



# Developmental Dyscalculia and Automatic Magnitudes Processing: Investigating Interference Effects between Area and Perimeter

Hili Eidlin-Levy\* and Orly Rubinsten

Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Department of Learning Disabilities, University of Haifa, Haifa, Israel

## OPEN ACCESS

### Edited by:

Frank Domahs,  
Philipps University of Marburg,  
Germany

### Reviewed by:

Ursula Fischer,  
University of Regensburg, Germany  
Annemie Desoete,  
Ghent University, Belgium

### \*Correspondence:

Hili Eidlin-Levy  
hileidlin@gmail.com

### Specialty section:

This article was submitted to  
Developmental Psychology,  
a section of the journal  
Frontiers in Psychology

Received: 24 July 2017

Accepted: 04 December 2017

Published: 21 December 2017

### Citation:

Eidlin-Levy H and Rubinsten O  
(2017) Developmental Dyscalculia  
and Automatic Magnitudes  
Processing: Investigating Interference  
Effects between Area and Perimeter.  
*Front. Psychol.* 8:2206.  
doi: 10.3389/fpsyg.2017.02206

The relationship between numbers and other magnitudes has been extensively investigated in the scientific literature. Here, the objectives were to examine whether two continuous magnitudes, area and perimeter, are automatically processed and whether adults with developmental dyscalculia (DD) are deficient in their ability to automatically process one or both of these magnitudes. Fifty-seven students (30 with DD and 27 with typical development) performed a novel Stroop-like task requiring estimation of one aspect (area or perimeter) while ignoring the other. In order to track possible changes in automaticity due to practice, we measured performance after initial and continuous exposure to stimuli. Similar to previous findings, current results show a significant group  $\times$  congruency interaction, evident beyond exposure level or magnitude type. That is, the DD group systematically showed larger Stroop effects. However, analysis of each exposure period showed that during initial exposure to stimuli the DD group showed larger Stroop effects in the perimeter and not in the area task. In contrast, during continuous exposure to stimuli no triple interaction was evident. It is concluded that both magnitudes are automatically processed. Nevertheless, individuals with DD are deficient in inhibiting irrelevant magnitude information in general and, specifically, struggle to inhibit salient area information after initial exposure to a perimeter comparison task. Accordingly, the findings support the assumption that DD involves a deficiency in multiple cognitive components, which include domain-specific and domain-general cognitive functions.

**Keywords:** developmental dyscalculia, magnitude processing, geometric processing Stroop task, inhibition processing

## INTRODUCTION

In the past few years, the relationship between numbers and other magnitudes, such as space and time, has evoked a great deal of interest (Walsh, 2003; Dehaene and Brannon, 2010; Gebuis and Reynvoet, 2012a; Newcombe, 2014; Leibovich et al., 2016; McCaskey et al., 2017). Walsh (2003) was the first to suggest the existence of a common processing mechanism for time, space, and quantity, and he established a Theory of Magnitude (ATOM). Evolutionarily, such a mechanism enabled simultaneous processing of numerical, temporal, and spatial features of the world in order to produce adaptive reactions. Following Walsh's theory, a wealth of knowledge appeared

with regards to magnitude and numerical associations (Pinel et al., 2004; Dakin et al., 2011; Ren et al., 2011; Stoianov and Zorzi, 2012; Tibber et al., 2012; Leibovich et al., 2016; McCaskey et al., 2017). Specifically, recent research implies the existence of a magnitude estimation mechanism representing both numerical and other continuous magnitudes, including area, perimeter, length, volume, time, etc. (Lourenco and Longo, 2010; Dakin et al., 2011; Gebuis and Reynvoet, 2012b; Leibovich and Henik, 2013, 2014; Newcombe, 2014; Leibovich et al., 2016; Lourenco and Bonny, 2016).

Accumulating evidence shows that magnitude information is probably more perceptually salient than numerical information, and thus may have developed earlier (Feigenson et al., 2002; Gebuis and Reynvoet, 2012a; Odic et al., 2012). As an example, infant research shows that continuous magnitudes, such as the contour length (perimeter) or size (area) of figures, are more salient to infants than the number of items (Clearfield and Mix, 1999; Feigenson et al., 2002). Despite the fact that continuous magnitude processing seems to be as important as numerical processing for daily performance (Gebuis and Reynvoet, 2012b; Leibovich and Henik, 2013, 2014), comparisons between different types of continuous magnitudes (e.g., area and perimeter) have received much less scientific attention.

The current study aims to deepen the knowledge about continuous magnitude processing and has three main purposes. The first purpose is to investigate automatic processing (namely, processed spontaneously and with no need for monitoring; Tzelgov, 1997) of two continuous magnitudes, area and perimeter. The second purpose is to explore general interactions between numerical cognition and magnitude processing by inclusion of a specific clinical population – participants with deficient numerical abilities (developmental dyscalculia or DD). The third purpose is to investigate whether the DD group benefits from continuous exposure to stimuli as much as the control group.

## Area and Perimeter Processing

As mentioned, the study aims to investigate the automaticity of area and perimeter magnitudes. The automaticity of magnitude processing is investigated here via a Stroop-like task. Generally, in Stroop tasks it is typically found that participants cannot ignore irrelevant dimensions of the task (e.g., the physical size of a digit), which are processed involuntarily and interfere with the processing of the relevant dimension (e.g., the actual quantity that the digit represents; Henik and Tzelgov, 1982). Here, we adopted a Stroop-like comparison task created by Babai et al. (2006), which compared the automaticity of area vs. perimeter processing. Previous findings demonstrated that in a perimeter comparison task, the figure's area was processed involuntarily and thus affected perimeter judgment. In the *congruent condition*, the figure with the greater perimeter also had a larger area, while in the *incongruent condition*, the perimeters of the two figures were equal, while one of them had a larger area. Results indicated that participants showed a clear *Stroop effect*, namely they responded faster to congruent than to incongruent trials (Stavy et al., 2006; Stavy and Babai, 2008, 2010). Thus, the irrelevant area aspect might be automatically processed and, despite being irrelevant

to the task, interfere with perimeter processing. Notably, the opposite task, that is when participants were asked to determine which figure has the largest area (and to ignore the irrelevant perimeter) was examined as well. Results showed that of the two tasks, area comparisons were significantly faster than perimeter comparisons. However, unlike the perimeter comparison task (where area is irrelevant to the task), in the area comparison task (where perimeter is irrelevant to the task) no differences were found between congruent and incongruent trials (Babai et al., 2006). The researchers inferred that area is more salient and more rapidly processed (i.e., more automatic) than perimeter.

To this end, it is not clear whether one magnitude (area) can be more perceptually salient than the other (perimeter). According to general magnitude mechanism theories, both magnitudes should be involuntarily processed and hence are supposed to interfere with each other to a similar degree (Newcombe, 2014; Leibovich et al., 2016). On the other hand one magnitude, in this case area (as found in Babai et al., 2006), can interfere with the processing of another magnitude (perimeter), but not vice versa. A possible solution for this contradiction is that magnitude saliency differs due to task demands (Spelke et al., 2010). Spelke et al. (2010) identified two core systems used in geometric processing: the first is important for navigation at large – scale surfaces (Cheng and Newcombe, 2005) while the second system is used for small object recognition, such as the figures in Babai et al.'s (2006) task. Each system relates to different neural and cognitive processing, with limited transference between them (Derdikman and Moser, 2010; Huang and Spelke, 2015). Accordingly, a particular magnitude can be automatically processed by one of the core systems but not by the other. For instance, the processing of a surface's layout (or, in other words, *perimeter*) is highly crucial for successful navigation (Hermer and Spelke, 1994; Hermer-Vazquez et al., 1999; Lee et al., 2012) but not as crucial for recognition of 2D figures (Lee and Spelke, 2011).

## Developmental Dyscalculia and Magnitude Processing

The current research also aims to expand the knowledge about the interactions between numerical cognition and magnitude processing by including a specific clinical population – participants with deficient numerical abilities (DD). DD is a specific deficit of numerical and mathematical abilities, with a neuro-anatomical source (Rotzer et al., 2008; Kaufmann et al., 2013), affecting about 3.6–6.5% of the population (Geary, 1993; Shalev and Gross-Tsur, 2001; Butterworth, 2005; Reigosa-Crespo et al., 2012). Individuals with DD fail to master common numerical and arithmetical skills, such as numerical comparisons (Rubinsten and Henik, 2005; Mussolin et al., 2010a), arithmetical fact retrieval (Mazzocco et al., 2008), and procedural knowledge (Desoete et al., 2009). Currently, there is a debate about the cognitive mechanisms that underlie numerical difficulties, and DD seems to be a heterogeneous disorder (Rubinsten and Henik, 2009; Kaufmann et al., 2013; Träff et al., 2016). One main explanation for the heterogeneity of DD is that numerical ability cannot be described as a single cognitive mechanism and arithmetic competence relies on domain-specific as well as domain-general skills (Rubinsten and Sury, 2011;

Kaufmann et al., 2013; Szűcs and Goswami, 2013; Bugden and Ansari, 2015; Huber et al., 2015; Noël et al., 2016). From one perspective, individuals with DD may have a domain-specific deficient “number sense,” manifested by difficulties with representing and manipulating all kinds of numerical notations: non-symbolic – such as dot arrays (Price et al., 2007; Mussolin et al., 2010b), or symbolic, namely Arabic numerals and number words (Rubinsten and Henik, 2005, 2006; Mussolin et al., 2010a). Concurrent with the general magnitude processing mechanism theory, they may also ineffectively process continuous magnitudes and have deficient “magnitude sense” (Leibovich et al., 2016). It is worth noticing though, that in a recent study adolescents with DD showed similar performance, on both behavioral and neuronal levels, as typical developing peers when performing non-symbolic numerical comparison tasks (McCaskey et al., 2017). However, participants with typical development activated domain-specific magnitude related areas while performing the task, while participants with DD activated domain-general frontal areas, related to inhibition and working memory. Accordingly, the authors inferred that domain-general deficits could also account for the development of DD.

Indeed, the role of inhibition, a domain-general mechanism, in intact and deficient numerical processing has also received scientific attention (Ashkenazi et al., 2009; Ashkenazi and Henik, 2010; Wang et al., 2012; Szucs et al., 2013a; Bugden and Ansari, 2015; Noël et al., 2016). From this perspective, individuals with DD show deficient performance on numerical tasks due to failure to inhibit irrelevant magnitude information, such as the overall area of dot arrays in a non-symbolic comparison task (Bugden and Ansari, 2015) or physical size in the numerical Stroop task (Szucs et al., 2013a). It is important to notice that these findings remained consistent after controlling for comorbidity (Wang et al., 2012).

## The Current Research

The discussion of a general magnitude processing mechanism as a basis for numerical cognition development seems to be more relevant than ever (Gebuis and Reynvoet, 2012b; Leibovich and Henik, 2014; Newcombe, 2014; Lourenco and Bonny, 2016; Leibovich et al., 2016). However, there exists no extensive work dealing with intact and deficient processing of this system (Skagerlund and Träff, 2014; McCaskey et al., 2017). Moreover, it is necessary to differentiate between “pure,” domain-specific, and domain-general mechanisms when performing magnitude comparison tasks, in order to define the source of DD difficulties.

Hence, we compared participants with DD and typically developing participants while performing a Stroop-like task (adopted from Babai et al., 2006). We expected to find group differences, as measured by Stroop effects (incongruent minus congruent trials) in both area and perimeter comparison tasks. According to domain-specific magnitude processing deficits (Newcombe, 2014; Leibovich et al., 2016), the DD group is expected to show smaller Stroop effects implying that they do not process the irrelevant magnitude (similar to the numerical Stroop task, Rubinsten and Henik, 2005). On the other hand, domain-general deficits (Wang et al., 2012; Szucs et al., 2013a; Bugden and Ansari, 2015) should result in larger Stroop effects in the DD

group, due to a deficit in the ability to ignore irrelevant magnitude information.

Another interesting specific question is whether participants with DD will show similar deficient processing on both the area and perimeter tasks. If the area component is indeed more perceptually salient (Babai et al., 2006), the DD group should show larger Stroop effects on the perimeter task (in which area is irrelevant to the task), namely they should find it harder to ignore the irrelevant area aspect.

The cognitive method as well as the statistical analysis of the current study enabled us not only to study the differences between automatic processing of area vs. perimeter (i.e., investigating magnitude sense in DD), but also to investigate whether DD participants indeed perform poorly or differently on continuous magnitude processing tasks in initial vs. proficiency stages of learning (i.e., to investigate learning functions in DD). Earlier studies showed that even a small number of rehearsals of numerical problems led to automatic processing and to changes in brain functions (Ischebeck et al., 2007; Aubin et al., 2016). For instance, Ischebeck et al. (2007) found that very short training (eight repetitions) in multiplication problems led to a decrease in the activity of fronto-parietal brain areas related to calculation and numerical processing (Menon et al., 2000; Dehaene et al., 2003). On the other hand, the training also resulted in increased activity in temporo-parietal regions known to be involved in arithmetic fact retrieval (Dehaene et al., 2003). Recently, it was proposed that DDs’ deficits in inhibition of irrelevant numerical information can also represent difficulties with consolidating learned information and with performing the shift from initial computing based processing to automatic retrieval based processing (Ischebeck et al., 2007; Aubin et al., 2016). Indeed, Aubin et al. (2016) suggested that people with DD may be less able to consolidate a numerical task within the frontal-parietal region and must instead rely on their working memory. Consequently, there would be limited progression to recalling numerical information and a continued dependence on working memory. Accordingly, it is predicted here that people with DD may need more rehearsals in order to attain proficiency in performing magnitude tasks. Such difficulty with consolidating learned knowledge and with using advanced retrieval strategies will produce consistent group differences predicted to be prominent in the perimeter task, evident even after continuous exposure to stimuli.

## MATERIALS AND METHODS

### Participants

Fifty-seven adults, 27 typically developing (i.e., control group; including 9 males, 18 females; mean age = 24.92 years,  $SD = 2.67$  years) and 30 with DD (2 males, 28 females; mean age = 24.43 years,  $SD = 2.74$  years) participated in the study. All participants were university students and had successfully completed math matriculation exams. Participants were recruited by advertisements distributed on campus and gave written consent to participate in the experiment. Some of them were paid about USD15 for their participation, while other received

credit points for academic courses. The study was carried out in accordance with the recommendations of the ethics committee of the University of Haifa with written informed consent from all participants. All participants gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the ethics committee of the University of Haifa (No. 108/09).

## Classification and Assessment Criteria

Participants were assigned to either a control or a DD group, using the “Israeli learning function diagnosis system” (also titled “MATAL” in Hebrew) for high school and higher education students (National Institute for Testing and Evaluation). This diagnostic tool is composed of a set of standardized computerized tests and questionnaires intended for diagnosing learning disabilities in high school and higher education students. All tests and questionnaires used are nationally normalized. All participants performed numerical tasks (simple calculation and procedural knowledge calculation) and reading related tasks (text reading, phoneme omission, and rapid naming). They also answered a questionnaire (based on the DSM) regarding their attention ability, and performed Raven’s Standard Progressive Matrices (SPM) test (Raven et al., 1995) in order to rule out non-verbal mental disabilities.

Participants were defined as having DD if their scores in either RT or accuracy (ACC) measures on the simple calculation and procedural knowledge tests were worse than mean  $-1$  SD and their scores in the reading and attention tests were more than  $+1$  SD above the mean. Participants in the control group scored better than mean  $-1$  SD in numerical, reading, and attention tests. All participants also scored above the 25th percentile in the SPM test. Independent  $t$ -tests were conducted on the different test results. The two groups differed significantly in both numerical tests (for mean test results and  $p$ -values of independent  $t$ -tests see **Table 1**). As significant group differences were also evident in reading related tasks, group results did not exceed 1SD lower than the norm score, hence it can be assumed that all participants were good readers.

## The Experimental Tasks – Area and Perimeter Tasks

The experiment was run on a PC using E-Prime 2.0 software, and contained two tasks, based on the comparison task of Babai et al. (2006). In each task, participants were requested to relate to a specific aspect of the stimuli (i.e., area or perimeter). Participants were presented with two polygons in each trial, and were asked to decide which polygon had the largest amount of one of the magnitude aspects (area/perimeter). Each block contained a similar number of congruent trials (where both area and perimeter are increased between stimuli) and incongruent trials (where one aspect is increased while the other is reduced between stimuli). Semi-neutral trials (where one aspect is equal and the other differs between stimuli) and equal trials (where both aspects are equal between stimuli) were added as fillers.

## Stimuli

Each stimulus was composed of two figures – a basic rectangle or a polygon derived from the basic rectangle by adding or subtracting one or two squares (the size of each square was  $1/12$  of the basic rectangle size). Pairing of polygons was based on their protrusion direction (into the basic rectangle, see **Figure 1B**, or protruding out of it, see **Figure 1A**) and congruency rules (described above). Each polygon was ascribed to two different stimuli in order to avoid visual bias. Moreover, in order to avoid visual bias and learning, each stimulus was presented eight times, four times on the left side of the fixation point and four times on the right side of the fixation point. Each stimulus was also presented twice in four directions – two on the vertical axis – original (turning up) or rotated by  $180^\circ$  (turning down) and two on the horizontal axis – rotated by  $90^\circ$  to the left (turning left) or to the right (turning right).

The task contained two blocks, one with filled figures and the other containing framed only figures, in order to avoid visual bias of one aspect (Stavy and Babai, 2008). Each block contained 32 “initial exposure” trials (following Ischebeck et al., 2007), representing the *practice* phase. We aimed to capture the reaction to unfamiliar stimuli – eight congruent trials (both area and perimeter were larger in one of the polygons), eight incongruent trials (where the area was larger for one polygon while the perimeter was larger in the other), and 16 filler trials (in which one or both aspects remained equal). After succeeding on more than 80% of the practice trials (in the first or second presentation), participants were able to continue to the continuous exposure, or *proficiency* phase, representing automatization levels. The proficiency phase contained 128 trials – 32 congruent trials (both area and perimeter were larger in one of the polygons), 32 incongruent trials (where the area was larger for one polygon while the perimeter was larger in the other), and 64 filler trials (in which one or both aspects remained equal). The filler trials enabled sufficient presentation of each stimulus and contained 32 equal trials, with the same polygons presented at different rotations, and 32 semi-neutral trials, where one aspect (area or perimeter) remained constant between polygons while the other differed. Since there were two experimental tasks – area and perimeter comparison – each experimental condition, congruent and incongruent, was represented in total by 32 trials for the practice phase ( $8 \times 2$  blocks, filled or framed  $\times 2$  tasks – area and perimeter) and 128 trials for the proficiency phase ( $32 \times 2$  blocks, filled or framed  $\times 2$  tasks – area and perimeter). For illustration of the experimental blocks see **Figure 2**.

## Procedure

Participants were seated about 60 cm from the computer screen. The participants’ goal was to decide, in two separate tasks, which polygon has the largest area (area task) or perimeter (perimeter task). Then, they were asked to convey their answer as quickly and accurately as possible by pressing one of three marked keys on a response box (right key for right polygon, left key for left polygon, central key for equal area or perimeter). Each participant was presented with

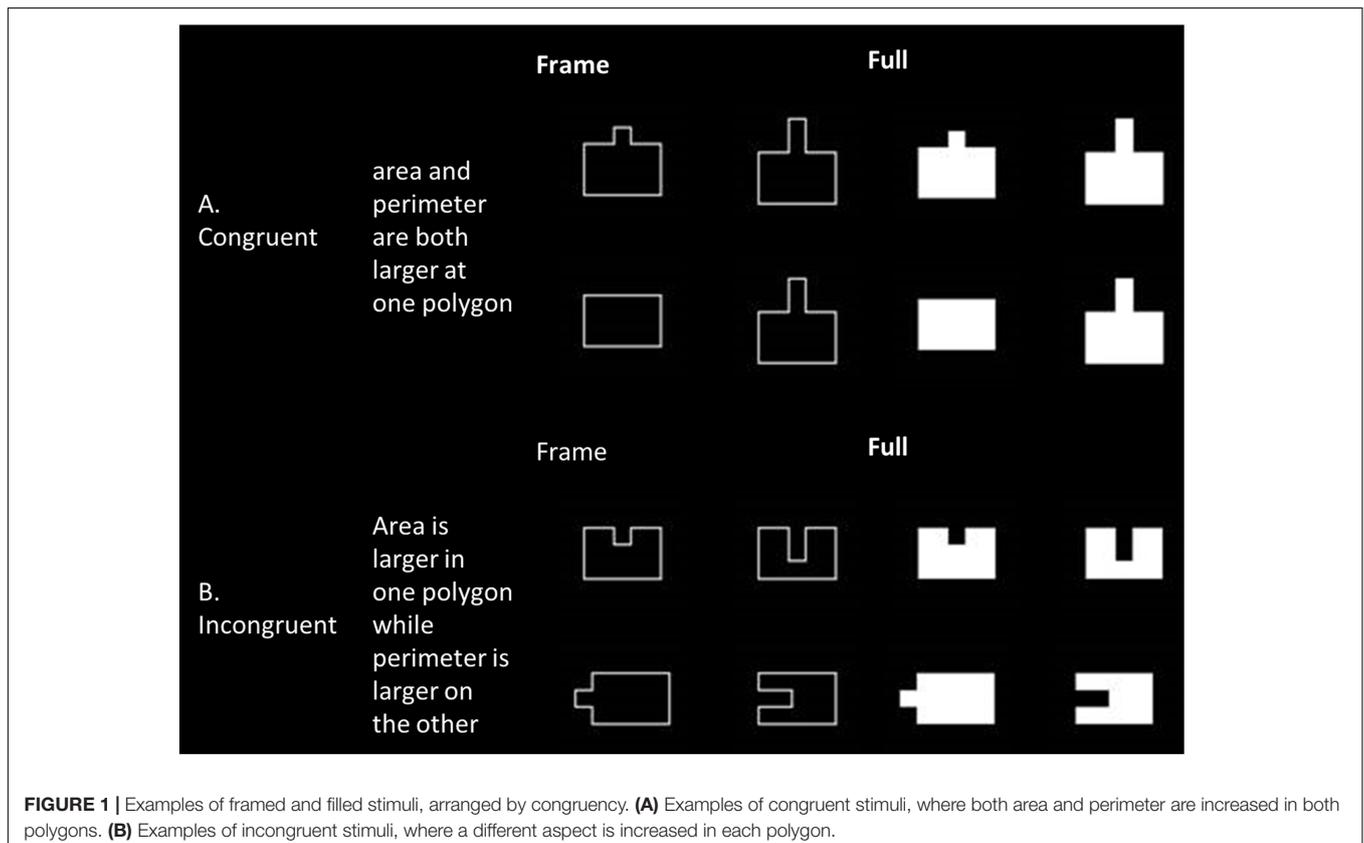
**TABLE 1** | Mean standard scores and group differences in the screening tests of the two groups.

Test	Control		DD		t-test	
	ACC	RT	ACC	RT	ACC	RT
Text reading	0.56	0.42	0.48	-0.31	0.37	2.78**
Phoneme omission	0.46	0.65	-0.05	0.01	2.34*	2.53*
Rapid naming		0.14		-0.32		1.94
	Objects					
	Letters	0.64		0.33		1.21
	Numbers	0.42		-0.15		2.08*
Questionnaire	Attention	0.09	0.12			-0.37
	Impulsiveness and hyperactivity	0.00	0.86			-1.57
Simple calculation	0.61	0.49	-1.3	-1.37	4.79**	5.93**
Procedural knowledge	0.46	0.59	-1.55	-1.26	6.05**	6.92**
SPM (percentiles)	58.4		52.43		1.25	

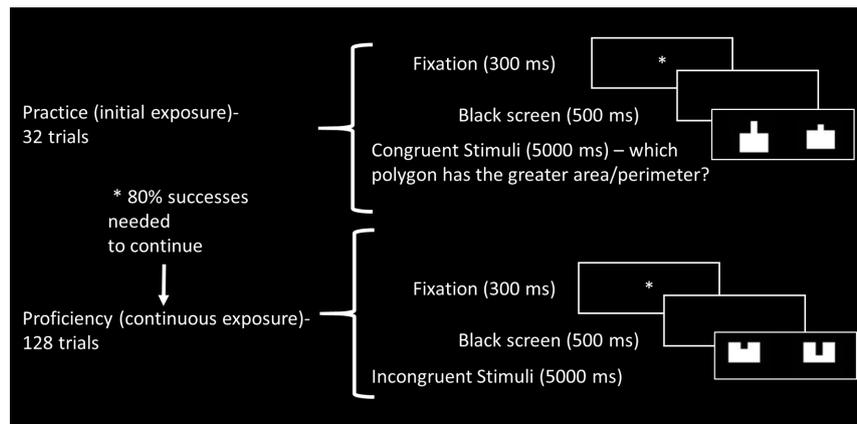
ACC = standard scores for accuracy rates; RT = standard scores for reaction times; SPM = Standard Progressive Matrices (Raven et al., 1995); Sig. = independent t-test significance between control and DD groups. Significance values: \* $p < 0.05$ ; \*\* $p < 0.01$ .

four blocks (two area blocks – filled and framed, and two perimeter blocks – filled and framed). Half the participants began with the area task (25% of participants began with a filled figures block and the rest with a framed figures block) and half with the perimeter task (25% began with a filled figures block and the rest with a framed figures block). Tasks and block performance order were counter balanced and were determined in advance according to participants' serial numbers.

The experiment included three breaks in each block, which were terminated when participants pressed a relevant key, as well as a break of a few minutes between the sections. The stimuli in each trial began with a fixation point (small white asterisk), which appeared for 300 ms and was followed by an empty black screen for 500 ms. Then the sample polygons appeared and remained in view until the participant pressed a key but no longer than 5,000 ms. The next trial began with the fixation point. Overall, the experimental tasks took about 1 h.



**FIGURE 1** | Examples of framed and filled stimuli, arranged by congruency. **(A)** Examples of congruent stimuli, where both area and perimeter are increased in both polygons. **(B)** Examples of incongruent stimuli, where a different aspect is increased in each polygon.



**FIGURE 2** | Illustration of experimental block, including initial (practice) and continuous exposure (proficiency) trials.

## Data Analysis

### Error Rates Analysis

Mean error rates (percentage of incorrect trials) were calculated for each participant for the practice (only the first practice was calculated for participants who failed to succeed in over 80% of the practice trials) and proficiency phases separately. Then, three-way repeated measures ANOVAs were used, with group (i.e., Control or DD) as the between-subject factor, and congruency (i.e., congruent or incongruent) and task (i.e., Area or Perimeter) as within-subject factors on the error rates of practice and proficiency trials. All of the following  $F$ -statistics were adjusted by the Greenhouse-Geisser correction.

In order to define the triple group  $\times$  task  $\times$  congruency interaction, Stroop effects of the error rates were calculated (by subtracting error rates of incongruent trials from error rates of congruent trials) for each group, in Area and Perimeter tasks separately. In order to test for a continuous exposure effect, we compared the Stroop effects of practice and proficiency phases of each task for each group separately, by using paired  $t$ -tests.

### Reaction Times (RT) Analysis

Reaction times analysis was similar to error rates analysis. Mean RTs (in ms) were calculated for each participant for practice and proficiency phases separately. Then, three-way repeated measures ANOVAs were used, with group as the between-subject factor and congruency and task as within-subject factors on the RTs of practice and proficiency trials. A continuous exposure effect was observed by comparing the Stroop effects of the practice and proficiency phases of each task for each group separately, using paired  $t$ -tests.

### Other Visual Features

We used repeated measures ANOVAs, similar to those described above, in order to investigate whether other visual features manipulated (i.e., figure filling, protrusion, and direction – see **Figure 1**) were possible confounders of Area and Perimeter processing.

### Gender Differences

Gender differences were tested using independent  $t$ -tests.

## RESULTS

### Gender Differences

The research sample contained a larger number of females. Accordingly, we tested gender differences (mean scores and  $t$ -tests are presented in **Table 2**). Females and males showed similar patterns across tasks and proficiency stages. Specifically, they showed similar Stroop effects in both tasks in both error rates and speed analyses.

### Visual Features

Several visual features were manipulated in order to track possible confounding with the experimental variables. The filling effect (filled vs. framed figures) was not significant in the practice phase, neither for error rates [ $F_{(1,55)} = 0.10, p = 0.746, \eta^2 = 0.002$ ] nor for RTs [ $F_{(1,55)} = 0.39, p = 0.533, \eta^2 = 0.008$ ]. Moreover, there was no filling effect for error rates in the proficiency phase [ $F_{(1,55)} = 0.11, p = 0.741, \eta^2 = 0.002$ ]. However, participants responded more slowly to framed figures ( $M = 977.95, SD = 229.58$ ) than to filled figures ( $M = 919.18, SD = 194.74$ ) in the proficiency phase [ $F_{(1,55)} = 6.31, p = 0.015, \eta^2 = 0.103$ ]. Furthermore, in this phase, a filling  $\times$  congruency interaction was evident [ $F_{(1,55)} = 5.28, p = 0.025, \eta^2 = 0.08$ ], such that the Stroop effect (incongruent minus congruent) was significantly greater [ $F_{(1,56)} = 4.58, p = 0.037, \eta^2 = 0.07$ ] for framed figures ( $M = 236.15, SD = 133.19$ ) than for full figures ( $M = 196.77, SD = 106.72$ ).

We further analyzed the effect of continuous exposure for each filling type separately. This was done in order to figure out whether stimuli rehearsal led to changes in the automaticity of the filling effect. On the error analysis, participants made more errors (in percentage) in congruent trials than in incongruent trials in framed figures after initial exposure ( $M$  Stroop effect =  $-0.05, SD = 0.07$ ) and this difference decreased after continuous exposure ( $M$  Stroop effect =  $-0.00, SD = 0.02$ ) to stimuli

$[t_{(56)} = -4.06, p < 0.001, d = 1.085]$ . For full figures, no difference between Stroop effects after initial ( $M = -0.03, SD = 0.07$ ) or continuous exposure ( $M = -0.01, SD = 0.04$ ) was evident [ $t_{(56)} = -1.51, p = 0.137, d = -0.403$ ]. Reaction times analysis show that the Stroop effects (in ms) were larger after initial exposure and decreased after continuous exposure for both framed [initial exposure:  $M = 326.46, SD = 201.91$ ; continuous exposure:  $M = 243.28, SD = 134.87$ ;  $t_{(56)} = 3.71, p = 0.001, d = 0.991$ ] and full figures [initial exposure:  $M = 290.99, SD = 213.14$ ; continuous exposure:  $M = 211.03, SD = 123.41$ ;  $t_{(56)} = 2.86, p = 0.006, d = 0.764$ ].

There were no other interactions involving the filling component. No significant findings were found for error rates or RTs for other visual features manipulated, namely protrusion (protruding in or out – see **Figure 1**) and vertical or horizontal direction.

## Error Rates Analysis

### Practice Phase

The current analysis aims to define whether Area interferes with Perimeter processing and vice versa, among participants with intact and deficient numerical processing after initial exposure to non-familiar stimuli.

Mean error rates (in percentage) in area and perimeter tasks across different exposure phases of the two groups are presented in **Table 3**. Results revealed a significant effect of group [ $F_{(1,55)} = 5.49, p = 0.023, \eta^2 = 0.091$ ], indicating that the DD group made more errors than the Control group. The main effect of congruency was also significant [ $F_{(1,55)} = 39.81, p < 0.001, \eta^2 = 0.420$ ], as congruent trials had smaller error rates than incongruent trials. No main effect of task was evident [ $F_{(1,55)} = 3.56, p = 0.064, \eta^2 = 0.061$ ], with similar error rates for both Area and Perimeter tasks.

However, the task  $\times$  congruency interaction was significant [ $F_{(1,55)} = 22.72, p < 0.001, \eta^2 = 0.292$ ]. Further analysis revealed that incongruent trials were less accurate than congruent trials on the Perimeter task [ $t_{(56)} = 6.7, p < 0.001, d = 1.79$ ]. No difference between error rates of congruent and incongruent trials was evident in the Area task [ $t_{(56)} = 1.32, p = 0.192, d = 0.352$ ]. Importantly, no order effect was evident, as participants who started with the Perimeter task showed similar Stroop effects as participants starting with the Area task in both tasks [Area task:

$t_{(55)} = 0.767, p = 0.446, d = 0.206$ ; Perimeter task:  $t_{(55)} = 0.004, p = 0.997, d = 0.001$ ].

Interestingly, and with high relevance to the current research questions, a triple interaction was evident [ $F_{(1,55)} = 6.69, p = 0.012, \eta^2 = 0.109$ ], as described in **Figure 3**. Further analysis of this interaction revealed significant effects for group [ $F_{(1,55)} = 4.73, p = 0.034, \eta^2 = 0.079$ ] and congruency [ $F_{(1,55)} = 47.72, p < 0.001, \eta^2 = 0.454$ ] in the Perimeter task (in which Area is irrelevant to the task). Moreover, in this Perimeter task, the group  $\times$  congruency interaction reached significance as well [ $F_{(1,55)} = 3.81, p = 0.050, \eta^2 = 0.065$ ]. The Stroop effect (incongruent minus congruent) of error rates was larger for the DD group ( $M = 0.10, SD = 0.09$ ) than for the Control group ( $M = 0.05, SD = 0.08$ ). In the Area task, no main effect of group [ $F_{(1,55)} = 1.87, p = 0.176, \eta^2 = 0.033$ ] or congruency [ $F_{(1,55)} = 2.04, p = 0.158, \eta^2 = 0.036$ ] was evident, nor group  $\times$  congruency interaction [ $F_{(1,55)} = 2.8, p = 0.098, \eta^2 = 0.049$ ]. In other words, in the Area task, there was no meaningful Stroop effect of error rates, and this pattern was similar for both Controls ( $M = 0.02, SD = 0.06$ ) and DD ( $M = 0.00, SD = 0.06$ ). In conclusion, while the DD group showed larger Stroop effects in the perimeter task, no group difference was evident in the area task.

### Proficiency Phase

Contrary to the former analysis, this analysis aims to find out whether magnitude interference and group differences are evident after continuous exposure, when participants are well familiar with the stimuli.

The main effect of task [ $F_{(1,55)} = 5.80, p = 0.019, \eta^2 = 0.096$ ] was minor but significant. Importantly, no order effect was evident, as participants who started with the Perimeter task showed similar Stroop effects as participants starting with the Area task in both tasks [Area task:  $t_{(55)} = 1.3, p = 0.196, d = 0.351$ ; Perimeter task:  $t_{(55)} = -0.55, p = 0.582, d = -0.148$ ]. The main effect of congruency reached significance as well [ $F_{(1,55)} = 18.63, p < 0.001, \eta^2 = 0.253$ ]. However, no main effect of group was evident [ $F_{(1,55)} = 0.03, p = 0.852, \eta^2 = 0.001$ ], nor any interaction. This suggests that there was a typical Stroop effect, which was similar in pattern across groups (DD and Controls) and tasks (Area and Perimeter).

**TABLE 2** | Mean scores and gender differences in area and perimeter tasks across different exposure phases.

	Task	Females <i>M</i> ( <i>SD</i> )	Males <i>M</i> ( <i>SD</i> )	<i>t</i> -test	<i>p</i> -value
Stroop effects of error rates (in percentage)	Practice phase- Area	0.01 (0.06)	0.01 (0.06)	0.468	0.642
	Practice phase- Perimeter	0.08 (0.09)	0.05 (0.07)	1.179	0.244
	Proficiency phase- Area	0.02 (0.04)	0.00 (0.03)	1.35	0.183
	Proficiency phase- Perimeter	0.01 (0.03)	0.01 (0.02)	-0.619	0.539
Stroop effects of reaction times (in ms)	Practice phase- Area	308.92 (217.90)	194.09 (205.24)	1.69	0.097
	Practice phase- Perimeter	378.85 (287.00)	266.23 (272.40)	1.25	0.214
	Proficiency phase- Area	227.66 (101.57)	270.58 (201.45)	-1.04	0.300
	Proficiency phase- Perimeter	242.73 (175.33)	193.76 (85.69)	0.969	0.337

*Stroop effects = incongruent minus congruent trials. Sig. = independent t-test between females and males.*

**TABLE 3** | Mean error rates (in percentage) in area and perimeter tasks across different exposure phases of the two groups.

Group	Practice phase				Proficiency phase			
	Area		Perimeter		Area		Perimeter	
	Congruent <i>M</i> ( <i>SD</i> )	Incongruent <i>M</i> ( <i>SD</i> )	Congruent <i>M</i> ( <i>SD</i> )	Incongruent <i>M</i> ( <i>SD</i> )	Congruent <i>M</i> ( <i>SD</i> )	Incongruent <i>M</i> ( <i>SD</i> )	Congruent <i>M</i> ( <i>SD</i> )	Incongruent <i>M</i> ( <i>SD</i> )
Control	0.02 (0.03)	0.04 (0.05)	0.01 (0.02)	0.07 (0.07)	0.03 (0.03)	0.04 (0.03)	0.02 (0.02)	0.03 (0.03)
DD	0.05 (0.06)	0.05 (0.05)	0.01 (0.02)	0.12 (0.09)	0.03 (0.01)	0.04 (0.02)	0.01 (0.01)	0.03 (0.04)

### Continuous Exposure Effect (i.e., The Statistical Differences between Stroop Effects in the Practice vs. Task Phases)

The effect of exposure was observed by comparing the Stroop effects in the practice and proficiency phases of each task for each group separately, using paired *t*-tests. This in order to figure out whether stimuli rehearsal led to changes in automaticity of one or both magnitudes among different participants.

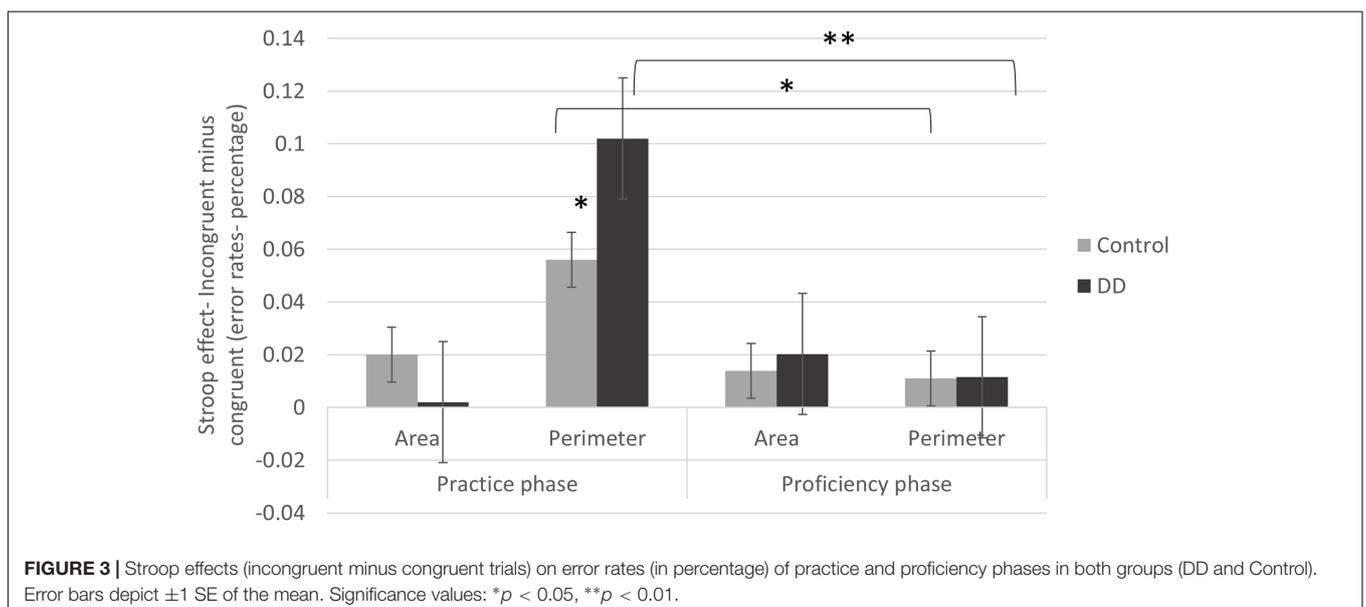
As described in **Figure 3**, in the Perimeter task a significant difference was found between the Stroop effects of the practice and proficiency phases for the DD group [ $t_{(29)} = 5.37$ ,  $p < 0.001$ ,  $d = 1.994$ ; practice phase:  $M = 0.10$ ,  $SD = 0.09$ ; proficiency phase:  $M = 0.015$ ,  $SD = 0.02$ ]. A similar significant difference was evident for the Control group [ $t_{(26)} = 2.62$ ,  $p = 0.015$ ,  $d = 1.027$ ; practice phase:  $M = 0.05$ ,  $SD = 0.08$ ; proficiency phase:  $M = 0.01$ ,  $SD = 0.03$ ]. Accordingly, the initial Stroop effects declined after practice and this pattern was evident in both groups. No difference was evident in the Stroop effects in the Area task neither for the DD group [ $t_{(29)} = -1.57$ ,  $d = -0.583$ ,  $p = 0.126$ ; practice phase:  $M = 0.002$ ,  $SD = 0.06$ ; proficiency phase:  $M = 0.02$ ,  $SD = 0.04$ ] nor for the Control group [ $t_{(26)} = 1.03$ ,  $p = 0.309$ ,  $d = 0.403$ ; practice phase:  $M = 0.02$ ,  $SD = 0.06$ ; proficiency phase:  $M = 0.01$ ,  $SD = 0.04$ ].

### Reaction Time Analysis (RT) Practice Phase

Mean reaction times (in ms) in area and perimeter tasks across different exposure phases of the two groups are presented in **Table 4**. When testing whether Area interferes with Perimeter processing and vice versa among participants with different numerical abilities after initial exposure to non-familiar stimuli, a main effect of congruency was found [ $F_{(1,55)} = 215.54$ ,  $p < 0.001$ ,  $\eta^2 = 0.797$ ]. Accordingly, participants responded more slowly to incongruent trials than to congruent trials. The main effect of group was not significant [ $F_{(1,55)} = 1.74$ ,  $p = 0.191$ ,  $\eta^2 = 0.031$ ], nor was the main effect of task [ $F_{(1,55)} = 2.17$ ,  $p = 0.146$ ,  $\eta^2 = 0.038$ ].

However, and following the experimental hypothesis, the group  $\times$  congruency interaction was found to be significant [ $F_{(1,55)} = 5.09$ ,  $p = 0.028$ ,  $\eta^2 = 0.085$ ]. The Stroop effect (incongruent minus congruent) was significantly larger [ $F_{(1,55)} = 2.25$ ,  $p = 0.028$ ] for the DD group ( $M = 363.9$ ,  $SD = 147.86$ ) than for the Control group ( $M = 266.89$ ,  $SD = 176.33$ ).

No triple interaction was found [ $F_{(1,55)} = 0.87$ ,  $p = 0.355$ ,  $\eta^2 = 0.016$ ]. However, following scientific background and interest, we further analyzed the triple interaction (see **Figure 4**).



**TABLE 4 |** Mean reaction times (in ms) in area and perimeter tasks across different exposure phases of the two groups.

Group	Practice phase				Proficiency phase			
	Area		Perimeter		Area		Perimeter	
	Congruent <i>M (SD)</i>	Incongruent <i>M(SD)</i>	Congruent <i>M (SD)</i>	Incongruent <i>M (SD)</i>	Congruent <i>M (SD)</i>	Incongruent <i>M (SD)</i>	Congruent <i>M (SD)</i>	Incongruent <i>M (SD)</i>
Control	1087.50 (246.43)	1344.09 (213.55)	1144.53 (273.77)	1421.73 (396.78)	787.55 (135.92)	995.53 (200.28)	812.73 (137.41)	1008.33 (269.53)
DD	1153.47 (289.83)	1459.74 (267.62)	1142.79 (296.73)	1564.34 (366.68)	898.98 (181.51)	1162.95 (238.72)	896.42 (249.88)	1160.36 (281.49)

In the Perimeter task, a significant congruency effect was found [ $F_{(1,55)} = 89.53, p < 0.001, \eta^2 = 0.619$ ], but no group effect [ $F_{(1,55)} = 0.74, p = 0.391, \eta^2 = 0.013$ ]. Furthermore, a marginally significant group  $\times$  congruency interaction was evident [ $F_{(1,55)} = 3.82, p = 0.056, \eta^2 = 0.065$ ]. The Stroop effect (incongruent minus congruent) tended to be larger in the DD group ( $M = 421.54, SD = 281.00$ ) than in the Control group ( $M = 277.19, SD = 275.51$ ). In the Area task, a significant congruency was found [ $F_{(1,55)} = 93.63, p < 0.001, \eta^2 = 0.630$ ], but no group effect [ $F_{(1,55)} = 2.16, p = 0.147, \eta^2 = 0.038$ ]. No group  $\times$  congruency interaction was found [ $F_{(1,55)} = 0.72, p = 0.397, \eta^2 = 0.013$ ]. In other words, in the Area task, there was a typical Stroop effect, and this pattern was similar for Controls ( $M = 256.59, SD = 198.83$ ) and DD ( $M = 306.27, SD = 236.10$ ). Similar to error rates analysis, the DD group tended to show larger Stroop effects (marginally significant) in the perimeter but not in the area task.

### Proficiency Phase

While aiming to find out whether magnitude interference and group differences were evident after continuous exposure, results revealed a main effect of group [ $F_{(1,55)} = 6.28, p = 0.015, \eta^2 = 0.103$ ]. DDs' RTs were significantly slower than Controls'. The main effect of congruency was also significant [ $F_{(1,55)} = 241.8, p < 0.001, \eta^2 = 0.815$ ], such that participants responded more slowly to incongruent trials than to congruent trials. However, the main effect of task did not reach significance [ $F_{(1,55)} = 0.16, p = 0.684, \eta^2 = 0.003$ ], such that participants responded similarly to Area and to Perimeter tasks.

Importantly and following the experimental hypothesis, the group  $\times$  congruency interaction was found to be significant [ $F_{(1,55)} = 04.30, p = 0.043, \eta^2 = 0.073$ ]. The Stroop effect (incongruent minus congruent) was significantly larger [ $F_{(1,55)} = 4.01, p = 0.043, \eta^2 = 0.068$ ] for the DD group ( $M = 263.95, SD = 107.74$ ) than for the Control group ( $M = 201.79, SD = 118.39$ ).

No triple interaction was found between group, congruency, and task in the proficiency phase [ $F_{(1,55)} = 0.06, p = 0.794, \eta^2 = 0.001$ ]. However, following scientific interest, the double interaction between congruency and group was analyzed separately in each task. In the Perimeter task, significant congruency [ $F_{(1,55)} = 120.68, p < 0.001, \eta^2 = 0.687$ ] and a marginally significant group effect [ $F_{(1,55)} = 3.74, p = 0.058,$

$\eta^2 = 0.064$ ] were found. However, no group  $\times$  congruency interaction was evident [ $F_{(1,55)} = 2.66, p = 0.108, \eta^2 = 0.046$ ] and both groups showed typical Stroop effects (DD:  $M = 263.93, SD = 150.83$ ; Control:  $M = 195.60, SD = 165.02$ ). Similarly, in the Area task, significant congruency [ $F_{(1,55)} = 192.45, p < 0.001, \eta^2 = 0.778$ ] and group effects [ $F_{(1,55)} = 8.25, p = 0.006, \eta^2 = 0.131$ ] were found. However, no group  $\times$  congruency interaction was found [ $F_{(1,55)} = 2.7, p = 0.106, \eta^2 = 0.047$ ] and both groups showed typical Stroop effects (DD:  $M = 263.97, SD = 144.12$ ; Control:  $M = 207.98, SD = 107.81$ ). Accordingly, after continuous exposure to stimuli, both groups showed typical Stroop effects in both tasks.

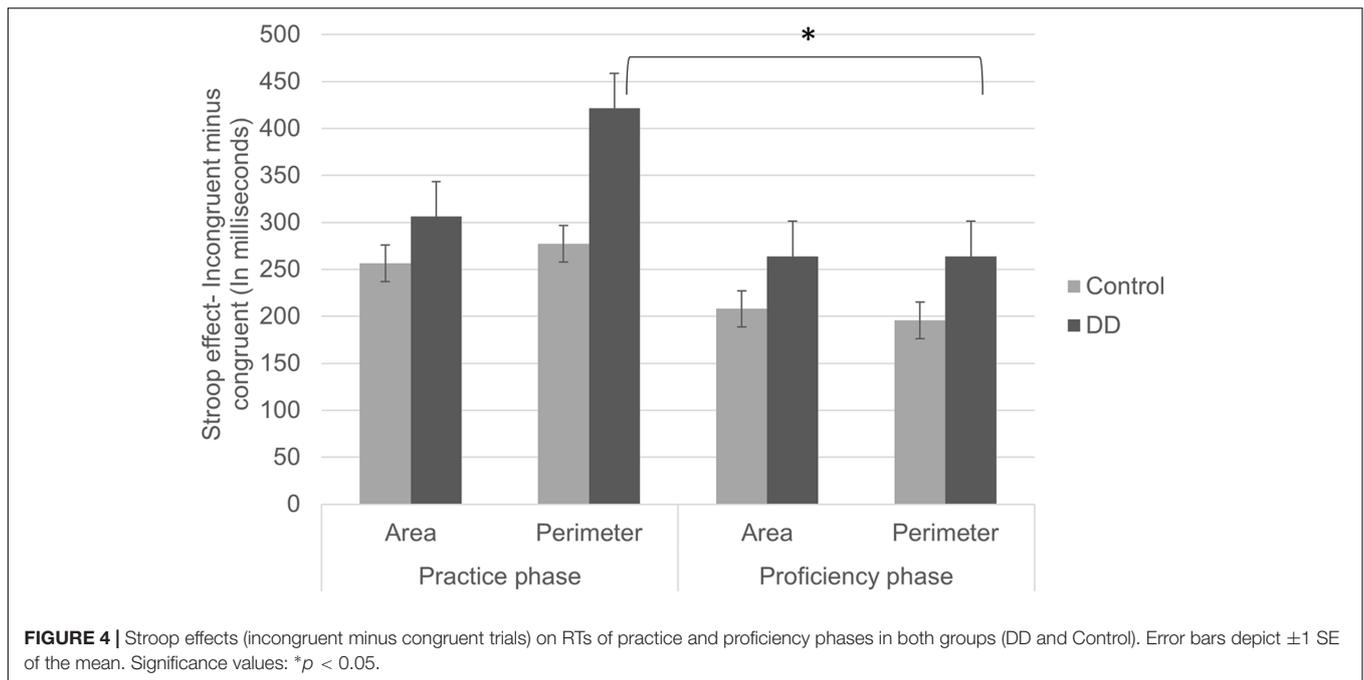
### Continuous Exposure Effect (i.e., The Statistical Differences between Stroop Effects in the Practice vs. Task Phases)

We analyzed the effect of exposure separately for each group in each task. This was done in order to figure out whether stimuli rehearsal led to changes in automaticity of one or both magnitudes among different participants. In the Perimeter task (in which Area is irrelevant to the task), a significant difference was found in the Stroop effects (congruent minus incongruent) between practice and proficiency phases for the DD group [ $t_{(29)} = 3.64, p = 0.001, d = 1.351$ ]. Accordingly, high initial Stroop effects ( $M = 421.54, SD = 281.006$ ) declined after practice ( $M = 263.93, SD = 150.83$ ). No significant difference was evident between Stroop effects in the practice and proficiency phases for the Control group [ $t_{(26)} = 1.62, p = 0.115, d = 0.635$ ; practice phase:  $M = 277.19, SD = 275.51$ ; proficiency phase:  $M = 195.60, SD = 165.02$ ]. No difference in Stroop effects was evident in the Area task (in which Perimeter was irrelevant to the task), neither for the DD group [ $t_{(29)} = 1.17, p = 0.248, d = 0.434$ ; Practice phase:  $M = 306.27, SD = 236.10$ ; Proficiency phase:  $M = 263.97, SD = 144.12$ ], nor for the Control group [ $t_{(26)} = 1.31, p = 0.202, d = 0.513$ ; practice phase:  $M = 256.59, SD = 198.83$ ; proficiency phase:  $M = 207.98, SD = 107.87$ ].

## DISCUSSION

The present study investigated the automaticity of area and perimeter processing at different exposure levels in adults with deficient and intact numerical abilities.

To the best of our knowledge, this is the first study to show that individuals with DD process area and perimeter



information differently and in a less automatic manner than peers with intact numerical competence. After initial exposure to stimuli, area processing was more automatic than perimeter processing for both groups, as represented by task  $\times$  congruency interaction, evident in error rate analysis. Specifically, significant Stroop effects (slower responses to incongruent vs. congruent trials) were evident in perimeter and not in area tasks. However, this pattern was more prominent among the DD group, appearing in both error rates and speed analyses, implying a magnitude processing deficit. Together with previous evidence (Babai et al., 2006), the current findings show that area interferes with perimeter processing but not vice versa. This pattern suggests that while area processing may be innate, perimeter processing is acquired. After continuous exposure, the difference between area and perimeter was no longer evident. Furthermore, significant Stroop effects (in speed analysis) show that both magnitudes were automatically processed and interfered with each other to a similar degree in both groups. However, we found that domain general learning and inhibition deficits are also involved in DD. Specifically, we found firm group  $\times$  congruency interactions in the speed dimension. Namely, the DD group showed larger Stroop effects, which were evident across all exposure levels. Overall, the findings imply that deficient performance of participants with DD may not be restricted to numerical processing.

## Automatic Processing of Area and Perimeter

In the current work, area and perimeter were both found to be automatically processed, as shown by the existence of significant Stroop effects (slower responses to incongruent vs. congruent

trials), evident from speed analysis in both area and perimeter tasks. In other words, the irrelevant aspect (area or perimeter) was involuntarily processed. As in the case of the numerical Stroop (Henik and Tzelgov, 1982), bidirectional effects were evident between the two magnitudes. The fact that both aspects are automatically processed is compatible with recent research highlighting the existence of a general magnitude processing mechanism (Gebuis and Reynvoet, 2012b; Newcombe, 2014; Leibovich et al., 2016).

Our findings are partially consistent with previous studies (Babai et al., 2006) showing that the area aspect interferes with perimeter judging but not vice versa. In the current study, perimeter processing seems to be automatic as well and to interfere with area processing after continuous exposure to the stimuli. However, some of our findings support the assumption that the area aspect is more perceptually salient. In the analysis of error rates in the practice phase we found Stroop effects (when congruent trials were more accurate than incongruent trials) in the perimeter but not in the area task. Area saliency probably relates to the fact that it occupies a larger space (in square meters) (Abbott, 1976). Moreover, magnitude saliency can differ due to task demands (Spelke et al., 2010). From this perspective, the processing of surface layouts (perimeter), which is crucial for navigation tasks (Cheng and Newcombe, 2005), is probably less vital than the figures' area in order to decide which small-scale 2D figure (Lee and Spelke, 2011) is larger, as required in the current task.

Since area saliency disappeared in the proficiency phase, we can assume that exposure levels also account for magnitude saliency. Accordingly, after practicing, participants were able to successfully inhibit irrelevant salient area magnitude information, and no difference was found between area and perimeter Stroop effects. Based on previous study that found

that as few as eight rehearsals trials, contributed to changes in brain functions (Ischebeck et al., 2007), we may argue, that in the current study which included larger number of stimuli, a meaningful learning indeed occurred. The findings are consistent with intervention studies indicating that attracting attention to the perimeter aspect (either by adding warnings or by presenting a polygon's perimeter in discrete units) improved participants' accuracy rates in similar geometric tasks (Babai et al., 2015, 2016). One should notice that the current research did not include direct comparison between numerical and continuous magnitudes. Accordingly, no conclusions regarding the interactions between area, perimeter, and numeracy in numerical tasks, such as dot arrays judgment tasks (as in Gebuis and Reynvoet, 2012a,b), can be made.

## DD and Automatic Processing of Magnitudes

When faced with unfamiliar stimuli in the practice phase, a triple group  $\times$  task  $\times$  congruency interaction was evident in error rates analysis. Participants with DD showed larger Stroop effects compared to the control group in the perimeter task but not in the area task (as in Babai et al., 2006). This pattern was also marginally significant for speed analysis. These findings indicate that individuals with deficient numerical processing also struggle with magnitude, non-numerical processing (as suggested by Leibovich et al., 2016). Moreover, group differences were significant when trying to ignore the irrelevant but salient area magnitude (Babai et al., 2006; Stavy et al., 2006; Stavy and Babai, 2008, 2010). Accordingly, the findings suggest that individuals with DD are deficient not only in the processing of numerical stimuli (e.g., Kaufmann and von Aster, 2012), but also in the processing of continuous magnitudes (in this specific case, perimeter).

Nevertheless, with practice and growing proficiency in the tasks' demands, no triple interactions were evident. The difference between perimeter and area tasks disappeared and both groups showed typical Stroop effects for both magnitudes. In other words, both magnitudes were processed automatically. Accordingly, group differences in magnitude comparison tasks may vary due to task familiarity or proficiency levels. Developmental studies regarding the numeric Stroop task indicate that one magnitude (physical size) seems to be more salient and to interfere with the other (symbolic numbers) among first graders who have no formal numerical education. With educational progress and repeated experience with numbers, both magnitudes are automatically processed and bidirectional effects are evident (Girelli et al., 2000; Rubinsten et al., 2002). Here we show a similar pattern, as task practice resulted in bidirectional effects for both groups and both magnitudes interfered with each other's judgment.

Further analysis revealed that the reduction of Stroop effects in the perimeter task was more prominent for the DD group. Accordingly, participants with DD showed automatic processing of the perimeter aspect only after continuous exposure to stimuli. Namely, they had to intentionally learn to process the perimeter aspect. The findings emphasize the need to summon intensive

exposure to magnitude, as well as to numerical information, in order to enhance compensation of DDs' core deficits (e.g., Kaufmann et al., 2003; Wilson et al., 2006).

Beyond the earlier discussion regarding magnitude saliency, a group  $\times$  congruency interaction was evident on the speed dimension across all exposure phases. Participants with DD systematically struggled to inhibit irrelevant magnitude information, manifested by larger Stroop effects, regardless of task type – area or perimeter. This pattern is compatible with inhibition deficits often associated with DD (Wang et al., 2012; Szucs et al., 2013a; Bugden and Ansari, 2015). As suggested by Aubin et al. (2016), individuals with DD might fail to consolidate magnitude knowledge and to produce the shift from the “slow” computing neural network to the “fast” retrieval network. Hence, they must invest more cognitive efforts in order to ignore irrelevant magnitude information. According to the magnitude mechanism hypothesis, a separate and more precise representation of each magnitude dimension occurs across development and becomes stronger with formal education (Newcombe, 2014). In line with this theory, people with DD struggle to effectively process different magnitude aspects of the stimulus in order to extract the proper magnitude on one hand and ignore the irrelevant magnitude on the other (Gebuis and Gevers, 2011). Hence, individuals with DD may show a developmental gap (Rubinsten and Henik, 2005), demonstrated by a difficulty in differentiating between magnitudes. Based on the current findings, we cannot specify whether individuals with DD have a deficient, domain-general inhibition mechanism (Soltész et al., 2007), or whether these inhibition deficits are exclusive to magnitude processing (De Visscher and Noël, 2014). It is also plausible that multiple neuro-cognitive components, domain-specific and domain-general, account for DD (Rubinsten and Henik, 2009; McCaskey et al., 2017). For instance, some researchers argue that DD relates to visual-spatial deficits (Ashkenazi et al., 2013; Szucs et al., 2013a; Bugden and Ansari, 2015). The fact that the DD group performed worse than the control group on reading tasks may strengthen this argument, as the reading process is known to involve visual-spatial processing (Cohen et al., 2000; Facoetti et al., 2000).

Since assessment tools and inclusion criteria vary between studies concerning numerical difficulties (Kaufmann et al., 2013), the debate on which domain-specific and domain-general deficits underlie numerical difficulties has not yet been fully addressed. There is also a need to develop a massive body of research involving developmental research and using other continuous magnitude assessment tasks. However, it is worth noting that, based on developmental research, Leibovich and Henik (2014), Leibovich et al. (2016) suggest a developmental model of the magnitude processing system: at the first step infants process mostly continuous magnitudes (for example, Clearfield and Mix, 1999). With the developmental process, they learn about the correlation between continuous and discrete magnitudes (numbers) and are subsequently able to process both types of magnitude separately (Leibovich and Henik, 2014). Development of cognitive control, including inhibition mechanisms (Davidson et al., 2006), is crucial in order to enable differentiation between

incongruent magnitudes (Leibovich et al., 2016). The proposed model has not yet been systematically validated, and the study of possible developmental gaps that might shed light on how numerical difficulties arise and develop, has not been established. However, the current study implies that DD deficits indeed relate to a failure to differentiate between magnitudes and to inhibit the irrelevant magnitude.

## Other Visual Properties

The uniqueness of the current Stroop-like task enabled us to look at the automaticity of framed vs. filled figures separately. Results show that participants showed larger Stroop effects in framed vs. filled trials, regardless of trial type (area or perimeter) and group (DD or control). These findings are not consistent with previous ones suggesting that framed figures emphasize the perimeter aspect and thus help ignore irrelevant area information (Stavy and Babai, 2008). Magnitudes tend to become confounded in a natural environment (Gebuis and Reynvoet, 2012a,b). Accordingly, it is possible that filled figures, occupying as they do larger magnitudes than framed figures, increased area and perimeter's natural confound.

Additional visual variables were examined in order to eliminate possible alternative explanations of the results. While no full control of other visual features is possible (Gebuis and Reynvoet, 2012a,b), changes in these variables appear to be good indicators that participants' performance was affected by the manipulation of experimental variables (i.e., area and perimeter). No group differences were found in any of the other visual features, including protrusion and vertical or horizontal direction. Hence, we can argue that group differences arise from magnitude processing rather than from interference of irrelevant visual features.

## Limitations

One limitation of the current study is that all participants were well-educated students and hence might not be representative of the entire DD demographic. Furthermore, the relationship between magnitude and numerical processing was not directly assessed in the current research, a fact that reduces our ability to generate firm conclusions about the role of magnitude processing in numerical processing.

## CONCLUSION AND IMPLICATIONS

The current study indicates that both area and perimeter magnitudes are automatically processed and that participants

with DD find it harder to ignore irrelevant magnitude information, especially salient area information. We assume that our findings derive from an inhibition deficit related to magnitude processing (Wang et al., 2012; De Visscher and Noël, 2014). The findings are also compatible with recent theories regarding the general magnitude processing mechanism (Leibovich and Henik, 2014; Newcombe, 2014; Lourenco and Bonny, 2016; Leibovich et al., 2016). On the other hand, we show that continuous exposure to stimuli was effective and resulted in similar patterns, namely typical Stroop effects for both magnitudes, for both groups. This fact is important for the planning of future intervention programs emphasizing the vitality of massive exposure to magnitude related stimuli in order to overcome the core deficits related to DD.

One main conclusion that can be deduced from the above is the need to develop other tasks assessing non-numerical magnitude processing, such as the task presented in the current work. Investigating multiple magnitude processing is crucial for both research and educational fields. On the one hand, there is a need for better knowledge about how multiple magnitude processing develops and occurs on the neuronal and behavioral levels in order to develop adaptive behavior. Traditionally, research methods have aimed to assess numerical processing exclusively and to block interfering magnitude information. On the other hand, the current and recent works (Gebuis and Reynvoet, 2012b; Szűcs et al., 2013b; Newcombe, 2014) stress that numerical and magnitude confounding should receive more attention in order to understand intact and impaired numerical processing. Furthermore, since non-numerical magnitude processing is crucial for higher education in science and mathematics (Wai et al., 2009; Marghetis et al., 2016), it is necessary to relate to non-numerical magnitude processing in educational research and in math curricula.

## AUTHOR CONTRIBUTIONS

HE-L and OR contributed significantly to research design, data interpretation, and manuscript drafting. HE-L collected the data.

## ACKNOWLEDGMENTS

We wish to acknowledge the help of Dr. Reuven Babai and Prof. Ruth Stavy in developing the Stroop-like tasks presented in the current study and adjusting the statistical model.

## REFERENCES

- Abbott, P. (1976). *Geometry*. London: Hodder and Stoughton.
- Ashkenazi, S., and Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behav. Brain Funct.* 6:2. doi: 10.1186/1744-9081-6-2
- Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W. S., Swigart, A. G., and Menon, V. (2013). Neuropsychologia visuo - spatial working memory is an important source of domain-general vulnerability in the development of arithmetic cognition. *Neuropsychologia* 51, 2305–2317. doi: 10.1016/j.neuropsychologia.2013.06.031
- Ashkenazi, S., Rubinsten, O., and Henik, A. (2009). Attention, automaticity, and developmental dyscalculia. *Neuropsychology* 23, 535–540. doi: 10.1037/a0015347
- Aubin, S., Voelker, A. R., and Eliasmith, C. (2016). Improving with practice: a neural model of mathematical development. *Top. Cogn. Sci.* 9, 6–20. doi: 10.1111/tops.12242
- Babai, R., Levyadun, T., Stavy, R., and Tirosh, D. (2006). Intuitive rules in science and mathematics: a reaction time study. *Int. J. Math. Educ. Sci. Technol.* 37, 913–924. doi: 10.1080/00207390600794958

- Babai, R., Nattiv, L., and Stavy, R. (2016). Comparison of perimeters: improving students' performance by increasing the salience of the relevant variable. *ZDM* 48, 367–378. doi: 10.1007/s11858-016-0766-z
- Babai, R., Shalev, E., and Stavy, R. (2015). A warning intervention improves students' ability to overcome intuitive interference. *ZDM* 47, 735–745. doi: 10.1007/s11858-015-0670-y
- Bugden, S., and Ansari, D. (2015). Probing the nature of deficits in the 'Approximate Number System' in children with persistent developmental dyscalculia. *Dev. Sci.* 19, 817–833. doi: 10.1111/desc.12324
- Butterworth, B. (2005). "Developmental dyscalculia," in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (New York, NY: Psychology Press), 455–467.
- Cheng, K., and Newcombe, N. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychon. Bull. Rev.* 12, 1–23. doi: 10.3758/BF03196346
- Clearfield, M. W., and Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychol. Sci.* 10, 408–411. doi: 10.1111/1467-9280.00177
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123, 291–307. doi: 10.1093/brain/123.2.291
- Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., and Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19552–19557. doi: 10.1073/pnas.1113195108
- Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006
- De Visscher, A., and Noël, M.-P. (2014). The detrimental effect of interference in multiplication facts storing: typical development and individual differences. *J. Exp. Psychol. Gen.* 143, 2380–2400. doi: 10.1037/xge0000029
- Dehaene, S., and Brannon, E. M. (2010). Space, time, and number: a Kantian research program. *Trends Cogn. Sci.* 14, 517–519. doi: 10.1016/j.tics.2010.09.009
- Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cogn. Neuropsychol.* 20, 487–506. doi: 10.1080/02643290244000239
- Derdikman, D., and Moser, E. I. (2010). A manifold of spatial maps in the brain. *Trends Cogn. Sci.* 14, 561–569. doi: 10.1016/j.tics.2010.09.004
- Desoete, A., Stock, P., Schepens, A., Baeyens, D., and Roeyers, H. (2009). Classification, seriation, and counting in grades 1, 2, and 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *J. Psychoeduc. Assess.* 27, 252–264. doi: 10.1177/0734282908330588
- Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., and Mascetti, G. G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex* 36, 109–123. doi: 10.1016/S0010-9452(08)70840-2
- Feigenson, L., Carey, S., and Spelke, E. S. (2002). Infants' discrimination of number vs. continuous extent. *Cogn. Psychol.* 44, 33–66. doi: 10.1006/cogp.2001.0760
- Geary, D. C. (1993). Mathematical disabilities: cognitive, neuropsychological, and genetic components. *Psychol. Bull.* 114, 345–362. doi: 10.1037/0033-2909.114.2.345
- Gebuis, T., and Gevers, W. (2011). Numerosities and space; indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). *Cognition* 121, 248–52; discussion 253–255. doi: 10.1016/j.cognition.2010.09.008
- Gebuis, T., and Reynvoet, B. (2012a). Continuous visual properties explain neural responses to nonsymbolic number. *Psychophysiology* 49, 1481–1491. doi: 10.1111/j.1469-8986.2012.01461.x
- Gebuis, T., and Reynvoet, B. (2012b). The neural mechanisms underlying passive and active processing of numerosity. *NeuroImage* 70:301–307. doi: 10.1016/j.neuroimage.2012.12.048
- Girelli, L., Lucangeli, D., and Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *J. Exp. Child Psychol.* 76, 104–122. doi: 10.1006/jecp.2000.2564
- Henik, A., and Tzelgov, J. (1982). Is three greater than five: the relation between physical and semantic size in comparison tasks. *Mem. Cogn.* 10, 389–395. doi: 10.3758/BF03202431
- Hermer, L., and Spelke, E. (1994). A geometric process for spatial representation in young children. *Nature* 370, 57–59. doi: 10.1016/j.jecp.2009.05.007
- Hermer-Vazquez, L., Spelke, E., and Katsnelson, A. (1999). Source of flexibility in human cognition: dual task studies of space and language. *Cognit. Psychol.* 39, 3–36. doi: 10.1006/cogp.1998.0713
- Huang, Y., and Spelke, E. S. (2015). Core knowledge and the emergence of symbols: the case of maps. *J. Cogn. Dev.* 6, 81–96. doi: 10.1080/15248372.2013.784975
- Huber, S., Sury, D., Moeller, K., Rubinsten, O., and Nuerk, H. C. (2015). A general number-to-space mapping deficit in developmental dyscalculia. *Res. Dev. Disabil.* 43, 32–42. doi: 10.1016/j.ridd.2015.06.003
- Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., and Delazer, M. (2007). Imaging early practice effects in arithmetic. *Neuroimage* 36, 993–1003. doi: 10.1016/j.neuroimage.2007.03.051
- Kaufmann, L., Handl, P., and Thöny, B. (2003). Evaluation of a numeracy intervention program focusing on basic numerical knowledge and conceptual knowledge: a pilot study. *J. Learn. Disabil.* 36, 564–573. doi: 10.1177/00222194030360060701
- Kaufmann, L., Mazzocco, M. M., Dowker, A., von Aster, M., Göbel, S. M., Grabner, R. H., et al. (2013). Dyscalculia from a developmental and differential perspective. *Front. Psychol.* 4:516. doi: 10.3389/fpsyg.2013.00516
- Kaufmann, L., and von Aster, M. (2012). The diagnosis and management of dyscalculia. *Deutsch. Ärztebl. Int.* 109, 767–777.
- Lee, S. A., Sovrano, V. A., and Spelke, E. S. (2012). Navigation as a source of geometric knowledge: young children's use of length, angle, distance, and direction in a reorientation task. *Cognition*, 123, 144–161. doi: 10.1016/j.cognition.2011.12.015
- Lee, S. A., and Spelke, E. S. (2011). Young children reorient by computing layout geometry, not by matching images of the environment. *Psychon. Bull. Rev.* 18, 192–198. doi: 10.3758/s13423-010-0035-z
- Leibovich, T., and Henik, A. (2013). Magnitude processing in non-symbolic stimuli. *Front. Psychol.* 4:375. doi: 10.3389/fpsyg.2013.00375
- Leibovich, T., and Henik, A. (2014). Comparing performance in discrete and continuous comparison tasks. *Q. J. Exp. Psychol.* 67, 899–917. doi: 10.1080/17470218.2013.837940
- Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2016). From 'sense of number' to 'sense of magnitude' – The role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* doi: 10.1017/S0140525X16000960 [Epub ahead of print].
- Lourenco, S. F., and Bonny, J. W. (2016). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Dev. Sci.* 20:e12418. doi: 10.1111/desc.12418
- Lourenco, S. F., and Longo, M. R. (2010). General magnitude representation in human infants. *Psychol. Sci.* 21, 873–881. doi: 10.1177/0956797610370158
- Marghetis, T., Landy, D., and Goldstone, R. L. (2016). Mastering algebra retrains the visual system to perceive hierarchical structure in equations. *Cogn. Res.* 1:25. doi: 10.1186/s41235-016-0020-9
- Mazzocco, M. M., Devlin, K. T., and McKenney, S. J. (2008). Is it a fact? Timed arithmetic performance of children with mathematical learning disabilities (MLD) varies as a function of how MLD is defined. *Dev. Neuropsychol.* 33, 318–344. doi: 10.1080/87565640801982403
- McCaskey, U., von Aster, M., O'Gorman Tuura, R., and Kucian, K. (2017). Adolescents with developmental dyscalculia do not have a generalized magnitude deficit-processing of discrete and continuous magnitudes. *Front. Hum. Neurosci.* 11:102. doi: 10.3389/fnhum.2017.00102
- Menon, V., Rivera, S. M., White, C. D., Glover, G. H., and Reiss, A. L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. *Neuroimage* 12, 357–365. doi: 10.1006/nimg.2000.0613
- Mussolin, C., De Volder, A., Grandin, C., Schölgel, X., Nassogne, M.-C., and Noël, M.-P. (2010a). Neural correlates of symbolic number processing in developmental dyscalculia. *J. Cogn. Neurosci.* 22, 860–874. doi: 10.1162/jocn.2009.21237
- Mussolin, C., Mejias, S., and Noël, M. P. (2010b). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition* 115, 10–25. doi: 10.1016/j.cognition.2009.10.006
- Newcombe, N. S. (2014). The origins and development of magnitude estimation. *Ecol. Psychol.* 26, 147–157. doi: 10.1080/10407413.2014.875333
- Noël, M. P., Rouselle, L., and De Visscher, A. (2016). Both specific and general cognitive factors account for dyscalculia. *Spec. Needs Math. Educ.* 8, 35–52. doi: 10.1177/0022219414547221

- Odic, D., Libertus, M. E., Feigenson, L., and Halberda, J. (2012). Developmental change in the acuity of approximate number and area representations. *Dev. Psychol.* 49, 1103–1112. doi: 10.1037/a0029472
- Pinel, P., Piazza, M., Le Bihan, D., and Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* 41, 983–993. doi: 10.1016/S0896-6273(04)00107-2
- Price, G. R., Holloway, I., Rasanen, P., Vesterinen, M., and Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Curr. Biol.* 17, 1042–1043. doi: 10.1016/j.cub.2007.10.013
- Raven, J. C., Court, J. H., and Raven, J. (1995). *Raven, Matrices Progresivas (Escalas: CPM, SPM, APM)*. Madrid: TEA Ediciones, S.A.
- Reigosa-Crespo, V., Valdés-Sosa, M., Butterworth, B., Estévez, N., Rodríguez, M., Santos, E., et al. (2012). Basic numerical capacities and prevalence of developmental dyscalculia: the Havana Survey. *Dev. Psychol.* 48, 123–135. doi: 10.1037/a0025356
- Ren, P., Nicholls, M. E. R., Ma, Y., and Chen, L. (2011). Size matters: non-numerical magnitude affects the spatial coding of response. *PLOS ONE* 6:e23553. doi: 10.1371/journal.pone.0023553
- Rotzer, S., Kucian, K., Martin, E., von Aster, M., Klaver, P., and Loenneker, T. (2008). Optimized voxel-based morphometry in children with developmental dyscalculia. *NeuroImage* 39, 417–422. doi: 10.1016/j.neuroimage.2007.08.045
- Rubinsten, O., and Henik, A. (2005). Automatic activation of internal magnitude: a study of developmental dyscalculia. *Neuropsychology* 19, 641–648. doi: 10.1037/0894-4105.19.5.641
- Rubinsten, O., and Henik, A. (2006). Double dissociation of functions in developmental dyslexia and dyscalculia. *J. Educ. Psychol.* 98, 854–867. doi: 10.1037/0022-0663.98.4.854
- Rubinsten, O., and Henik, A. (2009). Developmental dyscalculia: heterogeneity might not mean different mechanisms. *Trends Cogn. Sci.* 13, 92–99. doi: 10.1016/j.tics.2008.11.002
- Rubinsten, O., Henik, A., Berger, A., and Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *J. Exp. Child Psychol.* 81, 74–92. doi: 10.1006/jecp.2001.2645
- Rubinsten, O., and Sury, D. (2011). Processing ordinality and quantity: the case of developmental dyscalculia. *PLOS ONE* 6:e24079. doi: 10.1371/journal.pone.0024079
- Shalev, R. S., and Gross-Tsur, V. (2001). Developmental Dyscalculia. *Pediatr. Neurol.* 24, 337–342. doi: 10.1016/S0887-8994(00)00258-7
- Skagerlund, K., and Träff, U. (2014). Development of magnitude processing in children with developmental dyscalculia: space, time, and number. *Front. Psychol.* 5:675. doi: 10.3389/fpsyg.2014.00675
- Soltész, F., Szucs, D., Dékány, J., Márkus, A., and Csépe, V. (2007). A combined event-related potential and neuropsychological investigation of developmental dyscalculia. *Neurosci. Lett.* 417, 181–186. doi: 10.1016/j.neulet.2007.02.067
- Spelke, E., Lee, S. A., and Izard, V. (2010). Beyond core knowledge: natural geometry. *Cogn. Sci.* 34, 863–884. doi: 10.1111/j.1551-6709.2010.01110.x
- Stavy, R., and Babai, R. (2008). Complexity of shapes and quantitative reasoning in geometry. *Mind Brain Educ.* 2, 170–176. doi: 10.1111/j.1751-228X.2008.00051.x
- Stavy, R., and Babai, R. (2010). Overcoming intuitive interference in mathematics: insights from behavioral, brain imaging, and intervention studies. *Int. J. Math. Educ.* 42, 621–633. doi: 10.1007/s11858-010-0251-z
- Stavy, R., Goel V., Critchley H., and Dolan, R. (2006). Intuitive interference in quantitative reasoning. *Brain Res.* 1073–1074, 383–388. doi: 10.1016/j.brainres.2005.12.011
- Stoianov, I., and Zorzi, M. (2012). Emergence of a “visual number sense” in hierarchical generative models. *Nat. Neurosci.* 15, 194–196. doi: 10.1038/nn.2996
- Szucs, D., Devine, A., Soltész, F., Nobes, A., and Gabriel, F. (2013a). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex* 49, 2674–2688. doi: 10.1016/j.cortex.2013.06.007
- Szűcs, D., Nobes, A., Devine, A., Gabriel, F. C., and Gebuis, T. (2013b). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Front. Psychol.* 4:444. doi: 10.3389/fpsyg.2013.00444
- Szűcs, D., and Goswami, U. (2013). Developmental dyscalculia: fresh perspectives. *Trends Neurosci. Educ.* 2, 33–37. doi: 10.1016/j.tine.2013.06.004
- Tibber, M. S., Greenwood, J. A., and Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *J. Vis.* 12:8. doi: 10.1167/12.6.8
- Träff, U., Olsson, L., Östergren, R., and Skagerlund, K. (2016). Heterogeneity of developmental dyscalculia: cases with different deficit profiles. *Front. Psychol.* 7:2000. doi: 10.3389/fpsyg.2016.02000
- Tzelgov, J. (1997). “Automatic but conscious: that is how we act most of the time,” in *Advances in Social Cognition*, Vol. 10. ed. R. S. Wyer, Jr (Mahwah, NJ: Erlbaum), 217–230.
- Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. doi: 10.1037/a0016127
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space, and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Wang, L. C., Tasi, H. J., and Yang, H. M. (2012). Cognitive inhibition in students with and without dyslexia and dyscalculia. *Res. Dev. Disabil.* 33, 1453–1461. doi: 10.1016/j.ridd.2012.03.019
- Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., and Dehaene, S. (2006). An open trial assessment of “The Number Race,” an adaptive computer game for remediation of dyscalculia. *Behav. Brain Funct.* 2:20.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Eidlin-Levy and Rubinsten. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.