

---

This item was submitted to [Loughborough's Research Repository](#) by the author.  
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

## Inhibition in dot comparison tasks

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1007/s11858-014-0655-2>

PUBLISHER

© Springer

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at:  
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Clayton, Sarah, and Camilla K. Gilmore. 2019. "Inhibition in Dot Comparison Tasks". figshare.  
<https://hdl.handle.net/2134/17021>.

## Inhibition in Dot Comparison Tasks

Sarah Clayton & Camilla Gilmore

Mathematics Education Centre, Loughborough University, UK

*To appear in ZDM – The International Journal of Mathematics Education*

### Author Note

Sarah Clayton, Mathematics Education Centre, Loughborough University; Camilla Gilmore, Mathematics Education Centre, Loughborough University.

This research was supported by a Royal Society Dorothy Hodgkin Fellowship.

Correspondence concerning this paper should be addressed to Sarah Clayton, Mathematics Education Centre, Loughborough University, Loughborough, LE11 3TU, United Kingdom. Email: [s.clayton2@lboro.ac.uk](mailto:s.clayton2@lboro.ac.uk)

## Abstract

Dot comparison tasks are commonly used to index an individual's Approximate Number System (ANS) acuity, but the cognitive processes involved in completing these tasks are poorly understood. Here we investigated how factors including numerosity ratio, set size and visual cues influence task performance. Forty-four children aged 7-9 years completed a dot comparison task with a range of to-be-compared numerosities. We found that as the size of the numerosities increased, with ratios held constant, accuracy decreased due to the heightened salience of incongruent visual information.

Furthermore, in trials with larger numerosities participants' accuracies were influenced more by the convex hull of the array than the average dot size. The numerosity ratio between the arrays in each trial was an important predictor for all set sizes. We argue that these findings are consistent with a 'competing processes' inhibition-based account, where accuracy scores are influenced by individual differences in both ANS acuity and inhibitory control skills.

Key words: Inhibition, Approximate Number System, numerical cognition, dot comparison, visual cues

## **Inhibition in dot comparison tasks**

### **1. Introduction**

Competency in mathematics is important in every day life for most Western adults. Mathematics is used when calculating journey times, paying for goods in a shop, and even when planning to decorate a room. Research has shown that poor attainment in school level numeracy is correlated with real-world practical difficulties such as defaulting on mortgage payments (Gerardi et al. 2013) and poor budgeting for the future (Banks and Oldfield 2007). Nonetheless, approximately 6%-14% of school-age children have persistent difficulties with mathematics despite age-appropriate achievement in other domains (Barbaresi et al. 2005). In order to provide suitable educational support for individuals with mathematics difficulties, it is essential that researchers and educators first have a detailed and accurate understanding of the skills and underlying processes that are important for learning and performing mathematics.

The skills that are already known to be involved in mathematics can be categorised into two groups: domain-general processes and domain-specific processes. It is well-established within the psychology and education literatures that formal mathematics abilities are influenced by domain-general processes such as working memory and executive functioning skills, including inhibitory control, attention, task flexibility and switching (Bull et al. 2008; Friso-van den Bos et al. 2013; Gifford and Rockliffe 2012; LeFevre et al. 2010). Domain-specific skills also considered to be central to mathematics include procedural competence, conceptual understanding, counting, number fact knowledge and Approximate Number System (ANS) acuity or 'number sense' (Baroody 2003; Jordan et al 2003; Dehaene 1997). Despite piecemeal evidence concerning the skills that underlie mathematical learning and competence, it is not yet clear how these processes interact with each other. Very little research has focused on

the relationship between general processes such as inhibitory control and domain-specific factors, such the nature of ANS representations. Specifically, there is a gap in the literature regarding the extent to which performance on basic numerical tasks can be accounted for by general processing skills. In this manuscript we explore how a domain specific skill, ANS acuity, and the domain general ability of inhibitory control influence performance on a task aiming to measure numerical processing. In the sections below we first introduce the dominant model of the ANS and describe the task most commonly used to measure it. Then we discuss what is known about the numerical and general cognitive control factors that influence performance on this task, before introducing the present study.

### **1.1 The Approximate Number System (ANS)**

The Approximate Number System (ANS) is a cognitive system for representing quantity information, which has received a great deal of attention in recent years. The ANS is a universal system present in animals, children and adults that allows the comparison, addition and subtraction of quantities without counting them (Dehaene 1997). Instead of using exact calculations, tasks that draw on the ANS are solved using approximate representations of quantity. Studies have shown that children aged just six months old can reliably recruit their ANS to discriminate between large sets of items that differ by a ratio of 0.5, for example 8 vs. 16 dots (Xu and Spelke 2000). Similarly, many non-human primates and other animals can contrast approximate quantities that differ in numerosity, so long as the ratio difference between the two sets is sufficiently different from 1 (Emmerton 1998; Hauser et al. 2003). As humans develop, their ANS representations become more precise and adults have been found to reliably discriminate between quantities differing by up to a 0.9 ratio (Pica et al. 2004).

The examples above demonstrate that ANS performance is dependent on the ratio between the quantities to be compared. The larger the ratio, the easier it is to distinguish which set contains more items. According to the standard model of the ANS (Barth et al. 2005; Dehaene 1997), when an individual sees  $n$  objects they form an ANS representation of the quantity. This representation is drawn from a normal distribution with mean  $n$  and standard deviation  $wn$ . Here  $w$ , or the 'Weber fraction', is a parameter which indexes the acuity of each individual's ANS. When asked to compare two numerosities, say  $n$  and  $m$ , it is the ratio of these two quantities and the value of  $w$  that predicts an individual's probability of success. This is because where the  $n : m$  ratio is close to one, the distributions of possible  $n$  and  $m$  representations overlap to a greater extent, and so the probability of an individual generating incorrectly ordered representations is higher. Consequently, individuals are more likely to make an error comparing, for example, 29 vs. 30 items in comparison to 20 vs. 30 items. ANS acuities (i.e. Weber fractions) vary between individuals and influence magnitude comparison performance. Those with a more precise ANS (i.e. smaller  $w$ ) generate representations closer to the actual numerosity more often and so show superior performance on tasks used to measure the ANS (Barth et al. 2005; Dehaene 1997).

It has been proposed that the ANS may play a role in formal symbolic mathematics ability. A number of studies suggest that individuals with a more precise ANS representation also tend to have more advanced formal mathematical skills (Libertus et al. 2011; Mazzocco et al. 2011; Piazza et al. 2010, 2013; see Chen and Li 2014 for a meta-analysis). Mazzocco et al. (2011) found that poor performance on non-symbolic approximation tasks distinguishes children with mathematical learning disabilities from their typically performing peers. Similarly, links between ANS acuity and mathematics achievement in the general population have been found in numerous

studies by Halberda and colleagues (see Feigenson et al. 2013, for a review). In a study of typically developing adolescents, Halberda et al. (2008) found that participants' Weber fractions at age 14 correlated with TEMA-2 (Test of Early Mathematical Ability) scores and Woodcock-Johnson Calculation Subtest scores from the previous 10 years of schooling. The same research group later replicated the link between the TEMA-3 and dot comparison task performance with children as young as three years of age (Libertus et al. 2011). Likewise, corresponding results have been found with adult participants: Libertus et al. (2012) demonstrated a correlation between adults' ANS task performance and achievement on the quantitative section of standardised college-entrance exams in the USA. Other studies have reported further consistent findings from participants across a range of age groups (Halberda et al. 2012). Following this evidence, it has been hypothesised that training on ANS tasks may be suitable as an intervention to improve mathematics achievement (DeWind and Brannon 2012; Hyde et al. 2014; Park and Brannon 2013).

In contrast to this evidence of a relationship between ANS performance and mathematics, there is conflicting evidence from a number of studies that have failed to find a correlation between the ANS and mathematics achievement (see De Smedt et al. 2013, for a review). Holloway and Ansari (2009), for example, demonstrated that performance on non-symbolic comparison tasks was statistically unrelated to children's mathematics fluency or calculation scores on the Woodcock Johnson III Test of Achievement. Other studies have found a significant relationship in children but not in adult participants (Inglis et al. 2011). Several researchers have highlighted that the link between ANS task performance and mathematics is poorly understood (e.g. Chen and Li 2014; De Smedt et al 2013; Price, Palmer, Battista and Ansari 2012). This is important because interventions are being developed on the basis of a causal relationship between

ANS and mathematics, and yet at present we do not know what underlies the observed correlations or what contributes to performance on ANS tasks.

## **1.2 Dot comparison tasks**

An individual's ANS acuity can be measured empirically with a range of tasks. These include symbolic (e.g. digit) or non-symbolic (e.g. dot) approximate comparison and addition tasks or estimation tasks, which assess the mapping between symbols and non-symbolic representations. Non-symbolic tasks may involve visual representations including dots or flashes, auditory sequences or a combination of these. The most commonly used measure of ANS acuity is a dot comparison task, involving the comparison of two non-symbolic visual arrays of dots. During this task, participants see two dot arrays and must estimate which array they believe has more dots in it and respond either by key press, verbally, or by pointing, depending on the presentation methods employed and the age of the participants. Dot arrays are typically presented for only a brief period of time in order to prohibit counting of the dots. Performance on dot comparison tasks is often indexed by accuracy, i.e. how often the participant correctly selects the more numerous array. Other measures include estimates of participants' 'internal Weber fraction' or ' $w$ ' score, the numerical ratio effect (NRE), or reaction times, though some of these measures have been shown to be less reliable (Inglis and Gilmore 2014).

One reason why it is difficult to make sense of the conflicting literature concerning the links between ANS acuity and mathematics achievement is because the cognitive processes underlying success in dot comparison tasks are poorly understood. Without a comprehensive knowledge of the cognitive skills that are involved in completing a dot comparison task, we cannot begin to make sense of correlations between performance on the task and other cognitive abilities, such as mathematics



ability. We first need to understand whether dot comparison tasks are measuring more than just ANS acuity, and if differences in dot comparison methods influence the extent to which performance on the tasks reflect ANS acuity. There are many variables that can be manipulated in dot comparison tasks including whether the comparison stimuli are presented simultaneously or sequentially, the length of time the stimuli are displayed for and the numerosities represented. Price et al. (2012) highlighted the lack of consistency between tasks used in the literature and reported that the most robust method of presentation is to display dot arrays simultaneously on screen, so as to minimize extraneous cognitive processing demands. Inglis and Gilmore (2013) reported that the longer an individual is given to process a dot array, the more precise the formation of the resultant ANS representation. This has important implications for studies that use reaction time scores as a measure of ANS performance. If a participant is allowed to respond at their own pace, there is likely to be a trade-off between accuracy and speed: those who respond faster may not perform as accurately as they would if they chose to view the array for the full length of the presentation. Therefore, fixed presentation times and the use of accuracy scores may provide a more valid measure of the precision of ANS representations.

Gebuis and Reynvoet have highlighted the importance and influence of visual characteristics on dot comparison task performance. They suggested that individuals cannot extract quantity information from a visual scene independently of its environmental visual cues (Gebuis and Gevers 2011; Gebuis and Reynvoet 2012a, b). Instead, they proposed that when a participant is faced with an individual dot comparison trial, accuracy is influenced by their ability to attend to and 'weigh up' combinations of visual cues to make their choice (Gebuis and Reynvoet 2012a, b). Gebuis and Reynvoet argue that the existence of an ANS that can extract quantity

information independently from visual cues is unlikely. As researchers use divergent methods to create their stimuli, the control of visual cues is unstandardised across studies and therefore likely to influence performance differently in each case.

Finally, there is no consensus within the literature regarding the range of numerosities that are represented in the dot arrays. A review by De Smedt et al. (2013) highlights this lack of standardisation and suggests that inconsistencies in set size could play a role in the explanation of contrasting results within the dot comparison literature. Some previous research suggests that variation in set size does not influence accuracy on comparison tasks, however this effect has yet to be systematically explored. Barth et al. (2008) investigated whether set size affected comparison task performance by comparing the results from experiments that have used different absolute set sizes. They found that set size had no impact on accuracy scores or reaction times, however the size of the sets only differed marginally between tasks (e.g. 16 to 56 numerosity range compared with a 5 to 40 range). Another study by Barth and colleagues demonstrated similar results based on reaction time scores rather than accuracy scores (Barth et al. 2002). Revkin and colleagues found performance was higher for small sets (1- 4) compared to larger sets (10- 40) (Revkin et al. 2008). However, the authors suggested that the process of subitizing very small sets of items is characterised by different underlying mechanisms to ANS representations used for larger sets of items (Revkin et al. 2008) and so this result does not inform the question of set size effects within the ANS. Therefore the effect of set size on ANS performance remains unknown.

### **1.3 Inhibitory control and dot comparison tasks**

The variation in task procedures described above may play some role in explaining the conflicting evidence concerning the relationship between ANS performance and formal mathematics. These task variations are likely to impact on both

estimates of ANS acuity, but also, importantly, they may alter the domain-general factors that influence performance. Research has begun to emerge which suggests that inhibitory control is one such important factor (Fuhs and McNeil 2013; Cappelletti et al 2014; Gilmore et al. 2013; Nys and Content 2012).

Inhibitory control can be defined as an executive function mechanism that facilitates the suppression of prepotent responses in favour of efficient task processing (Dempster 1992). Inhibition involves the ability to focus on task relevant stimuli whilst resisting strong or automatic interference from task irrelevant information. Inhibition can be conceptually distinguished into many categories; in a review of constructs and related measurement paradigms Nigg (2000) specified four types of processing: interference control, cognitive inhibition, behavioural inhibition and oculomotor inhibition. The first type of inhibition— interference control— refers to the ability to maintain the execution of a primary motor response in the presence of distracting or competing stimuli pulling for a different response. This is classified as distinct from the ability to suppress intrusive, non-pertinent thoughts (cognitive inhibition), the ability to suppress prepotent behavioural responses (behavioural inhibition), and the effortful suppression of reflexive saccade (oculomotor inhibition). During interference control, competing stimuli draws attention away from the target response and interferes with the current operations of working memory, in turn slowing the primary motor response. A classic example of this type of inhibition is the Stroop task (Stroop 1935). A standard Stroop task involves stimuli comprising of rows of colour words, e.g. red/ blue/ black, written in different colour inks. Trials can be classified as either congruent or incongruent. Congruent trials are those in which the colour word matches the ink colour, for example the word blue written in blue ink. Incongruent trials are those in which the colour word differs from the ink colour, for example the word blue written in

red ink. The participant's task is to name aloud the ink colours written on the page without interference from the written word. Consistently it has been found that participants are slower and less accurate when completing incongruent trials in comparison to congruent Stroop task trials (MacLeod 1991).

It has been proposed that inhibition ability, specifically interference control skills, play an important role in dot comparison performance as a result of the way dot stimuli are created (Szűcs et al. 2013). In order to ensure that participants solve dot comparison tasks on the basis of the numerosity of the arrays, rather than visual characteristics such as dot size or convex hull, dot comparison tasks typically consist of both congruent and incongruent trials. On congruent trials visual cues such as the average dot size and convex hull of the array are positively correlated with numerosity i.e. the array with more dots is made up of larger dots and covers a greater area. Conversely, on incongruent trials average dot size and the convex hull of the array are negatively correlated with numerosity i.e. the array with fewer dots is made up of larger dots and covers a greater area. Some researchers have proposed that the congruency categories of dot comparison task trials are comparable to the different congruency categories present in Stroop task trials (Gilmore et al. 2013; Nys and Content 2012; Szűcs et al. 2013). Gilmore et al. suggested that for a participant to respond accurately to an incongruent dot comparison task trial they must inhibit the irrelevant and misleading visual information, such as dot size and convex hull, and respond solely based on numerosity estimations. There is a wealth of evidence to show that participants perform significantly slower and less accurately on trials where the continuous visual variables are not predictive of numerosity (Barth et al. 2006; Gebuis et al. 2009; Gilmore et al. 2013; Hurewitz et al. 2006; Nys and Content 2012). However,

not all studies have found evidence of congruency effects (e.g. Gebuis and Van der Smagt 2011; Odic, Hock and Halberda 2013; Odic, Libertus, Feigenson and Halberda 2013).

Gilmore and colleagues (2013) proposed that inhibitory control may be pertinent to the debate surrounding the relationship between the ANS and formal mathematics achievement. It is well-documented within the psychology and mathematics education literatures that there is a relationship between inhibitory control and formal mathematics ability (e.g. Blair and Razza 2007; Bull and Sherif 2001; Espy et al. 2004; St Clair-Thomson and Gathercole 2006; Stavy and Babai 2010). Individuals with better inhibition skills also tend to perform better on tasks measuring mathematical ability. In line with this, Gilmore et al. (2013) found that children's formal mathematics achievement scores were only correlated with performance on incongruent dot comparison task trials, and not with congruent trials. Gilmore et al. therefore proposed that the correlation often observed between mathematics achievement and ANS acuity may in fact arise from mutual correlations with inhibitory control. Indeed Gilmore et al. (2013) also reported that 7 to 10 year olds' overall dot comparison performance scores no longer significantly predicted mathematics achievement scores once inhibition skills were accounted for. Evidence from Fuhs and McNeil (2013) also supports this proposal with findings that low SES preschool children's dot comparison task performance was no longer a borderline predictor of mathematics achievement scores once scores on an inhibition task were taken into account. Recently, Cappelletti et al (2014) have suggested that a decline in dot comparison performance in an older population may reflect deterioration of inhibitory processes rather than impoverished numerical skills. Older participants were particularly impaired on tasks that required the inhibition of visual information

incongruent to numerosity, and moreover this difficulty was correlated with poorer inhibitory control performance on a Stroop task.

Put together, these findings provide strong evidence for an alternative competing processes account of dot comparison task performance by which both the ANS and other competing processes driven by visual cues, influence accuracy on non-symbolic comparison tasks (Fuhs and McNeil 2013; Gilmore et al. 2013; Nys and Content 2012). According to this account, when a participant is faced with an individual dot comparison trial, they engage their ANS to judge the difference between the two dot arrays. However, at the same time the visual characteristics of the stimuli such as average dot size and the convex hull may interfere with this process, and in some cases participants may have to inhibit a prepotent response based on these visual characteristics in order to respond correctly. Following this account, inhibitory control may be a key mediator of the link between dot comparison task performance and formal symbolic mathematics achievement. However, although these studies have indicated that inhibitory control skills may be involved in performance on dot comparison trials in general, they have failed to identify the characteristics of dot comparison trials that will increase inhibitory control demands. In order to understand the ways in which inhibitory control is involved in solving incongruent dot comparison trials we need to understand how visual characteristics influence dot comparison task performance. This is particularly important given that different researchers have employed different versions of dot comparison tasks. Understanding the influence of visual characteristics on dot comparison task performance is the focus of the present study.

#### **1.4 The present study**

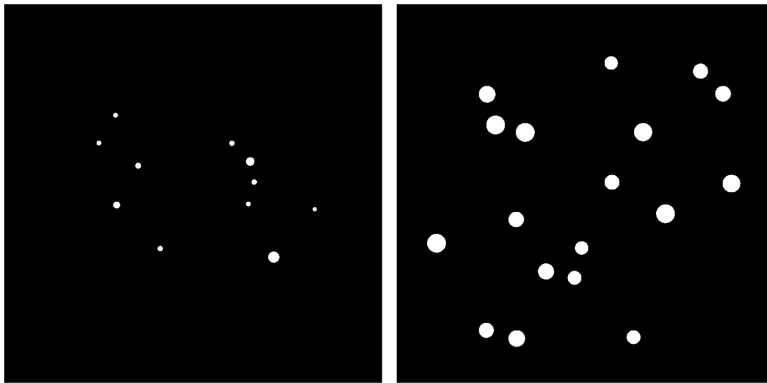
Previous research has suggested that inhibitory control may play a role in dot comparison performance due to the influence of visual characteristics on participants' performance. However, we do not yet know the ways in which the visual characteristics of dot arrays affect participants' responses. We therefore investigated the characteristics which influence congruency effects and explored whether some visual cues are harder to inhibit than others. We varied the numerical ratio between the arrays, the absolute set size of quantities and the visual characteristics of the stimuli, specifically average dot size and convex hull. We chose to focus only on these particular visual characteristics of the stimuli following previous research that demonstrated a high correlation between average dot size and other commonly reported visual characteristics, including cumulative surface area and dot density within the array (Gebuis and Reynvoet 2012a)

If both the ANS and inhibitory control skills influence performance on dot comparison tasks, as proposed by the competing processes account, then we would expect participants' accuracies to be related to all three factors: numerical ratio, visual characteristics and set size. Numerical ratio would influence performance due to the approximate nature of ANS representations. Visual characteristics and set size would influence performance by varying the inhibitory control demands of the incongruent trials. We predicted that the salience of visual cues of an array would change with absolute set size. In particular convex hull information is likely to become more salient with increasing numerosity as the density of the dots increases and creates a more prominent perimeter (see Figure 1 for an example). For an incongruent trial, this is likely to place a greater demand on inhibitory control skills to disregard the interfering information and focus on numerosity, hence we predicted that accuracy would decrease as set size increases. Moreover, we predicted that not only would convex hull

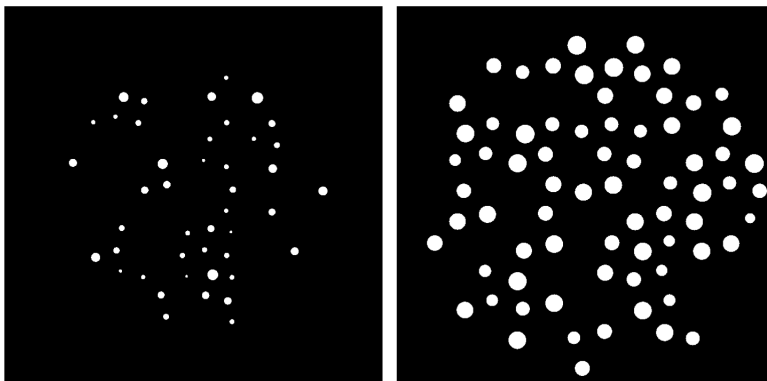
incongruent trials be more challenging than convex hull congruent trials, but that this congruency effect would also interact with set size. Participants are likely to perform less accurately on incongruent trials that are made up of larger absolute to-be-compared numerosities in comparison to smaller to-be-compared numerosities. We predicted that trials with dot size congruent visual cues would be more challenging than dot size incongruent trials, but did not make any predictions about interaction effects, as it is unclear whether dot size congruency effects would vary with increasing set size.

*Figure 1:* Examples of the same 0.61 numerosity ratio trial in a) the smallest set size (11 vs. 18 dots) and b) the largest set size (44 vs. 72 dots).

a)



b)





## 2. Method

### 2.1 Participants

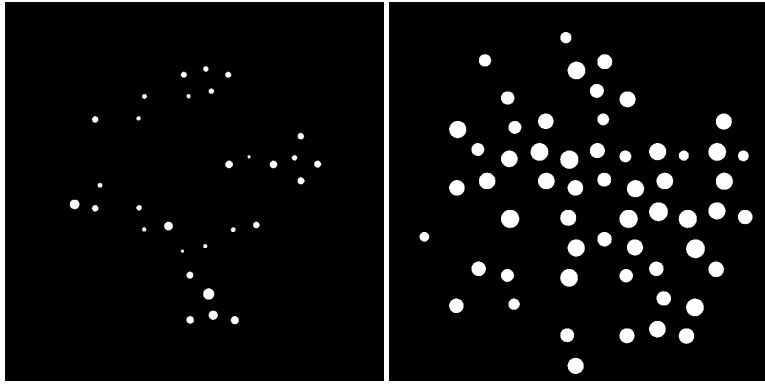
Participants were 44 children (22 male) aged between 7 and 9 years ( $M= 8.3$ ,  $SD= 0.59$  years). Seven children whose performance was not significantly above chance were excluded from the analysis<sup>1</sup>. This study was approved by the Loughborough University Ethics Approvals (Human Participants) Sub-Committee.

### 2.2 Task

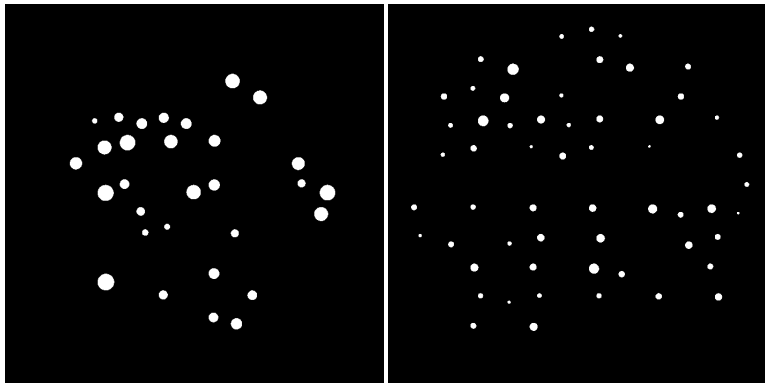
Participants completed a dot comparison task. Stimuli were arrays of white dots presented on a black background. The dots were generated using Gebuis and Reynvoet's (2011) method, which creates each pair of dot arrays four times with different visual characteristics in terms of average dot size and convex hull, resulting in four image types, shown in Figure 2. There were four set size conditions: small, medium, large and very large with the small numerosities ranging from 10 to 19, and the others being 2, 3, and 4 times as large respectively. The ratios between the numerosities displayed in the arrays were 0.53, 0.61, 0.71, 0.81, 0.88 and 0.93. The stimuli were presented simultaneously, side-by-side on a 15" laptop display.

*Figure 2:* Examples of the four image types used in a 30 vs. 57 dot comparison task trial. For each type of trial, the more numerous arrays contained a) larger dots and a larger convex hull (fully congruent) b) smaller dots and larger convex hull (dot incongruent, convex hull congruent) c) larger dots and a smaller convex hull (dot congruent, convex hull incongruent) d) smaller dots and a smaller convex hull (fully incongruent).

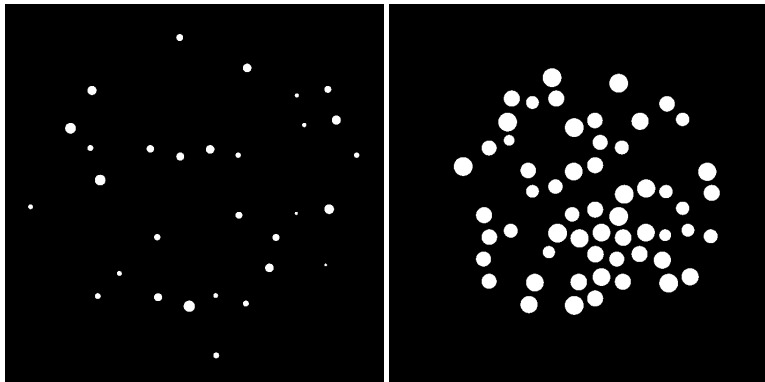
a)



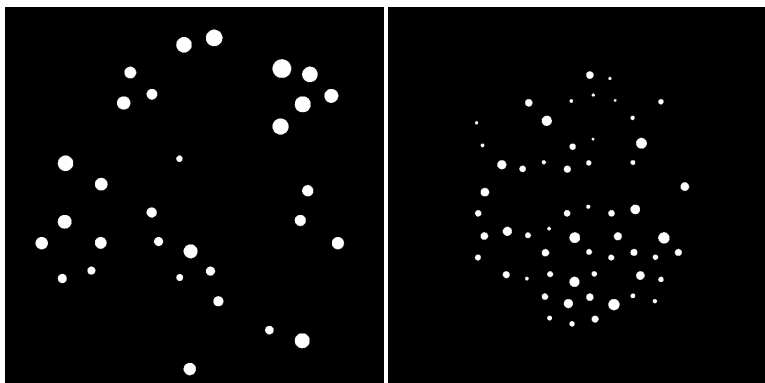
b)



c)



d)



### 2.3 Procedure

Children were tested individually in a quiet room at their school. Each trial began with a fixation point (600ms) followed by the presentation of the two arrays (600ms) and finally a screen with a question mark, which was displayed until the participant responded. Participants were asked to indicate which array was more numerous using left and right keys marked on the keyboard. There were eight practice trials and 184 experimental trials.

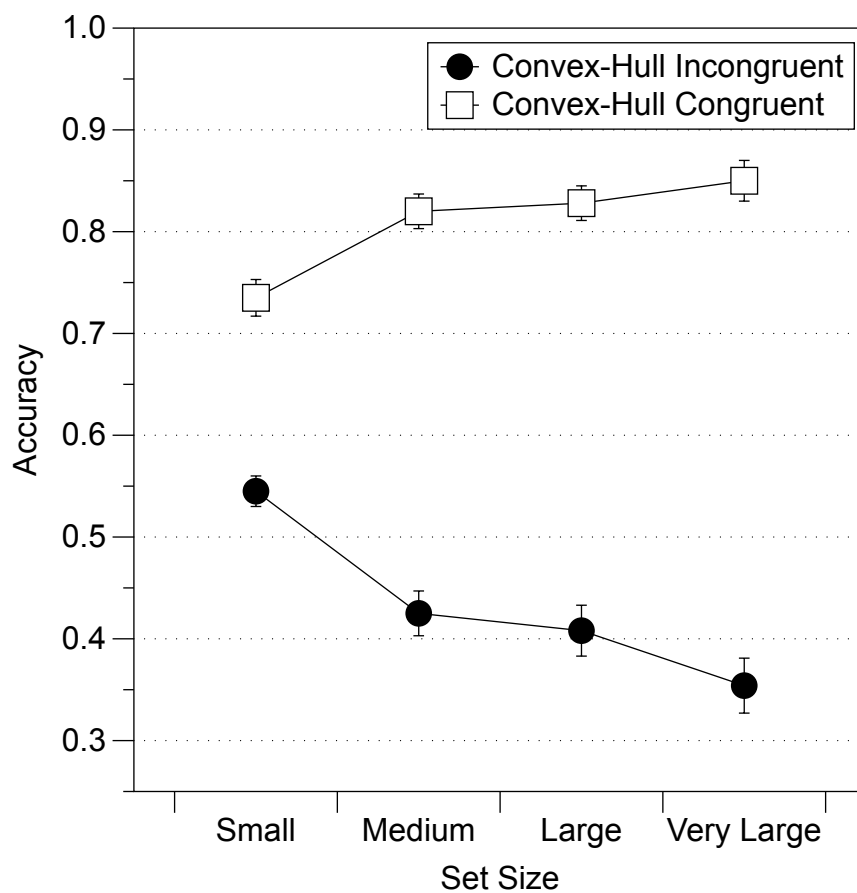
### 3. Results

First the effects of set size and congruency status of the images on performance were explored. Participants' accuracy data were subjected to a 4 (set size: small, medium, large, very large)  $\times$  2 (convex hull: congruent, incongruent)  $\times$  2 (average dot size: congruent, incongruent) within-subjects Analysis of Variance (ANOVA). As predicted, accuracy scores were highest in the small set size condition,  $M= 64.0\%$ , and declined with increasing numerosity,  $M_s= 62.2\%$ ,  $61.8\%$  and  $60.2\%$  for the medium, large and very large conditions respectively. This represented a significant linear trend,  $F(1, 36) = 6.6, p = .014, \eta_p^2 = .16$ .

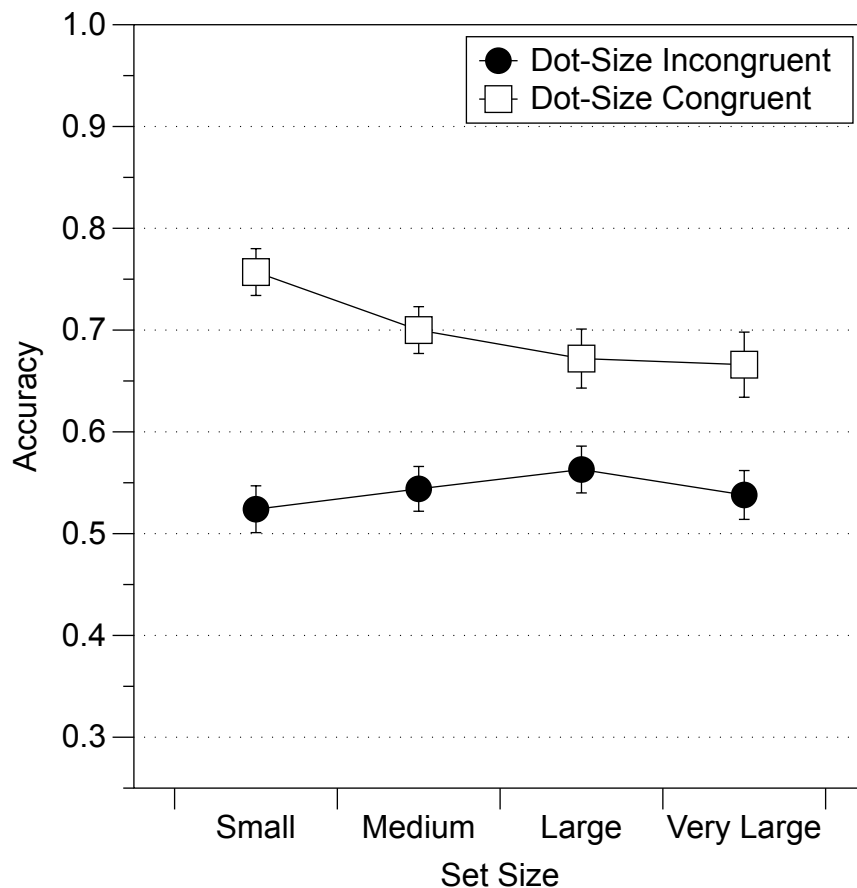
Accuracy was significantly higher for trials that were congruent in terms of convex hull ( $M= 80.8\%$ ) compared to incongruent trials ( $M= 43.3\%$ ),  $F(1, 36) = 158.18, p < .001, \eta_p^2 = .82$ . Similarly, accuracy was significantly higher for trials that were congruent in terms of dot size ( $M= 69.8\%$ ) than incongruent trials ( $M= 54.2\%$ ),  $F(1, 36) = 14.43, p < .001, \eta_p^2 = .29$ . Set size significantly interacted with both convex-hull congruency,  $F(3, 108) = 37.18, p < .001, \eta_p^2 = .51$  and dot-size congruency,  $F(3, 108) = 5.92, p < .001, \eta_p^2 = .14$ . As set size increased, the effect of convex-hull congruency

increased, shown in Figure 3, and the effect of dot-size congruency decreased, shown in Figure 4. Notably, performance drops significantly below chance on the convex-hull incongruent trials of medium,  $t(36) = -3.37, p = .002$ , large,  $t(36) = -3.67, p = .001$ , and very large set sizes,  $t(36) = -5.36, p < .001$ .

*Figure 3:* Accuracy scores for convex-hull congruent and incongruent trials with small, medium, large and very large set sizes.



*Figure 4:* Accuracy scores for dot-size congruent and incongruent trials with small, medium, large and very large set sizes.



In addition to this analysis, we conducted a series of binary logistic regressions to investigate more sensitively how the ratio between cues (numerosity, convex hull, dot size) on each trial affected individuals' performance, and how this differed between smaller and larger set size trials. We obtained values of the average dot size and the total convex hull of the arrays for each image and calculated the ratio for each trial. Then, for each individual, we conducted a binary logistic regression predicting trial response using the ratios between the trial's two numerosities, the two convex hulls, and the two mean dot sizes. This yielded odds ratios for convex-hull size, dot size and numerosity. These odds ratios were calculated for the full set of trials and then separately for trials that included smaller set sizes (small and medium set size groups) and larger set sizes (large and very large groups). We compared differences between

the odds ratios derived from these different set size groups using Wilcoxon Signed-ranks tests.

An odds ratio significantly greater than 1 indicates that the given predictor has had an effect on the individual's comparison performance. Overall, the odds ratios for numerosity ratio were higher than 1 for every participant, suggesting that all the participants used numerosity information to some extent to complete the task. A Wilcoxon signed-ranks test showed no difference between odds ratios for the numerosity ratio for smaller ( $Mdn = 2.99$ ) and larger ( $Mdn = 3.33$ ) set sizes,  $Z = .309$ ,  $p = .757$ , suggesting that participants focused on numerosity irrespective of the set size.

The odds ratios for the convex-hull ratio did, however, vary by set size. A Wilcoxon Signed-ranks test showed that odds ratios for convex-hull ratio were significantly lower,  $Mdn = 2.69$ , for smaller compared to larger,  $Mdn = 9.40$ , set sizes,  $Z = 5.21$ ,  $p < .001$ . For dot-size ratio, there was no difference between odds ratios for smaller,  $Mdn = 1.17$ , and larger,  $Mdn = 1.20$ , set size trials,  $Z = 1.74$ ,  $p = .083$ . This pattern of odds ratios is consistent with the picture that emerged from the ANOVA analysis. In particular, the findings suggest that the impact of visual characteristics of dot arrays varies with set size.

## 4. Discussion

### 4.1 Summary of Main Findings

Our goal in this study was to investigate how performance on a dot comparison task, a basic non-symbolic task designed to measure the ANS, was influenced by individual differences in numerosity judgement acuity but also by more wide-ranging domain-general cognitive skills. Specifically, the research examined whether dot comparison accuracy scores reflect participant's ANS acuity alone, indicated by significant effects of numerosity ratio, or whether visual characteristics of the arrays

also impacted on performance, which would suggest that inhibitory control is required to solve incongruent trials. To do this we manipulated the numerosity ratios, the visual characteristics and the absolute set sizes of the to-be-compared dot arrays, which in turn further affected the visual characteristics of the arrays.

As predicted, we found that as numerosities increased, accuracy scores decreased. Furthermore, we found that the types of visual cues that participants most attended to varied by set size. Specifically, as set size increased, it seems that participants are relatively more often influenced by the convex hull of the arrays, and less by the mean dot size. To elaborate, in larger set size conditions participants performed better on trials where convex hull size was predictive of numerosity and less accurately when it was incongruent with numerosity. In fact, performance was significantly below chance on the medium, large and very large set size convex-hull incongruent trials, suggesting that participants found it particularly difficult to ignore convex-hull cues on these trials. In these larger set size conditions average dot size had little influence on performance. In contrast, in smaller set size conditions, participants' performance was higher on trials where average dot size was predictive of numerosity and lower when it was incongruent with numerosity. In these smaller set size conditions, accuracy was less influenced by convex hull size. This shows that participants focused on different visual cues depending on absolute set size of the dot array. Importantly, regression analyses also showed that accuracy scores were not only influenced by visual cue processing, but also by numerosity processing.

#### ***4.2 Theoretical Implications***

The standard model of the ANS (Dehaene 1997) struggles to account for these findings. The standard model proposes that the participants' ANS precision and the ratio between the numerosities in each trial should be the only two predictors of

accuracy in dot comparison tasks (Barth et al. 2005; Dehaene 1997). Although we found that all participants focus on numerosity to some extent to complete the task, we found that set size and visual cues, such as dot size and convex hull, also interfere with task performance. This suggests that task success depends on more than just ANS processing.

Our results support the view of Gebuis and colleagues who argue that the visual characteristics of dot comparison tasks are of pivotal importance to performance on dot comparison tasks (Gebuis and Gevers 2011; Gebuis and Reynvoet 2012a, b). Both of the visual characteristics manipulated in this study influenced accuracy scores. Notably, the congruency status of the convex hull in the dot arrays had a stronger influence on performance than average dot size, as reflected by the larger overall congruency effects reported in the ANOVA. Correspondingly, as set size varied, the influence of convex hull on participants' accuracy scores varied significantly, however despite the significant ANOVA interaction, the odds ratio analyses showed the influence of average dot size to be less prominent. Gebuis and colleagues suggest that judgements are a function of weighing up multiple visual cues simultaneously, and that it is unlikely that ANS representations of numerosity can be extracted independently of this. Our results suggest that ANS representations are in fact employed alongside visual judgements, although as this study could not test how all the visual parameters were simultaneously used to judge numerosity, this is a hypothesis that warrants further investigation.

Our findings provide support the competing processes account of dot comparison task performance (Fuhs and McNeil 2013; Gilmore et al. 2013; Nys and Content 2012). Gilmore et al. (2013) suggested that the successful completion of a dot comparison task relies partially on ANS abilities and partially on the ability to inhibit salient visual features of the array. Our results are consistent with the suggestion that



participants do use their ANS to perform dot comparison tasks, and show that the congruency of the visual stimuli, in terms of dot size and convex hull, also influences performance. Participants found trials where it was necessary to inhibit the incongruent visual characteristics of an array significantly more difficult than congruent trials. We do not suggest that ANS processing itself is influenced by interfering processes, but rather that the interference competes with ANS processing, and consequently influences dot comparison performance.

From this evidence, it seems likely that dot comparison tasks may also measure inhibitory control skills, specifically individual differences in interference control. Participants with better inhibitory control skills are therefore likely to perform better on dot comparison task trials that contain incongruent visual cue information. This finding is consistent with previous research that has found a correlation between performance on dot comparison trials where visual information was incongruent to numerosity, and inhibitory control tasks (Cappelletti et al. 2014; Gilmore et al. 2013). Following this, further research should focus on to what extent performance on dot comparison tasks can be accounted for by inhibitory control skills rather than individual differences in ANS acuity.

The support this study provides for the competing processes account has implications for the idea that there is a causal relationship between ANS acuity and formal mathematics achievement (Libertus et al. 2011; Mazzocco et al. 2011; Piazza et al. 2010; Piazza et al. 2013). The competing processes account instead proposes that inhibition could be the key mediator of this link (Fuhs and McNeil 2013; Gilmore et al. 2013). If this is the case, then developing successful interventions to improve participants' formal mathematical skills through dot comparison training may be unachievable. It will be advantageous for future research to develop a more extensive

and connected picture of the cognitive processes that underlie performance on dot comparison tasks before the task's educational applications can be appropriately considered.

### **4.3 Methodological Implications**

Our findings underscore the significance of the many procedural differences in dot comparison task protocols present within the literature. Currently, there is no consensus concerning the range of numerosities included in comparison tasks and many studies that are cited and reviewed in the literature involve diverse ranges of numerosities. Similarly, there are no established recommendations on how visual characteristics should be controlled in the stimuli. Many researchers only control for the cumulative surface area of the dots (e.g. Halberda et al. 2008, 2012; Libertus et al. 2011; Mazzocco et al. 2011) and so fail to investigate how performance is influenced by other important visual cues such as convex hull. As shown by the present results, variations in set size, convex hull and average dot size all influence accuracy scores and so should be considered carefully when designing, analysing and comparing non-symbolic dot comparison experiments.

In order to advance our understanding of the ANS and its correlates, future research may benefit from a shift in focus towards the development of alternative protocols to measure ANS acuity. Due to the unwanted influence of visual cues, the use of dot comparison tasks may not be appropriate. Cross-modal methods involving a mixture of visual dot arrays and auditory stimuli have successfully been used in previous research. Barth et al. (2005) found that children were able to integrate quantity information from these two different modalities, demonstrating that performance in a dual-modality task was not significantly different to performance in a single visual modality task. It is possible this method of non-symbolic comparison may

require less inhibitory control demands than standard dot comparison tasks, and therefore be a more valid measure of ANS acuity.

#### **4.4 Conclusion**

We have demonstrated that children's performance on a dot comparison task designed to measure the ANS was not only influenced by individual differences in ANS acuity, but also by the size of the numerosities involved and the visual characteristics of the stimuli. Results follow a pattern consistent with the hypothesis that inhibitory control may have been recruited to account for visual cues that were incongruent with numerosity information. This finding strengthens evidence for the crucial role of inhibition in dot comparison tasks. Accordingly, it will be important for future studies measuring the relationship between dot comparison task performance and mathematics achievement to control for individual differences in inhibitory control.

## Footnote

1 The same pattern of results were found when these participants were included. However, because we were not certain that these participants understood the task instructions we chose to exclude them.

## References

- Banks, J., & Oldfield, Z. (2007). Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving. *Fiscal Studies*, *28* (2), 143–170.
- Barbarese, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: Incidence in a population-based birth cohort, 1976–82, Rochester, Minn. *Ambulatory Pediatrics*, *5*(5), 281–289.
- Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise studies* (pp. 1–33). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Barth, H., Beckmann, L., & Spelke, E. S. (2008). Nonsymbolic, approximate arithmetic in children: Abstract addition prior to instruction. *Developmental Psychology*, *44*(5), 1466–1477.
- Barth, H., Kanwisher, N., & Spelke, E. S. (2002). The construction of large number representations in adults. *Cognition*, *86*, 201–221.
- Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222.
- Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic

- in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 14116–14121.
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663.
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental neuropsychology*, *33*(3), 205–228.
- Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching and working memory. *Developmental Neuropsychology*, *19*(3), 273-293.
- Cappelletti, M., Didino, D., Stoianov, I., & Zorzi, M. (2014). Number skills are maintained in healthy ageing. *Cognitive Psychology*, *69*, 25-45.
- Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, *148*, 163-172.
- Dehaene, S. (1997) *The number sense*. Oxford: Oxford University Press.
- Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental review*, *12*(1), 45–75.
- De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*. *2*, 48-55.
- DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: effects of feedback and training. *Frontiers in Human Neuroscience*, *6*(68), 1-10.

- Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning & Behavior*, 26(3), 243–256.
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, 26(1), 465–486.
- Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Development Perspectives*, 7(2), 74–79.
- Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, 10, 29–44.
- Fuhs, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low - income homes: contributions of inhibitory control. *Developmental Science*, 16, 136–148.
- Gebuis, T., & Gevers, W. (2011). Numerosities and space; indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). *Cognition*, 121(2), 248– 252.
- Gebuis, T., Kadosh, R. C., de Haan, E., & Henik, A. (2009). Automatic quantity processing in 5-year olds and adults. *Cognitive Processing*, 10(2), 133–142.
- Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods*, 43, 981–986.
- Gebuis, T., & Reynvoet, B. (2012a). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, 141(4), 642–648.
- Gebuis, T., & Reynvoet, B. (2012b). The Role of Visual Information in Numerosity Estimation. *PLoS ONE*, 7(5), e37426.

- Gebuis, T., & van der Smagt, M. J. (2011). False approximations of the approximate number system. *PLoS ONE*, *6*(10), e25405.
- Gerardi, K., Goette, L., & Meier, S. (2013). Numerical ability predicts mortgage default. *Proceedings of the National Academy of Sciences*, *110*(28), 11267–11271.
- Gifford, S., & Rockliffe, F. (2012). Mathematics difficulties: Does one approach fit all? *Research in Mathematics Education*, *14*(1), 1–15.
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS ONE* *8*(6): e67374.
- Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116– 11120.
- Halberda, J., Mazocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668.
- Hauser, M. D., Tsao, F., Garcia, P., & Spelke, E. S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton-top tamarins. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1523), 1441–1446.
- Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29.
- Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*(51), 19599–19604.
- Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number

- practice enhances subsequent exact symbolic arithmetic in children. *Cognition*, 131(1), 92-107.
- Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review*, 18(6), 1222–1229.
- Inglis, M., & Gilmore, C. (2013). Sampling from the mental number line: How are Approximate Number System representations formed? *Cognition*, 129, 63-69.
- Inglis, M., & Gilmore, C. (2014). Indexing the Approximate Number System. *Acta Psychologica*, 145, 147-155.
- Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development*, 74(3), 834–850.
- LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith Chant, B. L., Bisanz, J., Kamawar, D., & Penner Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, 81(6), 1753– 1767.
- Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, 14, 1292–1300.
- Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica*, 141(3), 373–379.
- MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin*, 109(2), 163–203.
- Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the



- approximate number system predicts later school mathematics performance. *PLoS ONE*, 6(9): e23749.
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological bulletin*, 126(2), 220.
- Nys, J., & Content, A. (2012). Judgement of discrete and continuous quantity in adults: Number counts! *The Quarterly Journal of Experimental Psychology*, 65(4), 675–690.
- Odic, D., Hock, H., & Halberda, J. (2013) Hysteresis affects number discrimination in young children. *Journal of Experimental Psychology: General*. 143(1), 255-265.
- Odic, D., Libertus, M., Feigenson, L., & Halberda, J. (2013) Developmental change in the acuity of approximating area and approximating number. *Developmental Psychology*, 49, 1103-1112.
- Park, J., & Brannon, E. M. (2013). The approximate number system improves math proficiency. *Psychological Science*, 24(10), 2013-2019.
- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., ... Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, 116, 33–41.
- Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal Approximate Number System. *Psychological Science*, 24, 1037-1043.
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 306(5695), 499–503.
- Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and

outcome measures, and their relationship to arithmetic achievement in adults.

*Acta Psychologica*, 140(1), 50–57.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, 19(6), 607–614.

St Clair-Thomson, H. L., Gathercole, S. E. (2006). Executive functions and achievement in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, 59(4), 745-759.

Stavy, R., & Babai, R. (2010). Overcoming intuitive interference in mathematics: insights from behavioral, brain imaging and intervention studies. *ZDM Mathematics Education*, 42(6), 621-633.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General*, 18(6), 643–662.

Szűcs, D., Nobes, A., Devine, A., Gabriel, F., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of Approximate Number System acuity and comparative effects between adults and children. *Frontiers in Psychology*, 4(444).

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74(1), B1–B11.