scores broken down by subtest (Calculation, Math fluency, and Applied problems) and time of testing (T_1, T_2). Error bars represent bootstrapped 95% confidence intervals. *, $p < 0.05$; **, $p < 0.01$; L., left; R., right.

several previous neuroimaging studies that have also found that transitive reasoning mainly relies on regions in and around the IPS (Goel and Dolan, 2001; Acuna et al., 2002; Knauff et al., 2002; Knauff et al., 2003; Heckers et al., 2004; Fangmeier et al., 2006; Hinton et al., 2010; Prado et al., 2010a, 2010b; Prado et al., 2013). For example, quantitative meta-analyses of the literature show that transitive reasoning consistently activates the left and right IPS across studies in adults (Prado et al., 2011a; Wertheim and Ragni, 2018). A recent study in children also points to the involvement of the parietal cortex in transitive reasoning (Mathieu et al., 2015). Thus, together with these previous studies, our results suggest that the IPS houses mechanisms that are important for representing and processing transitive relations. This may be because transitive relations are mapped onto spatial mental models (Goodwin and Johnson-Laird, 2005) that are encoded and represented in the IPS (Alfred et al., 2018).

Second, a major novel finding from the present study is that IPS activity associated with the contrast of transitive versus non-transitive relations was predictive of math skills at both T_1 and T_2. This predictive power, however, was limited to math calculation skills as measured by the Calculation subtest of the Woodcock-Johnson III battery. That is, it was not observed with the Math fluency subtest (reflecting arithmetic fluency) and the Applied problems subtest (reflecting broader problem-solving skills). Why would neural activity in the (right) IPS during the processing of transitive relations specifically predict current and future calculation skills? One possibility is that math calculation recruits

mechanisms supporting the processing of transitive relations in the IPS because numerical operations are inherently relational (i.e., they involve relating quantities to each other). Thus, activity related to transitive reasoning in the IPS may reflect the way this region processes relational information in children (Wendelken, 2015), which may in turn relate to their calculation skills. This possibility, however, is undermined by the fact that activity associated with transitive reasoning does not predict arithmetic fluency (as reflected by knowledge of single-digit problems in the Math fluency subtest). Even though they are largely thought to be retrieved from long-term memory rather than calculated (Campbell and Xue, 2001), single-digit arithmetic problems are also arguably relational and should therefore be related to IPS activity as it reflects relational processing.

Another possibility, which may be more likely, is that both transitive reasoning and calculation skills may rely on a set of IPS mechanisms involved in ordering information in working-memory. The role of the IPS in processing numerical and non-numerical ordinal information has been suggested by several studies (Marshuetz, 2005; Majerus et al., 2006). Individuals with math learning disability, who exhibit anatomical and functional impairment in the IPS, have also deficits when processing ordinal information (Attout and Majerus, 2015; De Visscher et al., 2015). A prominent cognitive theory of human reasoning (i.e., the Mental Models theory) claims that integrating transitive relations relies on the representation and manipulation of ordered mental models in working memory (Goodwin and Johnson-Laird, 2005). The ability to
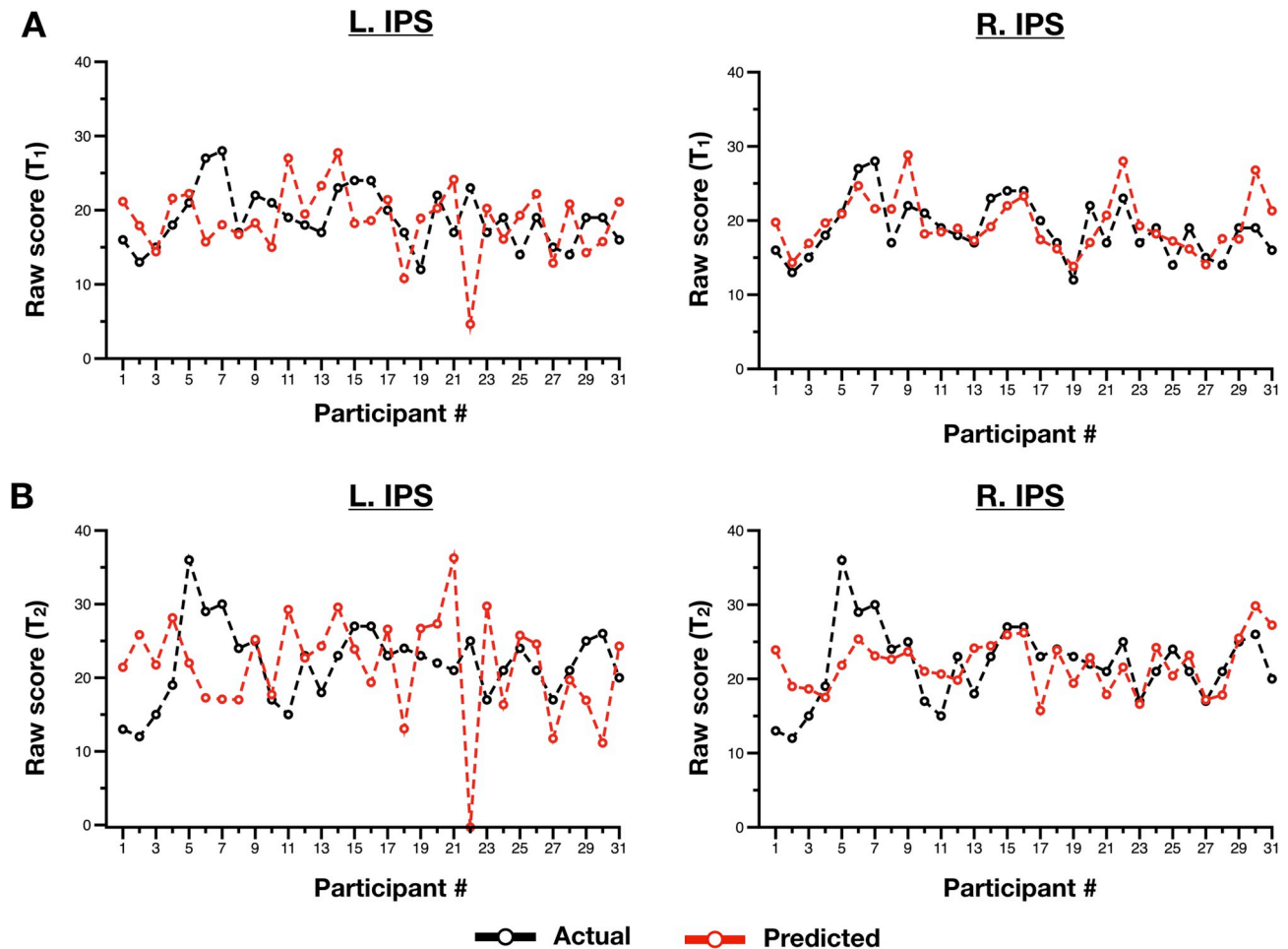
Fig. 4. Individual predictions of Calculation scores. (A) Line plots showing the consistency between the actual and the predicted Calculation score for each participant in the left and right IPS at T_1. (B) Line plots showing the consistency between the actual and the predicted Calculation score for each participant in the left and right IPS at T_2.

order numerical information is also fundamental to arithmetic skills (Knops and Willmes, 2014; Lyons et al., 2014). Therefore, it is possible that a common reliance on ordinal working memory in the IPS underlies the relationship between the processing of transitive relations and arithmetic learning. Interestingly, this possibility would account for the fact that IPS activity associated with transitive reasoning is not associated with Math fluency scores. As mentioned above, single-digit arithmetic problems are typically thought to be either directly retrieved from long-term memory or automatized through procedural knowledge by the end of elementary school (Thevenot et al., 2016). Thus, these problems are likely to rely to a lesser extent on working memory resources and involve brain mechanisms supporting the retrieval of verbal information in the left temporo-parietal cortex, rather than right IPS mechanisms typically associated with working memory and more effortful calculation strategies (Zamarian et al., 2009; Prado et al., 2011b).

More generally, it is also noteworthy that the predictive relationship between neural activity associated with transitive reasoning and calculation skills was observed in the right rather than the left IPS. Both the right and left IPS are involved in math cognition (Sokolowski et al., 2017; Arsalidou et al., 2018), but a dissociation may emerge during development. It has notably been suggested that the neural mechanisms supporting math skills may develop earlier in the right than in the left IPS, notably because they are involved in non-symbolic processing (Cantlon et al., 2006; Emerson and Cantlon, 2015; Vogel et al., 2015). Thus, although transitive reasoning has typically been associated with

bilateral IPS activity (Goel and Dolan, 2003; Knauff et al., 2003; Fangmeier et al., 2006; Prado et al., 2010a, 2010b; Mathieu et al., 2015; Schwartz et al., 2018), its interaction with math processing may be more readily observed in the right than left hemisphere in children.

It is interesting to note that we did not find any relationship between IPS activity associated with transitive reasoning and Applied problems scores. One possibility is that transitive reasoning does not critically support math problem-solving skills (as measured by the Applied problems subtest) in the developmental range considered. For instance, transitive reasoning may be particularly important earlier on when acquiring foundational math concepts such as measurement, serial-order and categorization (Bryant and Kopytynska, 1976; Rabinowitz and Howe, 1994; Wright, 2001). Although this suggests that transitive reasoning and math problem-solving may be observed in children younger than the age range considered here, accuracy to reasoning questions was related to score on the Applied problems subtest in our sample (at least at the time of testing). Thus, it is likely that brain activity during transitive reasoning also relates to Applied problems scores on some levels, even though this relation may be observed in brain regions that were not identified here. For instance, the Applied problems subtest involves math word problems that may rely on brain regions involved in verbal processing and executive control in the prefrontal cortex (Prabhakaran et al., 2001; Newman et al., 2011). Unfortunately, we were not able to localize frontal brain regions involved in transitive reasoning at the whole-brain level. This might be investigated in future studies.

Finally, we found that both univariate and multivariate activity associated with the neural processing of transitive relations in the IPS predicted calculation skills at T_1 and T_2, even though there was no behavioral relationship between transitive reasoning performance and calculation skills (over and above differences in age and verbal IQ). As pointed out by Evans et al. (2015), this may indicate that neurobiological measures are more sensitive than behavioral measurements in predicting learning outcomes. For example, activity in the right inferior frontal gyrus has been found to predict long-term reading gains in children with dyslexia even though no behavioral measures of reading ability could do so (Hoeft et al., 2011). In the domain of math cognition, previous neuroimaging studies have found that neither measures of executive functions nor measures of numerical abilities could accurately predict short-term (Supekar et al., 2013) or long-term (Evans et al., 2015) changes in math performance, whereas functional and anatomical brain measurements were predictive of learning. Therefore, our findings add to the growing evidence that brain measures have a unique potential for forecasting academic skills.

In summary, the present longitudinal study indicates that the neural processing of transitive relations may predict current and future calculation skills in children. These results add to growing evidence that math learning does not only rely on the ability to process numerical quantities, but also critically involves the ability to understand and process logical relations (Singley and Bunge, 2014). Because calculation skills were not related to transitive reasoning performance, our findings also suggest that fMRI measures may be more sensitive than behavioral measures in predicting academic skills. Overall, the present results call for future studies investigating the relationship between neurobiological measures of relational reasoning and different types of math learning in younger children as well as in children with math learning disability.

Declaration of competing interest

CRediT authorship contribution statement

Flora Schwartz: Methodology, Investigation, Data curation, Formal analysis, Writing - review & editing. **Justine Epinat-Duclos:** Investigation, Data curation. **Jessica Léone:** Investigation, Data curation. **Alice Poisson:** Supervision. **Jérôme Prado:** Funding acquisition, Conceptualization, Methodology, Supervision, Formal analysis, Writing - review & editing.

Acknowledgments:

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2020.107410.

References

Acuna, B., Eliassen, J., Donoghue, J., Sanes, J., 2002. Frontal and parietal lobe activation during transitive inference in humans. Cerebr. Cortex 12 (12), 1312–1321.

Alfred, K.L., Connolly, A.C., Kraemer, D.J.M., 2018. Putting the pieces together: generating a novel representational space through deductive reasoning. Neuroimage 183, 99–111.

Ansari, D., 2008. Effects of development and enculturation on number representation in the brain. Nat. Rev. Neurosci. 9 (4), 278–291.

Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., Pascual-Leone, J., 2018. Brain areas associated with numbers and calculations in children: meta-analyses of fMRI studies. Dev. Cog. Neurosci. 30, 239–250.

Attout, L., Majerus, S., 2015. Working memory deficits in developmental dyscalculia: the importance of serial order. Child Neuropsychol. 21 (4), 432–450.

Bryant, P., Kopytynska, H., 1976. Spontaneous measurement by young children. Nature 260, 773.

Campbell, J.I., Xue, Q., 2001. Cognitive arithmetic across cultures. J. Exp. Psychol. Gen. 130 (2), 299–315.

Cantlon, J.F., Brannon, E.M., Carter, E.J., Pelphrey, K.A., 2006. Functional imaging of numerical processing in adults and 4-y-Old children. PLoS Biol. 4 (5), e125.

Choi, H.-J., Zilles, K., Mohlberg, H., Schleicher, A., Fink, G.R., Armstrong, E., Amunts, K., 2006. Cytoarchitectonic identification and probabilistic mapping of two distinct areas within the anterior ventral bank of the human intraparietal sulcus. J. Comp. Neurol. 495 (1), 53–69.

Chu, C., Ni, Y., Tan, G., Saunders, C.J., Ashburner, J., 2011. Kernel regression for fMRI pattern prediction. Neuroimage 56 (2), 662–673.

Cognet, G., 2006. NEMI-2. ECPA, Paris, France.

De Visscher, A., Szmalec, A., Van Der Linden, L., Noel, M.P., 2015. Serial-order learning impairment and hypersensitivity-to-interference in dyscalculia. Cognition 144, 38–48.

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM Toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25 (4), 1325–1335.

Emerson, R.W., Cantlon, J.F., 2015. Continuity and change in children's longitudinal neural responses to numbers. Dev. Sci. 18 (2), 314–326.

Evans, T.M., Kochalka, J., Ngoon, T.J., Wu, S.S., Qin, S., Battista, C., Menon, V., 2015. Brain structural integrity and intrinsic functional connectivity forecast 6 year longitudinal growth in children's numerical abilities. J. Neurosci. 35 (33), 11743–11750.

Fangmeier, T., Knauff, M., Ruff, C.C., Sloutsky, V., 2006. FMRI evidence for a three-stage model of deductive reasoning. J. Cognit. Neurosci. 18 (3), 320–334.

Feigenson, L., Dehaene, S., Spelke, E., 2004. Core systems of number. Trends Cognit. Sci. 8 (7), 307–314.

Fias, W., Lammertyn, J., Caessens, B., Orban, G.A., 2007. Processing of abstract ordinal knowledge in the horizontal segment of the intraparietal sulcus. J. Neurosci. 27 (33), 8952–8956.

Gabrieli, J.D.E., Ghosh, S.S., Whitfield-Gabrieli, S., 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. Neuron 85 (1).

Goel, V., Dolan, R.J., 2001. Functional neuroanatomy of three-term relational reasoning. Neuropsychologia 39 (9), 901–909.

Goel, V., Dolan, R.J., 2003. Explaining modulation of reasoning by belief. Cognition 87 (1), B11–B22.

Goodwin, G.P., Johnson-Laird, P.N., 2005. Reasoning about relations. Psychol. Rev. 112 (2), 468–493.

Handley, S., Capon, A., Beveridge, M., Dennis, I., Evans, J.S.B.T., 2004. Working memory, inhibitory control, and the development of children's reasoning. Think. Reas. 10, 175–195.

Hebart, M.N., Baker, C.I., 2018. Deconstructing multivariate decoding for the study of brain function. Neuroimage 180, 4–18.

Heckers, S., Zalesak, M., Weiss, A., Ditman, T., Titone, D., 2004. Hippocampal activation during transitive inference in humans. Hippocampus 14 (2), 153–162.

Hinton, E.C., Dymond, S., von Hecker, U., Evans, C.J., 2010. Neural correlates of relational reasoning and the symbolic distance effect: involvement of parietal cortex. Neuroscience 168 (1), 138–148.

Hoeft, F., McCandliss, B.D., Black, J.M., Gantman, A., Zakerani, N., Hulme, C., Lyytinen, H., Whitfield-Gabrieli, S., Glover, G.H., Reiss, A.L., Gabrieli, J.D., 2011. Neural systems predicting long-term outcome in dyslexia. Proc. Natl. Acad. Sci. U. S. A. 108 (1), 361–366.

Hubbard, E.M., Piazza, M., Pinel, P., Dehaene, S., 2005. Interactions between number and space in parietal cortex. Nat. Rev. Neurosci. 6 (6), 435–448.

Inglis, M., Attridge, N., 2017. Does Mathematical Study Develop Logical Thinking?: Testing the Theory of Formal Discipline. World Scientific, Singapore.

Inhelder, B., Piaget, J., 1958. The Growth of Logical Thinking from Childhood to Adolescence: an Essay on the Construction of Formal Operational Structures. Basic Books, New York, NY.

Knauff, M., Fangmeier, T., Ruff, C.C., Johnson-Laird, P.N., 2003. Reasoning, models, and images: behavioral measures and cortical activity. J. Cognit. Neurosci. 15 (4), 559–573.

Knauff, M., Mulack, T., Kassubek, J., Salih, H.R., Greenlee, M.W., 2002. Spatial imagery in deductive reasoning: a functional MRI study. Brain Res. Cogn. Brain Res. 13 (2), 203–212.

Knops, A., Willmes, K., 2014. Numerical ordering and symbolic arithmetic share frontal and parietal circuits in the right hemisphere. Neuroimage 84, 786–795.

Lyons, I.M., Price, G.R., Vaessen, A., Blomert, L., Ansari, D., 2014. Numerical predictors of arithmetic success in grades 1-6. Dev. Sci. 17 (5), 714–726.

Majerus, S., Poncelet, M., Van Der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., Vandewalle, G., Collette, F., Maquet, P., 2006. The left intraparietal sulcus and verbal short-term memory: focus of attention or serial order? Neuroimage 32 (2), 880–891.

Marshuetz, C., 2005. Order information in working memory: an integrative review of evidence from brain and behavior. Psychol. Bull. 131 (3), 323–339.

Mathieu, R., Booth, J.R., Prado, J., 2015. Distributed neural representations of logical arguments in school-age children. Hum. Brain Mapp. 36 (3), 996–1009.

Mathieu, R., Epinat-Duclos, J., Leone, J., Fayol, M., Thevenot, C., Prado, J., 2018. Hippocampal spatial mechanisms relate to the development of arithmetic symbol processing in children. Dev. Cogn. Neurosci. 30, 324–332.

Mazaika, P., Hoeft, F., Glover, G., Reiss, A., 2009. Methods and software for fMRI analysis for clinical subjects. In: The Organization of Human Brain Mapping, 15th Annual Meeting. The Organization of Human Brain Mapping, San Francisco, CA.

Morsanyi, K., Devine, A., Nobes, A., Szücs, D., 2013. The link between logic, mathematics and imagination: evidence from children with developmental dyscalculia and mathematically gifted children. Dev. Sci. 16 (4), 542–553.

Morsanyi, K., Kahl, T., Rooney, R., 2017a. The link between math and logic in adolescence: the effect of argument form. In: Toplak, M.E., Hove, J. Weller (Eds.), Individual Differences in Judgment and Decision Making from a Developmental Context. Psychology Press, UK, Hove, pp. 166–185.

Morsanyi, K., McCormack, T., O'Mahony, E., 2017b. The link between deductive reasoning and mathematics. Think. Reas. (1–24).

Moyer, R.S., Landauer, T.K., 1967. Time required for judgements of numerical inequality. Nature 215 (5109), 1519.

Mussolin, C., Mejias, S., Noel, M.P., 2010. Symbolic and nonsymbolic number comparison in children with and without dyscalculia. Cognition 115 (1), 10–25.

Newman, S.D., Willoughby, G., Pruce, B., 2011. The effect of problem structure on problem-solving: an fMRI study of word versus number problems. Brain Res. 1410, 77–88.

Nieder, A., Dehaene, S., 2009. Representation of number in the brain. Annu. Rev. Neurosci. 32, 185–208.

Piaget, J., 1952. The Child's Conception of Number. Routledge & Kegan Paul, London, England.

Pinel, P., Dehaene, S., Riviere, D., LeBihan, D., 2001. Modulation of parietal activation by semantic distance in a number comparison task. Neuroimage 14 (5), 1013–1026.

Pinel, P., Piazza, M., Le Bihan, D., Dehaene, S., 2004. Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. Neuron 41 (6), 983–993.

Portugal, L.C., Rosa, M.J., Rao, A., Bebko, G., Bertocci, M.A., Hinze, A.K., Bonar, L., Almeida, J.R., Perlman, S.B., Versace, A., Schirda, C., 2016. Can emotional and behavioral dysregulation in youth be decoded from functional neuroimaging? PloS One 11 (1), e0117603.

Prabhakaran, V., Rypma, B., Gabrieli, J.D., 2001. Neural substrates of mathematical reasoning: a functional magnetic resonance imaging study of neocortical activation during performance of the necessary arithmetic operations test. Neuropsychology 15 (1), 115–127.

Prado, J., Chadha, A., Booth, J.R., 2011a. The brain network for deductive reasoning: a quantitative meta-analysis of 28 neuroimaging studies. J. Cognit. Neurosci. 23 (11), 3483–3497.

Prado, J., Mutreja, R., Booth, J.R., 2013. Fractionating the neural substrates of transitive reasoning: task-dependent contributions of spatial and verbal representations. Cerebr. Cortex 23 (3), 499–507.

Prado, J., Mutreja, R., Zhang, H., Mehta, R., Desroches, A.S., Minas, J.E., Booth, J.R., 2011b. Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. Hum. Brain Mapp. 32 (11), 1932–1947.

Prado, J., Noveck, I.A., Van Der Henst, J.-B., 2010a. Overlapping and distinct neural representations of numbers and verbal transitive series. Cerebr. Cortex 20 (3), 720–729.

Prado, J., Van Der Henst, J.-B., Noveck, I.A., 2010b. Recomposing a fragmented literature: how conditional and relational arguments engage different neural systems for deductive reasoning. Neuroimage 51 (3), 1213–1221.

Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D.C., Menon, V., 2015. Hippocampal-neocortical functional reorganization underlies children's cognitive development. Nat. Neurosci. 17 (9), 1263–1269.

Rabinowitz, F.M., Howe, M.L., 1994. Development of the middle concept. J. Exp. Child Psychol. 57, 418–449.

Scheperjans, F., Eickhoff, S.B., Homke, L., Mohlberg, H., Hermann, K., Amunts, K., Zilles, K., 2008a. Probabilistic maps, morphometry, and variability of cytoarchitectonic areas in the human superior parietal cortex. Cerebr. Cortex 18 (9), 2141–2157.

Scheperjans, F., Hermann, K., Eickhoff, S.B., Amunts, K., Schleicher, A., Zilles, K., 2008b. Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. Cerebr. Cortex 18 (4), 846–867.

Schrouff, J., Rosa, M.J., Rondina, J.M., Marquand, A.F., Chu, C., Ashburner, J., Phillips, C., Richiardi, J., Mourão-Miranda, J., 2013. PRoNTo: pattern recognition for neuroimaging Toolbox. Neuroinformatics 11 (3), 319–337.

Schwartz, F., Epinat-Duclos, J., Léone, J., Poisson, A., Prado, J., 2018. Impaired neural processing of transitive relations in children with math learning difficulty. Neuroimage: Clinical 20, 1255–1265.

Schwartz, F., Epinat-Duclos, J., Leone, J., Prado, J., 2017. The neural development of conditional reasoning in children: different mechanisms for assessing the logical validity and likelihood of conclusions. Neuroimage 163, 264–275.

Shawe-Taylor, J., Cristianini, N., 2004. Kernel Methods for Pattern Analysis. Cambridge University Press.

Singley, A.T.M., Bunge, S.A., 2014. Neurodevelopment of relational reasoning: implications for mathematical pedagogy. Trends Neurosci. Edu. 3 (2), 33–37.

Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44 (1), 83–98.

Sokolowski, H.M., Fias, W., Mousa, A., Ansari, D., 2017. Common and distinct brain regions in both parietal and frontal cortex support symbolic and nonsymbolic number processing in humans: a functional neuroimaging meta-analysis. Neuroimage 146, 376–394.

Supekar, K., Swigart, A.G., Tenison, C., Jolles, D.D., Rosenberg-Lee, M., Fuchs, L., Menon, V., 2013. Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proc. Natl. Acad. Sci. U. S. A. 110 (20), 8230–8235.

Taylor, J.A., Matthews, N., Michie, P.T., Rosa, M.J., Garrido, M.I., 2017. Auditory prediction errors as individual biomarkers of schizophrenia. Neuroimage: Clinical 15, 264–273.

Thevenot, C., Barrouillet, P., Castel, C., Uittenhove, K., 2016. Ten-year-old children strategies in mental addition: a counting model account. Cognition 146, 48–57.

Vogel, S.E., Goffin, C., Ansari, D., 2015. Developmental specialization of the left parietal cortex for the semantic representation of Arabic numerals: an fMR-adaptation study. Dev. Cogn. Neurosci. 12, 61–73.

Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Perspect. Psychol. Sci. 4, 274–290.

Wendelken, C., 2015. Meta-analysis: how does posterior parietal cortex contribute to reasoning? Front. Hum. Neurosci. 8 (1042).

Wertheim, J., Ragni, M., 2018. The neural correlates of relational reasoning: a meta-analysis of 47 functional magnetic resonance studies. J. Cognit. Neurosci. 30 (11), 1734–1748.

Woodcock, R.W., McGrew, K.S., Mather, N., 2001. Woodcock-Johnson III Tests of Achievement. The Riverside Publishing Company, Itasca, IL.

Wright, B.C., 2001. Reconceptualizing the transitive inference ability: a framework for existing and future research. Dev. Rev. 21 (4), 375–422.

Zamarian, L., Ischebeck, A., Delazer, M., 2009. Neuroscience of learning arithmetic-evidence from brain imaging studies. Neurosci. Biobehav. Rev. 33 (6), 909–925.