Word counts:

Abstract: 179

Main text: 13403

References: 5422

Total: 20029

Title:

Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review

Titia Gebuis*; Department of Molecular and Cellular Neurobiology, Center for Neurogenomics and Cognitive Research, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands; e-mail: T.Gebuis@vu.nl; phone number: +31 20 5987111

Roi Cohen Kadosh; Department of experimental psychology, South Parks Road, Oxford, OX1 3UD, United Kingdom; e-mail: Roi.CohenKadosh@psy.ox.ac.uk; phone number: +44 1865 271385; fax number +44 1865 310447

Wim Gevers; Center for Research in Cognition and Neurosciences (CRCN), Université Libre de Bruxelles and UNI – ULB Neurosciences institute, Brussels, Belgium, CP122, avenue F.D. Roosevelt 50, 1050 Bruxelles; e-mail: Wim.Gevers@ulb.ac.be; phone number +32 26 504228

*Corresponding author

ABSTRACT

It is widely accepted that human and nonhuman species possess a specialized system to process large approximate numerosities. The theory of an evolutionarily ancient *approximate number system* (ANS) has received converging support from developmental studies, comparative experiments, neuroimaging, and computational modelling, and it is one of the most dominant and influential theories in numerical cognition. The existence of an ANS system is significant, as it is believed to be the building block of numerical development in general. The acuity of the ANS is related to future arithmetic achievements, and intervention strategies therefore aim to improve the ANS. Here we critically review current evidence supporting the existence of an ANS. We show that important shortcomings and confounds exist in the empirical studies on human and non-human animals as well as the logic used to build computational models that support the ANS theory. We conclude that rather than taking the ANS theory for granted, a more comprehensive explanation might be provided by a sensory-integration system that compares or estimates large approximate numerosities by integrating the different sensory cues comprising number stimuli.

KEYWORDS: Approximate number system, cognitive development, evolution, sensory integration, brain, numerosity, cognitive control

1. INTRODUCTION

The human brain is equipped with a symbolic number system that allows humans to strive in their environment. From computers to shopping, from sports events to trading, symbolic numbers play a fundamental role for these and many other activities. To understand the concept of symbolic number, it is argued that humans rely on two different innate *non-symbolic* number systems (Cantlon & Brannon, 2006; but see: Meck & Church, 1983; Rugani, Vallortigara, & Regolin, 2014). Both systems are suggested to be present across different human societies as well as other animal species (e.g. Cantlon, 2012; Feigenson, Dehaene, & Spelke, 2004; Gomez-Laplaza & Gerlai, 2013). One of these systems represents large numbers of objects (larger than 4 or 5 items), whereas the other system serves the exact representation of a small number of objects (smaller than 4 or 5 items) (Cantlon, Platt, & Brannon, 2009; Feigenson, Dehaene, & Spelke, 2004). In the former case, the number of objects can only be approximated while in the latter case a rapid and more accurate estimate can be made called subitizing. In the present work, we will focus on the system that allows the representation of large numbers of objects, the so-called *approximate number system* (ANS), as it is considered the foundation of more complex mathematics. From here on, we will use the term *numerosity* to describe large quantities (>5) presented in a non-symbolic manner, such as visual arrays of dots or acoustic beeps.

The ANS system is often described as a logarithmically compressed number line, which represents numerosities as partially overlapping Gaussian curves (Dehaene, 2003; but see also Cantlon, Cordes, Libertus, & Brannon, 2009; Gallistel & Gelman, 2000). This overlap increases when numerosities closer to each other are selected. Studies in humans and nonhuman animal species support this idea (e.g. Agrillo, Piffer, & Bisazza, 2010; Agrillo, Piffer, Bisazza, & Butterworth, 2012; Gebuis & Van der

Smagt, 2011; Gilmore, Attridge, & Inglis, 2011; Halberda & Feigenson, 2008; Libertus, Woldorff, & Brannon, 2007; Xu & Spelke, 2000). Furthermore, they show that this overlap in activation pattern between neighbouring numbers also increases with increasing numerosity (see Figure 1). The ability to discriminate between two numerosities is therefore best described by the relative difference (i.e. ratio) between both numerosities (Piazza et al., 2010). On the basis of several behavioral and imaging studies it was suggested that this ratio dependence is the core mechanism of the ANS system (Cantlon, Brannon, Carter, & Pelphrey, 2006; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Roggeman, Santens, Fias, & Verguts, 2011).

The ANS is believed to be important for elementary numerosity processes and it is assumed to be the building block for our numerical development in general (Butterworth, 2005; Dehaene, 2009; Piazza, 2010), as well as more advanced mathematics in particular (Halberda & Feigenson, 2008). Several lines of evidence exist for these claims. First, the ANS increases in precision with increasing age until young adulthood (Brannon, Suanda, & Libertus, 2007; Halberda & Feigenson, 2008; Wood & Spelke, 2005; Xu, Spelke, & Goddard, 2005). Second, this increase as a function of age is impaired in children with arithmetic deficiencies (Piazza et al., 2010). Third, several studies have shown a relationship between the ANS and more complex arithmetic skills (e.g. Halberda, Mazzocco, & Feigenson, 2008; Inglis, Attridge, Batchelor, & Gilmore, 2011; Mazzocco, Feigenson, & Halberda, 2011; Park & Brannon, 2013; M. Piazza, Pica, Izard, Spelke, & Dehaene, 2013; A. Starr, M. E. Libertus, & E. M. Brannon, 2013). For instance, in a recent study, the acuity of the ANS was shown to be a stable predictor for math ability in children between 7 and 9 years of age (Inglis et al., 2011). Even after controlling for age and nonverbal IQ the relationship between the ANS and math achievement persisted. This means that the performance on arithmetic tasks such as addition and subtraction improves when the ANS becomes more accurate. These results have a significant impact on the field of education and intervention. For example, intervention studies try to improve ANS acuity and thus performance in arithmetic in general (Räsänen, Wilson, Aunio, & Dehaene, 2009; Wilson, Dehaene, et al., 2006; Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006).

Recently, Gebuis and Reynvoet (2015) noted that the relationship between non-symbolic number representation and arithmetic might be spurious. Indeed, upon closer inspection of the data, contradictory results become apparent. Some studies do find a relationship between non-symbolic numerosity performance and arithmetic whereas others do not (De Smedt & Gilmore, 2011; Gilmore et al., 2013; Holloway & Ansari, 2009; Rousselle & Noel, 2007; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Soltesz, Szucs, & Szucs, 2010). These contradictory results could partly be attributed to methodological differences across studies (e.g., range in numerosities included or the measure for ratio dependence) (see for further methodological explanations: Feigenson, Libertus, & Halberda, 2013; Guillaume, Gevers, & Content, 2015; Price, Palmer, Battista, & Ansari, 2012; Smets, Gebuis, Defever, & Reynvoet, 2014; Smets, Gebuis, & Reynvoet, 2013).

In this review we propose that a more complete explanation of the contradictory results in the literature might be provided by the sensory properties that comprise the numerosity stimuli. We define *sensory properties* as each non-numerical cue that comprises numerosity stimuli, examples in the visual domain are: density, circumference, surface, diameter and convex hull. We propose that all sensory properties present in the stimulus contribute to our ability to compare the numerosity of two sets of objects. This proposal is based on the notion that sensory properties are confounded with numerosity.

In numerosity studies, sensory properties are typically manipulated to exclude the participant's reliance on these confounding sensory properties. For example: 10 small dots spread over a large area have to be compared to 5 large dots spread over an equally large area. Obviously, when an increasing number of sensory properties are manipulated, the more difficult it becomes for the participant to rely on the sensory cues to judge numerosity. It is therefore assumed that in such cases, participants do not rely on sensory cues but instead on numerosity. The data however suggests otherwise: sensory effects remain visible even when sensory controls are used. This is problematic since the reliance on sensory cues can differ between studies due to differences in methods used to control the sensory cues. Consequently, differences in performance can be unrelated to the task at hand. This difference in performance induced by the sensory cues could explain previously mentioned conflicting results (Clayton, Gilmore, & Inglis, 2015; Gebuis & Reynvoet, 2012b). Such a role for the sensory cues in numerosity judgments contradicts the central tenet of the ANS theory, which proposes that the confounding sensory properties are first normalized (i.e. removed) to allow in the next stage, the extraction of numerosity independent of these previously confounding sensory properties (Dehaene, 1997; Stoianov & Zorzi, 2012; Verguts, Fias, & Stevens, 2005; Whalen, Gallistel, & Gelman, 1999).

The critical question, which we aim to answer in this review, is the following: "What mechanism underlies our ability to compare or estimate large numerosities?" One possibility is that the ability to compare or estimate numerosities is supported by an ANS system where numerosity estimates are derived independent of their confounding sensory cues. As an alternative we propose a sensory-integration system.

In short, this theory conveys that different sensory cues are used to derive an estimate of numerosity. Each sensory cue gets a weight; consequently those that are more prominent will have a larger weight and thus will also play a larger role in determining which set of items is more numerous. Sensory cues that point in the opposite direction (i.e. when the smaller number has larger dots or vice versa) will have a negative weight and bias the person towards choosing the incorrect stimulus. The final "balance" of the weights given to the different sensory cues determines which stimulus is considered more or less numerous. Although the current consensus position is that of the ANS, we will present findings from different domains to suggest that the sensory properties of numerosity stimuli play a larger role in numerosity processing than currently assumed.

Even though some researchers in the field might acknowledge that sensory cues play a role in numerosity processing, the precise role of these sensory cues on current results is hardly investigated and often completely ignored. By providing a systematic and comprehensive review of the literature, we aim to shift the current view towards a better understanding of the role of the sensory properties in numerosity judgments. This in turn would have substantive implications for research on the precursor for numerical abilities and the cognitive architectures and neural substrates in non-human animals as well as human children and adults.

In this review, first the ANS theory is outlined. Next, some shortcomings of the ANS theory are described both from an evolutionary as well as from an empirical and computational perspective. Throughout the whole review, starting points are provided for building a framework that includes a role for the sensory properties of which numerosity stimuli are constructed. In the final part of the review, an alternative sensory-integration theory will be presented in detail and a parallel between numerosity judgments and conservation abilities (Piaget, 1965) will be proposed. Drawing such a parallel can provide future studies with a starting point to validate the sensory-integration theory.

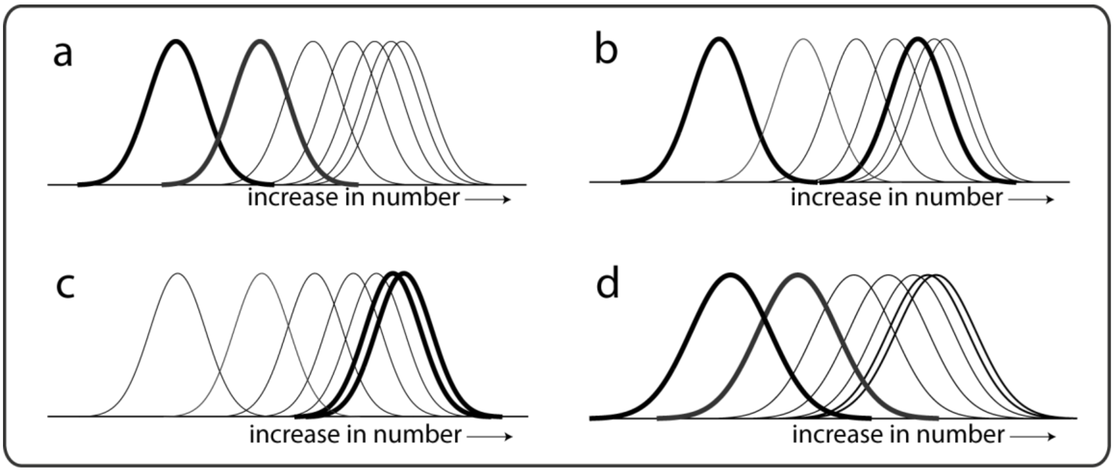

Figure 1. A schematic representation of the neural representation of number. The processing of a number results in the activation of the neural representation of that number as well as its neighboring numbers. This activation of neighboring numbers decreases with increasing distance from the number that is processed. Consequently, a larger overlap in neural representation exists between two neighboring numbers (a) than two numbers further apart (b). Also, the overlap in neural representation increases with increasing number. It is therefore easier to compare two numbers that are further apart than two numbers closer to each other (a versus b). This phenomenon is called the distance effect (Moyer & Landauer, 1967). Similarly it is easier to compare two relatively small numbers than two relatively large numbers, even when their number distance is kept constant (a versus c). This is the so-called size effect (Moyer & Landauer, 1967). The approximate number system hypothesis suggests that the neural representation of numbers in humans who are less skilled in mathematics will be less accurate (the width of the Gaussian curve will be larger) (a versus d).

1.1 The Approximate Number System

Using habituation or violation of expectation paradigms it has repeatedly been shown that infants of 6-month-old can differentiate groups of items that differ with a ratio of 1:2 (Lipton & Spelke, 2003; Wood & Spelke, 2005; Xu & Spelke, 2000; Xu et al., 2005) and that this precision increases with age up to a ratio of around 11:12 in adulthood (Gebuis & Van der Smagt, 2011; Halberda & Feigenson, 2008; Holloway & Ansari, 2009; Pica, Lemer, Izard, & Dehaene, 2004; van Oeffelen & Vos, 1982). The ANS is argued to be the mechanism that supports the development of numerosity processing.

The ANS is specific for the comparison or estimation of large approximate quantities. Researchers have argued that the ANS processes large approximate numerosities *independent of their confounding sensory properties* (e.g. surface or density). This idea comes from a large number of studies showing that human (e.g. Halberda & Feigenson, 2008; Libertus & Brannon, 2009; Piazza & Izard, 2009) and non-human animals (e.g. Brannon & Terrace, 1998; Nieder, Freedman, & Miller, 2002) can compare, estimate, and even perform simple calculations on numerosities. In these studies, the sensory properties are manipulated in such a manner that a single visual variable is not a helpful cue in judging numerosity across all trials. For example, when diameter is kept constant in half of the trials and surface in the other half of the trials, reliance on either of the two cues is suggested to result in performance at chance (Dehaene, Izard, & Piazza, 2005). Above chance performance is therefore considered indicative of a numerosity system that processes number independently of its sensory properties.

In these ANS studies, performance is *ratio dependent.* Participants become more accurate and faster to judge which of two numerosities is larger with increasing relative difference between the two numerosities (Gebuis & Van der Smagt, 2011; Gilmore et al., 2011). It is hypothesized that the underlying mechanism of this ratio dependence effect is the amount of overlap in neural activation patterns of neurons representing numerosities. When the relative difference between two numerosities increases, the overlap in neural activation pattern decreases, resulting in faster and more accurate judgments of numerosity (Dehaene, 2003) (see Figure 1). More specifically, the presentation of a numerosity results in an activation that is maximal for that numerosity but decreases with increasing numerical distance from that numerosity. For instance, neurons firing for numerosity 12 will fire maximally when 12 items are presented but the activation of these neurons will be less for 10 items and much less or even absent for 8 items. Hence, when a numerosity is perceived, not only this numerosity but also the neighboring numerosities are activated, albeit to a lesser degree. In numerosity comparison tasks, two numerosities are presented and the overlap between both their Gaussian neural activation patterns determines the ease with which the two numerosities are differentiated. This idea is supported by several neuroimaging studies (Cohen Kadosh, Bahrami, et al., 2011; Harvey, Klein, Petridou, & Dumoulin, 2013; Piazza et al., 2004; M. Piazza, Pinel, Le Bihan, & Dehaene, 2007). For instance, Piazza et al. (2004, 2007) showed that neural adaptation effects are dependent on the relative numerical distance between the adaptation and test stimuli. Additionally, single cell recordings in monkeys have documented neurons that are maximally responsive to their preferred number. These neurons also respond to the neighbouring numbers and this activation decreases with increasing distance from the preferred number (Nieder & Merten, 2007). The change in performance or neural response with respect to the distance between numerosities that need to be compared is called the numerical distance effect and is considered indicative of the ANS.

Different non-human animals such as bears, monkeys, chicks, fish and pigeons, also show a spontaneous representation of numerosity (e.g. Dadda, Piffer, Agrillo, & Bisazza, 2009; Emmerton & Renner, 2009; Hauser, Carey, & Hauser, 2000; Rugani et al., 2014; Vonk & Beran, 2012). Remarkably, these studies showed that even without intensive training sessions, animals could compare different sets of items and respond differently to correct compared to incorrect outcomes of simple arithmetic problems (e.g. Agrillo et al., 2010; Cantlon & Brannon, 2005; Flombaum, Junge, & Hauser, 2005; Hauser, Tsao, Garcia, & Spelke, 2003). It appears logical that the ability to discriminate or estimate numerosity is necessary for foraging and fight or flight decisions. For instance, it would be beneficial for fish to join the largest shoal as this decreases the chance of being caught by a predator (Agrillo, Dadda, Serena, & Bisazza, 2008). The ANS has therefore been suggested an *evolutionarily ancient system* (e.g. Cantlon, Platt, & Brannon, 2009; Dehaene, 1997; Hauser et al., 2003; Piazza et al., 2004).

Together, a large number of empirical and modelling studies argued to provide evidence for an evolutionarily ancient ANS that processes large numerosities independent of the confounding sensory cues in a ratio dependent manner. The ANS has been described in great detail and forms the basis for a diverse array of theories in different fields of research including developmental psychology, cognitive psychology, neuropsychology, evolution, neuroscience, and education (Dehaene & Brannon, 2010). In particular, the ANS is suggested to play a fundamental role in learning mathematics (Halberda et al., 2008), which has great impact on everyday life functioning (Ansari, 2008; Butterworth, 1999; Rubinsten & Henik, 2009).

1.2. An evolutionarily ancient number system

The ANS theory proposes that the ANS is an evolutionarily ancient system (Agrillo, 2015; Beran, Perdue, & Evans, 2015). Most studies investigating the ability of animals to judge numerosities suggest that this is not surprising as it is of utmost importance for animals to have the capacity to judge numerosity. In the absence of such an ability, core survival-related abilities including foraging and fight or flight decisions, are compromised, thereby reducing fitness (e.g. Agrillo, 2015; Ansari, 2008; Beran et al., 2015; Cantlon, Platt, et al., 2009; Dehaene, 1997; Hauser et al., 2003; Piazza et al., 2004).

Although it is intuitively appealing that numerosity itself plays a central role in survival-related abilities, a closer look at the behavior of animals and their natural environment suggests otherwise. Biologists proposed that an evolutionary stable strategy for animal decision-making in conflict and social settings is more complicated than simply looking at numbers as it depends on several factors that contribute to the fitness of and potential costs for the animal (Cresswell & Quinn, 2004; Lee, 2008; Smith, 1974). Prey responds dynamically to their predators and rapidly adapts their behavior to change the probability of being caught (Lima & Dill, 1990). An example is the escape speed of redshanks, which depends on the distance between the birds in a group (Hilton, Cresswell, & Ruxton, 1999) or how wind speed affects attacking behaviour (Hilton, Ruxton, & Cresswell, 1999). When choosing which of 2 groups of redshanks to attack, a sparrowhawk may not have enough time to determine which group is the most vulnerable because the group of redshanks moves frequently and changes its size, etcetera (Cresswell & Quinn, 2004).

Similarly, it has been shown that tigers as well as leopards choose their prey on the basis of size (Ullas Karanth & Sunquist, 1995). Size-dependent responses to potential prey have also been demonstrated in zebrafish larvae and frogs (Bianco, Kampff, & Engert, 2011; Lettvin, Maturana, McCulloch, & Pitts, 1959). Recently, it was concluded that size discrimination in the visual system is a central classification process for rapid response selection with regard to prey in animals (Preuss, Trivedi, Vom Berg-Maurer, Ryu, & Bollmann, 2014). The same line of reasoning observed in predation also holds for foraging: chimpanzees (Pan troglodytes) prefer fewer items but a larger amount of food to more items and a smaller amount of food (Boysen, Berntson, & Mukobi, 2001). In sum, fight or flight decisions cannot simply be explained on the basis of numerosity alone. Rather, other variables such as visual properties (e.g. size) seem to be important factors in the decision process.

1.2.1. Large approximate numerosities.

Although various variables other than numerosity appear to play a role in animal fight or flight behavior, several studies suggest that animals can judge numerosity. Of these, a large number of studies investigated numerosity processing within the subitizing range (<4 or 5) (Agrillo et al., 2010; Beran, 2007; Beran, Decker, Schwartz, & Schultz, 2011; Brannon & Terrace, 1998; Cantlon & Brannon, 2005, 2007; Nieder & Merten, 2007; Nieder & Miller, 2004). In the current review the focus does not lie on these studies because the majority of researchers consider only large numerosities as the hallmark of the ANS but more importantly controlling the sensory cues in small numerosities is problematic if not impossible (Gebuis & Reynvoet, 2011). Gebuis and Reynvoet (2011) showed that numerosities 1 to 3 have too few degrees of freedom to decorrelate numerosity and sensory cues. It is also difficult to talk about density or convex hull when only one or two dots are presented. Of greater value are therefore studies that not only included small numerosities but also large numerosities. Such studies show a discrepancy between results (Jones & Brannon, 2012). About half of the studies showed an upper limit for the animals around 3 or 4 items (Hauser et al., 2000; Rugani, Regolin, & Vallortigara, 2008; Stancher, Sovrano, Potrich, & Vallortigara, 2013; Wood, Hauser, Glynn, & Barner, 2008) whereas the other half of studies observed that animals could respond to both small and large numerosities (Agrillo et al., 2008; Dadda et al., 2009; Flombaum et al., 2005; Franks et al., 2006; Gomez-Laplaza & Gerlai, 2012, 2013; Rugani, Cavazzana, Vallortigara, & Regolin, 2013).

Interestingly, the studies that show that animals can perform comparisons across or above the subitizing range did not control for the sensory cues that are confounded with numerosity (Agrillo, Dadda, & Bisazza, 2007; Agrillo et al., 2012; Uller, Jaeger, Guidry, & Martin, 2003), only used minor controls (Dadda et al., 2009; Flombaum et al., 2005; Rugani, Cavazzana, et al., 2013; Rugani, Vallortigara, & Regolin, 2013), or showed that this behavior is only observed when sensory cues are confounded with numerosity (Agrillo et al., 2008; Franks et al., 2006; Gomez-Laplaza & Gerlai, 2012, 2013; Rugani et al., 2008; Rugani, Regolin, & Vallortigara, 2011). For instance, Agrillo et al. (2008) and Gomez-Laplaza and Gerlai (2012; 2013) showed that fish could dissociate small and large numerosities only when surface and movement were confounded with numerosity. Similarly, Franks et al. (2006) showed that ants can make numerosity judgments when numerosity is confounded with luminance. The observation that animals cannot judge numerosities that cross the subitizing range appears odd in light of an evolutionary ancient numerosity system that supports survival decisions. If sensory cues are necessary for a correct discrimination, then the idea of spontaneous representation of number becomes questionable.

Studies that only include large numerosities and no training phase to assess spontaneous numerosity representation can provide strong evidence in favor of an ANS. The studies that fulfil these criteria strongly support the pivotal role of non-numerical cues. For example, Krusche et al. (2010) showed that salamanders base their decisions on movement. They could dissociate different sets of flies when the flies could move around but when movement was controlled, performance dropped to chance level. Similarly, Frommen et al. (2009) showed that fish can dissociate 8 from 12 fish when density or surface are not controlled. Importantly, the fish did not have a preference anymore for the larger shoal when density was equalized between both shoals.

It should be noted also that the majority of studies do not test whether sensory cues can explain their results. Generally, it is assumed on the basis of several controls applied that sensory cues cannot be used to judge numerosity. However, the necessity to test the role of sensory cues becomes apparent in the following example. Monkeys can judge numerosities 1-9 when either color, shape, diameter or surface area is controlled for (Brannon & Terrace, 1998). Judge, Evans, and Vyas (2005) replicated this study and confirmed that monkeys can indeed judge numerosities 1-9, and that overall performance is comparable to other studies. This was however only the case when data was pooled across different surface area controls. When the different surface area conditions were analyzed separately it became apparent that the monkeys were probably using surface area as a cue. Based on this study and an additional experiment it was concluded that the monkeys most likely used multiple strategies to solve the task, amongst which reliance on sensory cues.

Together, studies that investigated large approximate numerosities failed to show spontaneous numerosity representation while sensory cues were controlled to some extent. Rather, these studies point out that animals require sensory cues to make everyday decisions pertaining to fight or flight responses.

1.3. Mechanisms of the ANS model

The mechanism of the ANS is not fully understood but the most influential models suggest that the ANS consists of a normalization and an accumulator stage (see Figure 2) (Dehaene, 1997; Dehaene & Changeux, 1993; Piazza & Izard, 2009; Stoianov & Zorzi, 2012; Verguts & Fias, 2004; Whalen et al., 1999). These models suggest that the first stage of the ANS is a sensory stage where the incoming sensory properties of stimuli are processed (stage 1). For a visually presented stimulus this could be the size of the different elements, the density, the convex hull, etc. These sensory properties increase or decrease with increasing numerosity and are therefore confounded with numerosity. To account for this confound, the sensory information is transferred to the next stage, the so-called *normalization stage* where the confounding sensory information is removed and pure numerosity information is retained (stage 2). This numerosity information is then fed to the accumulator (stage 3). To describe this stage, researchers often use an analogy of a beaker that pours water into a cylinder for each item perceived (Dehaene, 1997). Because the beaker does not always contain exactly the same amount of water, each time that water is poured into the cylinder the total amount of water becomes increasingly inaccurate. This, in the end, results in an approximate and not an accurate estimate of numerosity (stage 4). The decrease in accuracy of the final amount of water with an increasing number of scoops of water is consistent with the observation that numerosity discrimination performance decreases with increasing numerosities (i.e. ratio dependent performance) (Gebuis & Reynvoet, 2012c; Izard & Dehaene, 2008). A direct reflection of this process is described in the computational model presented by Stoianov and Zorzi (2012). This model could compare numerosities with similar accuracy as human adults can (Piazza et al., 2010), suggesting that the mechanism described in the ANS model could underlie our numerosity processes. Also, results from neuroimaging and single cell recording studies are in line with that model (Piazza et al., 2004; Roggeman et al., 2011; Roitman, Brannon, & Platt, 2007).

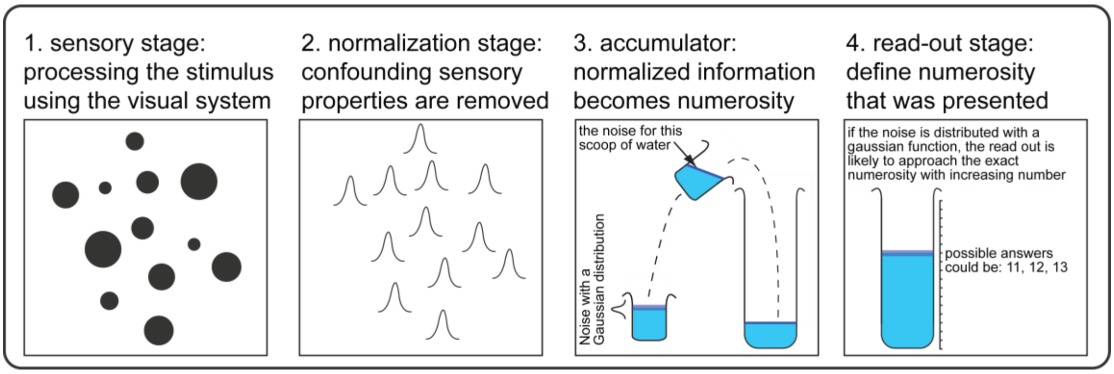

Figure 2. The accumulator model. The first stage is the sensory stage where the stimulus is processed by for instance the visual system. The second stage is the normalization stage. Here, the curves represent the neuronal response to the presence of an object without its sensory characteristics. The third stage is the accumulator. Here a scoop of water is added to a beaker for every item detected in the normalization stage. The beaker is not 100% accurate. The Gaussian curve represents the distribution of the noise present for each scoop of water. The fourth stage shows the final result.

An observation in the literature that is difficult to reconcile with the current models concerns the normalization stage, which eliminates the confounding sensory cues of numerosity stimuli. If this stage did not exist, it would not be possible to extract "pure" numerosity and consequently mistakes would be made in situations where sensory cues are misleading. For instance, the situation can arise where 8 large oranges make up a larger pile than 10 small pears. Without the normalization stage we would falsely conclude that the physically larger pile of oranges contains more fruits.

However, studies investigating the effect of sensory cues on numerosity comparison actually show that the responses of humans and other animal species are affected by the sensory properties of numerosity stimuli (e.g. Gebuis, Herfs, Kenemans, de Haan, & van der Smagt, 2009; Gebuis, Kenemans, de Haan, & van der Smagt, 2010; Gilmore et al., 2011; Inglis et al., 2011). For instance, performance on numerosity comparison tasks can strongly differ between so-called congruent and incongruent trials (see Figure 3). However, it is not agreed upon what the source of these congruency effects is nor is it accounted for by any model.

If indeed an ANS that processes numerosity independent of its sensory cues exists and at the same time is affected by the sensory cues then an adjustment has to be made to such accumulator models to keep the idea of a normalization stage intact. One explanation is provided by the model of Dehaene and Changeux (1993), which does allow sensory noise to enter the normalization stage. This sensory noise is allowed based on the argument that "the operation of normalization of size and location cannot be performed with arbitrary precision by real neuronal systems, the numerosity estimates will therefore be variable and obey Fechner's law" (last paragraph page 401). In other words, sensory information cannot be fully removed from the normalization stage. Consequently, there will always be some influence of sensory cues on numerosity processes. An alternative theory to account for the sensory interference on numerosity would be to assume that sensory cues and numerosity are processed in parallel from the sensory stage onwards and interfere with each other only after normalization occurred (see Figure 4) (see for a similar reasoning: Gilmore et al., 2013; Inglis & Gilmore, 2013; Rousselle & Noel, 2008; Santens & Verguts, 2011).

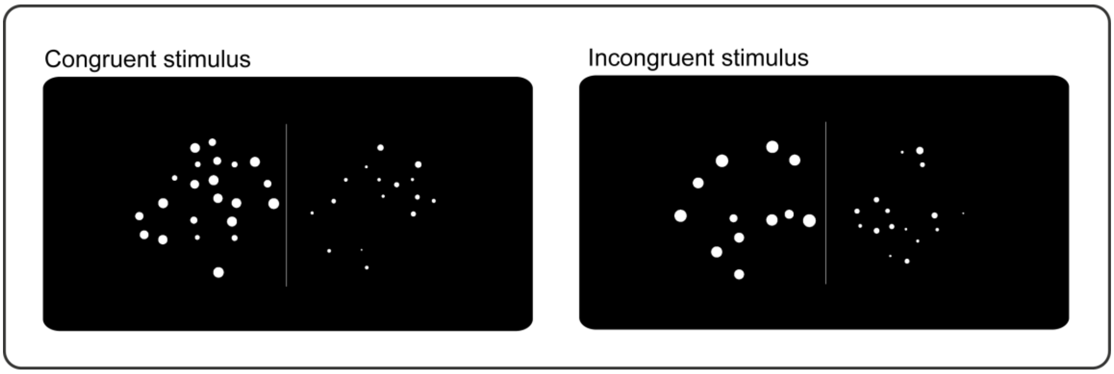

Figure 3. Congruent and incongruent numerosity stimuli. The congruent stimulus consists of two dot arrays of which one contains more dots and larger sensory cues than the other. In contrast, the incongruent consists of two dot arrays of which one contains more dots but smaller sensory cues than the other.

The latter idea of a late interference stage is comparable to the well-known color-word congruency effect (Stroop, 1935). Here, participants are asked to report the ink color of a color word. Participants are slower to respond when the ink color (e.g. red) is incongruent with the color word (e.g. “GREEN”) than when it is congruent with the color word (“RED”). Because colors and words are initially not processed by the same neural mechanisms, it is hypothesized that colors and words are processed in parallel up to a certain stage where both interfere with each other. Note that color-word, but also numerical Stroop (or the so-called size congruity effect, Henik & Tzelgov, 1982; Tzelgov, Henik, & Berger, 1992) studies support this idea (e.g. Cohen Kadosh et al., 2007; Cohen Kadosh, Gevers, & Notebaert, 2011; Gebuis et al., 2010; Lansbergen

& Kenemans, 2008; Santens & Verguts, 2011; Szucs & Soltesz, 2008). Even though these studies concern symbolic content and processing, a similar process could underlie non-symbolic numerosity and sensory information. Thus, according to this late interference reasoning, the ANS on its own would be able to make a "pure numerosity" estimate, it is only afterwards that the sensory cues start to interfere.

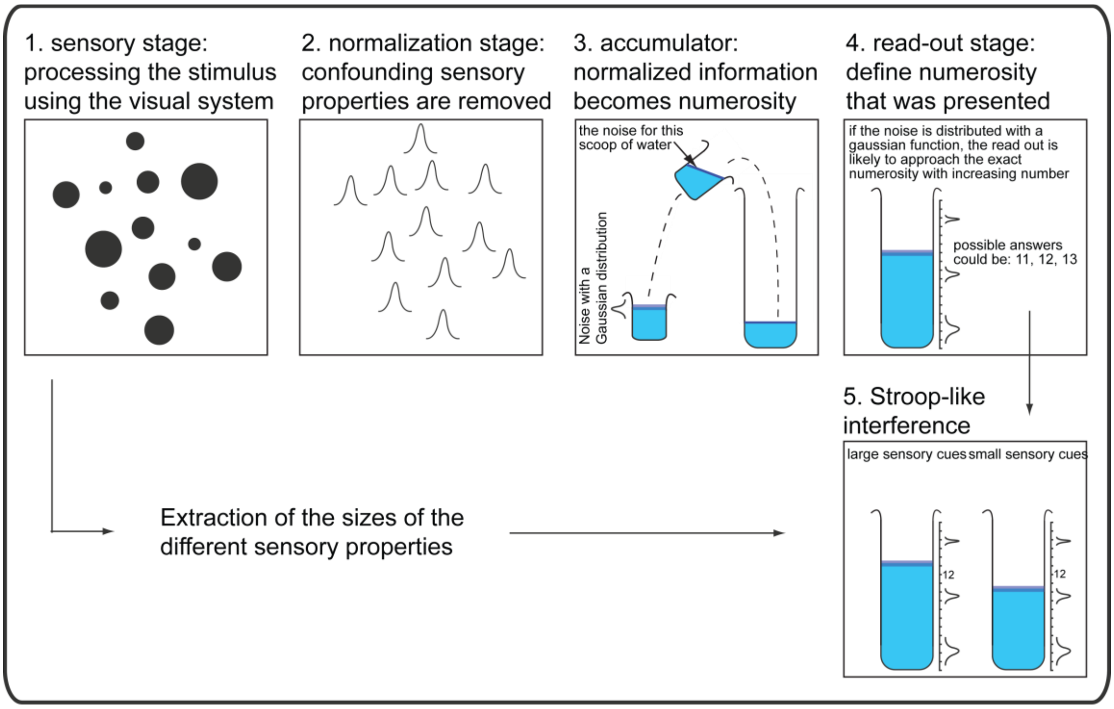

Figure 4. The accumulator model and Stroop-like interference. For an explanation of the first four stages see Figures 1 and 2. The fifth stage shows the parallel processing of the sensory cues from the sensory stage onwards and their influence on the final numerosity estimate. For instance, the final estimate increases or decreases in size in the fifth compared to the fourth stage relative the size of the different sensory cues.

Both theories: 1) improper extraction of numerosity at the normalization stage (Dehaene & Changeux, 1993) and 2) post-processing interference are not explained in detail in the ANS model and it can therefore only be speculated how they would affect the pure numerosity estimate of the ANS. The first theory (Dehaene and Changeux 1993), suggests that the improper extraction of numerosity at the normalization stage is dependent on the size and the location of the items, more specifically, the diameter and the density of the dots. Thus, it could be that when the size of the individual items and density increase, it becomes more difficult to disentangle the different items (i.e. the items being closer to each other makes it more difficult to properly extract them as separate entities), leading to an underestimation of numerosity. In other words, the number of items in a dense set would be estimated as less numerous than the number of items in a less dense set. However, studies investigating the effect of density on numerosity judgments showed the opposite pattern: participants judge a less dense array as less numerous (Gebuis & Reynvoet, 2012c; Sophian, 2007). This theory of improper extraction cannot unambiguously account for the observed influence of sensory cues on numerosity processing.

The second theory suggests that sensory cues are processed in parallel and in the end interfere with the abstract numerosity. For example, a stimulus with large sensory cues could bias the final estimate towards a larger number, while a stimulus with small sensory cues could bias the final estimate towards a smaller number. Note that irrespective of how the sensory cues influence the numerosity estimate, the direction of the influence is always the same. It is not logical that one time a larger stimulus causes an overestimation of numerosity and the other time an underestimation. Although the majority of studies show that larger sensory cues cause an overestimation and smaller sensory cues an underestimation, exceptions exist (Gebuis & Reynvoet, 2012c; Gebuis & Van der Smagt, 2011; Izard, Dehaene-Lambertz, & Dehaene, 2008). Gebuis and Reynvoet (2012d) show that dot-arrays consisting of a smaller surface are estimated as larger, while Izard and Dehaene (2008) and Guillaume et al. (2015) found a general underestimation.

Together, both theories that account for the observed influence of the sensory cues do not seem to hold. Still, if either explanation were to hold or a similar alternative existed; it would imply that the influence of sensory cues is inevitable. In such a case, not a single method to control sensory cues would be able to prevent sensory cues to influence the final numerosity estimate. Thus measuring “pure numerosity” processes as is often suggested in manuscripts that “control all sensory properties” would not be possible. Still, if either of the two or both mechanisms existed, why would there be an ANS system that can extract “pure numerosity”? What would be the use of having a system that can tell us exactly which cue at the passport control contains less people when it in the end adjusts this accurate answer in a possibly incorrect answer when for instance the length of the people in the cue is taken into account?

The idea that an influence of sensory cues on numerosity judgments is inevitable was taken into account in a recent model of Leibovich and Henik (2016). These authors proposed a magnitude system that separately extracts the numerosity as well as the sensory cues from the stimulus. In a next phase, these numerosity and sensory cues can interact with each other and the result of this interaction would feed into a non-symbolic processing system (see Figure 7, page 913 Leibovich & Henik, 2014). This non-symbolic system again consists of two subsystems, the object tracking system for small numerosities (see Piazza, 2010) and the ANS for large numerosities. Even though this model takes into account sensory cues, in reality it is highly similar to the model proposed by Dehaene and Changeux (1993) with a late stage interference of sensory cues (see above). The most important difference between both theories lies in the terminology. Dehaene and Changeux (1993) termed the extraction of numerosity independent of sensory cues the ANS system and allows a late stage interference of sensory cues. Leibovich and Henik (2016) call the initial extraction of numerosity the discrete system and only after a possible interaction with numerosity it is called the ANS. The same reasoning as applied above to the ANS model also holds for the model of Leibovich and Henik (2016) model: why having a system that can extract numerosity independent of sensory cues but subsequently receives interference from sensory cues? A novelty introduced in the model of Leibovich and Henik (2016) is the role attributed to cognitive control. Cognitive control would be necessary to “disambiguate numerical from continuous variables during non-symbolic number processing” (p. 10). This would imply that numerosity could be extracted independent of the sensory cues, which is at odds with their own proposal that the influence of sensory cues is inevitable in numerosity judgment tasks (Leibovich & Ansari, 2016; Leibovich & Henik, 2013, 2014).

Together, it appears that the inclusion of a normalization stage, which is a critical stage for models that describe the ANS, is not in line with empirical observations showing clear interactions between numerosity judgments and sensory cues. Adding a parallel processing stage for sensory cues, where the sensory cues interfere with the final numerosity estimate, could solve this problem. But this still leaves us with a system where effect of sensory properties are present and therefore questions the reason of having a normalization stage in the first place. Based on Occam’s Razor the less parsimonious explanation for the observation of the sensory effects in numerosity estimation, i.e., the sensory-integration theory, should be preferred (Cohen & Dehaene, 2001). Here, the sensory cues do not induce noise to the final estimate but instead are used to make the estimate.

2. CONTROLLING VERSUS MANIPULATING SENSORY CUES.

Interestingly, even though the presence of sensory effects are not accounted for in the ANS model, it is acknowledged that they can interfere with number processes. This is evident from the kind of paradigms used in studies on numerosity processing. Roughly two streams of research can be dissociated. On the one hand studies exist that try to *control* the sensory cues to investigate numerosity processing independent of its sensory cues (see Figure 5) (e.g. Ansari, Fugelsang, Dhital, & Venkatraman, 2006; Burr & Ross, 2008; Cantlon et al., 2006; de Hevia & Spelke, 2009; Izard, Sann, Spelke, & Streri, 2009; Piazza et al., 2004; Roggeman et al., 2011; Santens, Roggeman, Fias, & Verguts, 2010). In these studies, the sensory cues are manipulated so that each single sensory cue is not helpful to judge numerosity throughout the experiment. Consequently, if participants judge numerosity above chance level it is argued that numerosity can be judged independent of sensory cues.

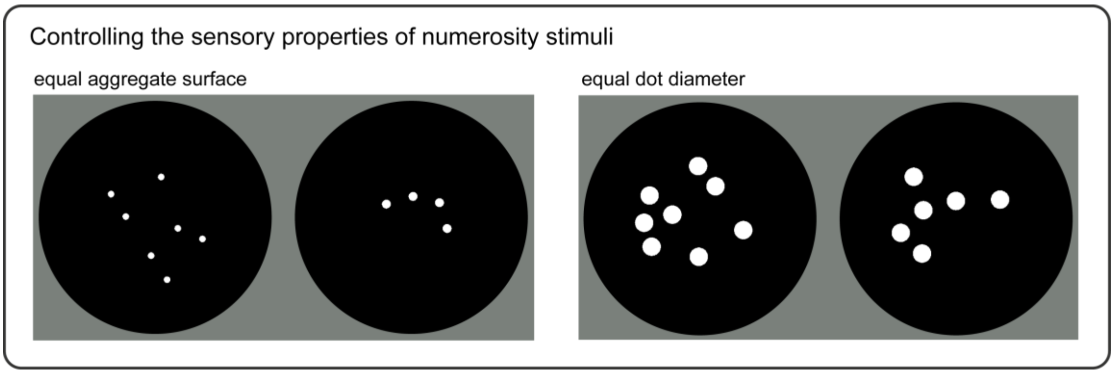

Figure 5. Examples of numerosity stimuli that are controlled for sensory cues. The program of Dehaene et al. (2005) controls for aggregate surface in half of the trials (left image) and for dot diameter in the other half of the trials (right image). Across all trials, aggregate surface and dot diameter are not helpful cues to judge which dot array contains more dots.

On the other hand studies exist that *manipulate* sensory cues to investigate their role in numerosity processes (see Figure 6) (Allik & Tuulmets, 1991; Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2012; Frith & Frith, 1972; Gebuis & Reynvoet, 2012a; Ginsburg, 1991; Ginsburg & Nicholls, 1988; Sophian, 2007; Sophian & Chu, 2008; Szucs & Soltesz, 2008; Szucs, Soltesz, Jarmi, & Csepe, 2007). In these studies, one or more sensory cues of dot arrays that represent the same numerosity are manipulated. If the participant’s decision is influenced by the different sensory properties of dot arrays representing the same numerosity it can be concluded that numerosity judgments are not independent of the sensory properties.

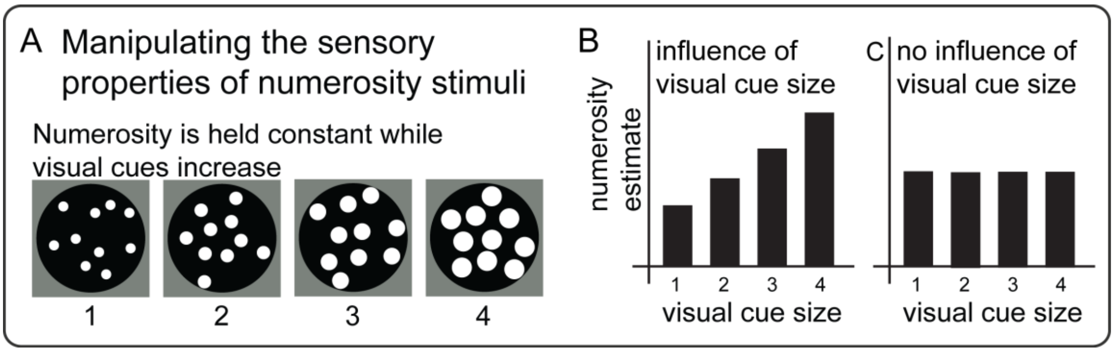

Figure 6. Manipulating the sensory cues. a) The 10 dots presented differ in dot diameter. b) When participants estimate the number of dots presented on the basis of the sensory cues, the numerosity estimate will be correlated with the size of the sensory cues. c) However, when numerosity is processed independently of its sensory cues, the numerosity estimate will be the same for the different dot arrays.

Generally studies controlling for sensory cues and studies manipulating sensory cues reach different conclusions. Those controlling for sensory cues suggest the existence of an ANS that processes numerosity independent of the sensory cues. In contrast, studies that manipulate the sensory cues, suggest that sensory cues play an important role in numerosity processes. Clearly, both opposing viewpoints cannot exist at the same time. Resolving this discrepancy can elucidate the mechanism(s) supporting our numerosity judgment abilities. In the next paragraphs, it is explained how this discrepancy could arise.

2.1. Controlling the sensory cues

Central to the discussion of whether sensory cues can or cannot explain observed numerosity processes is the method that is used to control the sensory cues. Unfortunately, studies that aimed to control sensory cues neglected the most important fact: a perfect control for sensory cues is impossible. Namely, two sets of items can never consist of exactly the same sensory cues while differing in numerosity. It is therefore inevitable that two sets of items that differ in numerosity also differ in one or more sensory cues. The sensory cues that remain present in each and every trial can therefore form the basis for the numerosity judgment. Thus it remains to be investigated whether current results can be attributed to numerosity processes or to the confounding sensory cues.

2.1.1. Sensory controls in behavioral studies

The paradigm created by Dehaene et al. (2005) offers a seemingly elegant solution to control the sensory cues and is therefore one of the most frequently used methods to control the sensory cues in numerical cognition studies. In this paradigm, surface is kept constant in half of the trials and diameter is kept constant in the other half of the trials (see Figure 5). Another frequently used procedure is to make half of the trials congruent and the other half of the trials incongruent (see Figure 3) and to present these in separate arrays or a single array using different colors[1] (Halberda et al., 2008; Vonk & Beran, 2012)[2]. In either case, it is reasoned that each sensory property is not a helpful cue across all trials. In these cases, it is correct to conclude that participants did not rely on *a single sensory cue* throughout the experiment. However, it cannot be concluded that participants do not rely on *any sensory cue*.

This is the case because the methods for controlling the sensory cues rest upon two assumptions that are questionable. First, it is assumed that participants only take

---

[1] This method is suggested superior to presenting stimuli in two separate arrays, as it would exclude the effect of density and inter-item spacing. However, this method controls diameter and surface by using congruent and incongruent trials therefore the arguments presented in the following paragraphs are also applicable to this method.

[2] Other paradigms exist as well, such as violation of expectation or item-by-item paradigms that are predominantly used in infants or nonhuman animal studies. We do not discuss these procedures in the current manuscript as they are mainly used in studies including small exact numerosities, which are generally not considered a part of the ANS, and are therefore beyond the scope of the current review.

into account a *single sensory variable* at the time and second, do not *switch between different sensory cues* throughout the experiment. These assumptions do not hold as previous studies showed that multiple sensory cues are used at the same time to judge numerosity (Allik & Tuulmets, 1991; Gebuis & Reynvoet, 2012b; Smets, Sasanguie, Szucs, & Reynvoet, 2015). For instance, Gebuis and Reynvoet (2012b) performed a numerosity comparison task and created four different sensory cue conditions. These conditions differed in the number and kind of sensory cues that were kept constant and were manipulated. The results showed that when more sensory cues were available to the participants, the congruency effect increased. Thus the combination of several sensory cues resulted in a congruency that was of comparable size as their separate effects combined. This can of course only occur if participants take into account multiple sensory cues at the same time to judge numerosity. Similarly, Smets et al. (2015) showed that with increasing complexity of the sensory controls performance can decrease with even more than 20%.

The few studies that investigate congruency effects report opposite congruency effects: better performance on (partially) incongruent compared to congruent trials (Gebuis & Reynvoet, 2012b; Gebuis & Van der Smagt, 2011; Hurewitz, Gelman, & Schnitzer, 2006). This can be explained using a sensory-integration explanation. Namely, when multiple sensory cues play a role in numerosity processes, the different sensory cues can compete with each other in the weight given to each sensory cue. For instance, if there is a stimulus that is incongruent in surface but congruent in diameter, more weight could be given to diameter if this cue is more salient (e.g. differs to a larger extent between the two stimuli that need to be compared). In this case better performance for surface incongruent trials can be observed. Two sensory cues can also cancel each other’s effect. For instance the congruency effect for convex hull (which is the smallest contour that can be drawn around a set of items) can cancel out the incongruence for surface (Gebuis & Reynvoet, 2012b). In contrast, the ANS theory cannot explain the presence of sensory biases let alone opposite biases for similar sensory cues.

Another elegant method to control the sensory cues is applied in infant habituation studies (Libertus & Brannon, 2009; Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu et al., 2005). Generally, a subset of sensory cues is manipulated in the habituation phase and kept constant in the test phase and vice versa. In this manner a single sensory cue cannot be used throughout the whole experiment. Here, similar as for the Dehaene et al. (2005) method, it is assumed that infants do not take into account multiple sensory cues at the same time nor switch between them.

Possible confounds and difficulties with interpretation of earlier infant habituation studies (e.g. use of small number, confound with non-numerical information, the presentation of multiple changing variables at the same time, etc…) have been clearly outlined in earlier work (Cantrell & Smith, 2013; Clearfield & Mix, 1999; Mix, Huttenlocher, & Levine, 2002). However, the more recently developed change detection task (Libertus & Brannon, 2010; Libertus, Starr, & Brannon, 2014; A. B. Starr, M. E. Libertus, & E. M. Brannon, 2013) still requires consideration. In this paradigm, two streams of stimuli are presented at the same time of which one consists of images presenting the same number while the other stream consists of images that alternate in number (e.g. 6-6-6-6 versus 6-24-6-24). When infants spend more time looking at the alternating stream it is concluded that they noticed the changes in number. The sensory cues are controlled by equalizing them between the two streams (i.e. in a third of the trials is surface equal between both streams, while in another third of the trials diameter is equal, etc.). Problematic in such studies is that the sensory cues are controlled *between* but not *within* one stream. Consequently, sensory cues can still change to a larger extent in the numerosity different than the numerosity same trials. Indeed, post hoc calculations of the Libertus et al. (2010) stimuli showed that the change in convex hull could even be 5-fold larger in the changing stream compared to the non-changing stream in the 6-24 comparison condition. This effect decreased to a difference around 1.5-fold in the 8-16 comparison. The decrease in numerosity distance together with the decrease in sensory change can again nicely explain the distance effects observed. That the infants might have responded to convex hull rather than numerosity is likely because convex hull has been shown a stronger cue than numerosity before (Cleland & Bull, 2015; Gebuis & Gevers, 2011). Also, when sensory cues are controlled within one stream, performance drastically decreases (Smets et al., 2014). Thus, infants might be looking longer to the changing stream simply because within that stream sensory cues are changing to a larger extent. Overall, irrespective of the kind of sensory control used, sensory cues are inevitably present in a numerosity stimulus, meaning that sensory cues can be used even when sensory controls are applied (see our last paragraph: “reliance on sensory cues” for a more detailed explanation about how this can be achieved).

Besides the troublesome assumptions underlying the methods to control sensory cues, several studies did not control all sensory cues in the first place. For instance, some studies did not control for convex hull (de Hevia & Spelke, 2009; Libertus et al., 2014; Nieder & Merten, 2007). This is problematic because recent studies (Gebuis & Gevers, 2011; Gebuis & Reynvoet, 2012a) showed that the convex hull can explain previous behavioral (de Hevia & Spelke, 2009) and neuroimaging results (Hyde & Spelke, 2009, 2011). The example of neglecting visual cues present in the dot arrays such as convex hull is not an exception; other studies have overlooked, for instance, density (Brannon, Cantlon, & Terrace, 2006; Burr & Ross, 2008; Libertus et al., 2007), dot diameter (Ansari & Dhital, 2006; Lourenco, Bonny, Fernandez, & Rao, 2012), contour length (Ansari & Dhital, 2006; Holloway, Price, & Ansari, 2010; Nieder & Merten, 2007) or brightness contrast (Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Halberda et al., 2008).

In sum, sensory cues are generally not properly controlled. More importantly, controlling the sensory cues is not the same as removing the sensory cues. The studies that control sensory cues rely on several assumptions that have been proven incorrect. This means that the current view about the existence of an ANS could be based on false assumptions. These problems are not only visible in behavioral, but also in neuroimaging, single-cell recording, and modeling studies. Due to differences in study designs we discuss the neuroimaging and modeling studies separately in the next sections.

2.1.2. Sensory controls in neuroimaging studies

Most of the problems associated with methods to control sensory cues in behavioral studies also apply to neuroimaging studies. However, for neuroimaging studies an additional problem exists. The methods to control the sensory cues are intended to manipulate the sensory cues in such a manner that the participant cannot rely on a single sensory cue to judge numerosity throughout the experiment. In other words the participant cannot develop a *strategy* other than numerosity to perform the task at hand. However, a distinction should be made between participant strategies and neural responses of the brain. A participant can consciously decide to ignore one aspect of the image (e.g. the surface of the dots) and only attend to another aspect of the image (e.g.

numerosity). In contrast, the brain processes each sensory property that is present in the field of view. If for instance a set of dots is presented, the surface, contour length, etc. are processed by visual brain regions, and even by associative cortical brain regions (Silver & Kastner, 2009). The brain does not “ignore” sensory (e.g., surface) information because it is uninformative about numerosity across trials. To deal with this problem, neuroimaging studies cannot rely on the standard sensory controls. Namely in these studies, the problem is that often trials are averaged that belong to the same numerosity condition (Ansari & Dhital, 2006; Cohen Kadosh, Bahrami, et al., 2011; Libertus et al., 2007; Piazza et al., 2004). For instance, if trials are presented with 4, 8, 12 or 16 dots, averages are created for all trials containing 4 dots, 8 dots, 12 dots and 16 dots separately. To control for the sensory cues, the majority of these studies keep a single sensory cue constant across 50% of the trials and increase this sensory cue with number in the other 50% of the trials. Now, an increase in one or more sensory cues with increasing numerosity would be present when all trials are averaged together (see Figure 7 for an example). Thus, the observation of an increase in neural response is not necessarily an indication of numerosity processes as is generally concluded (e.g. Hyde & Spelke, 2009; Temple & Posner, 1998).

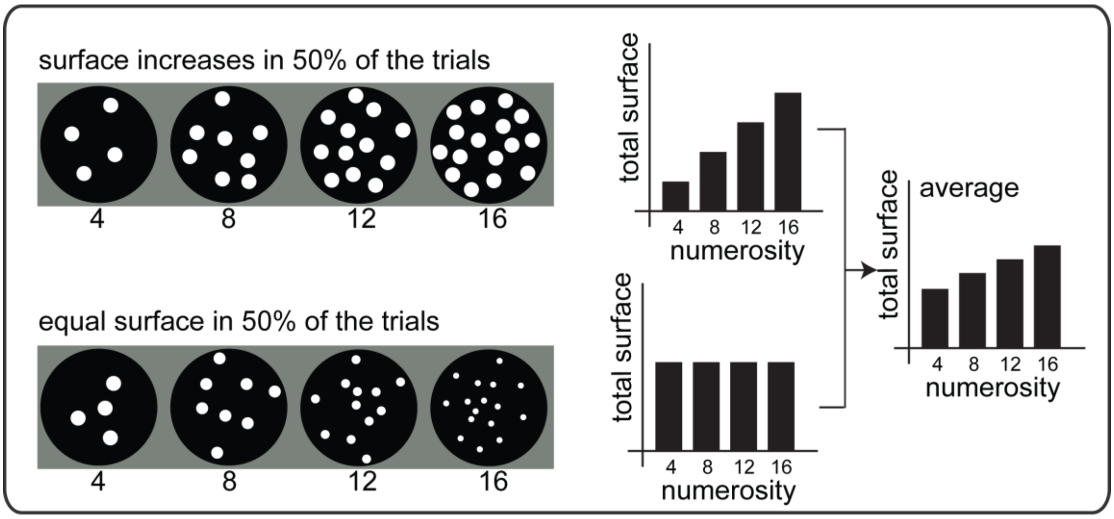

Figure 7. The confound of sensory cues in imaging studies. In imaging studies, numerosity stimuli representing the same numerosity are often averaged. Although different variations of controls exist, the problem pointed out in this example is likely to be present in most studies. If for instance in half of the trials surface increases with increasing numerosity (upper panels) and in the other half of the trials equalized (lower panels) then on average surface still increases with increasing numerosity. Now neurons responsive to surface will on average reveal an increase in activation with increasing numerosity.

Another illustration of this problem can be found in the electroencephalography (EEG) literature where opposite effects were observed when similar numerosity studies were performed but different sensory controls applied. For instance, Temple and Posner (1998) showed that the P2 component decreases with increasing numerosity ratio. Libertus et al, (2007) replicated these findings when the same stimuli were used. However, when the sensory properties of the stimuli were changed, an opposite pattern for the P2 component was found. Libertus et al. (2007) attributed this change in the P2 component to the influence that sensory properties have on the EEG signal preceding the P2. This effect was suggested to be specific for small numerosities because the usage of larger numerosities resulted in the same findings as originally found in the Temple and Posner (1998) study. However, it is premature to conclude that the results for larger numerosities were unrelated to sensory cues given that no second set of large approximate numerosities consisting of different sensory properties was tested. Indeed, from the study of Hyde and Spelke (2009) it can be concluded that also the sensory cues present in large numerosities can influence early event related potentials (ERPs). In experiment 1 it was observed that the N1 became less negative with increasing large approximate numerosities. When changing the sensory properties of the stimuli in experiment 2, the N1 effect disappeared. As only the sensory parameters changed, the changes in the N1 component can only be attributed to sensory cues present in large numerosities. Taking the results of both studies into account, it can be concluded that sensory properties of both small and large numerosities have an influence on early (N1 and P2) processes. These studies nicely demonstrate how easily results can change when the sensory properties of the stimuli change and underline the necessity of a thorough investigation of what exactly the role of the sensory properties is.

Recently, studies directly addressed this issue and showed that changes in the sensory cues, but not numerosity, affect ERP components in a similar manner as is expected from numerosity (Gebuis & Reynvoet, 2012a, 2013; Soltesz & Szucs, 2014). For example, Soltesz and Szucs (2014) used an adaptation paradigm to disentangle sensory and numerosity related effects. They found neural adaptation to sensory properties at the early ERP components. In two other studies Gebuis and Reynvoet (2012a, 2013) reported similar results using a different paradigm. Here, the sensory cues were manipulated in such a manner that they did not linearly increase or decrease with increasing numerosity. In this manner, a possible increase in the EEG peak amplitude of the N1 and or P2 component with increasing numerosity could not be attributed to sensory cues. Crucially, under these circumstances, no numerosity related effects were present for small and large numerosities. However, when the stimuli were reanalyzed according to visual cue size, ERP components appeared that were comparable to those previously attributed to numerosity (Hyde & Spelke, 2009, 2011; Libertus et al., 2007; Mazza, Pagano, & Caramazza, 2013; Temple & Posner, 1998).

2.1.3. Sensory controls in computational modelling

The problem of controlling sensory cues is also applicable to computational modelling studies. Stoianov and Zorzi (2012) presented a computational model that investigated large numerosity processes, which is the hallmark of the ANS. The model consisted of a visible layer encoding the sensory data and two hidden layers. The model first performed a training phase and then a numerosity comparison task. During the training phase, several example stimuli were shown but no task was performed (i.e. training of the deep network). Second, a linear classifier (i.e. a linear combination of the stimulus characteristics are used to decide to which group a stimulus belongs) was used to test the nature of the hidden layer units (i.e. numerosity versus surface area processing). For each training phase, only a single sensory cue was kept constant across all trials. The training phase was conducted several times to test the different sensory cues. For each run, the models invariance to the sensory cue that was equalized was tested. The model appeared largely invariant to the sensory cue that was equalized but no information was provided about the sensory cues that did vary with numerosity. As a result of this, the model's output was independent of sensory effects. Regardless of whether the sensory cues were congruent or incongruent with the numerosity, the same level of performance was maintained. As outlined above, this does not correspond to what is observed in human studies where typically a strong influence of visual cues is observed (e.g. Gebuis, Herfs, et al., 2009; Gebuis et al., 2010; Gilmore et al., 2011; Inglis et al., 2011). On the contrary, it was suggested that because the model could perform the numerosity comparison task in all these control conditions separately, the results could not be explained by sensory cues.

Again the same troublesome assumptions as discussed earlier were made: the possibility that different sensory cues can be combined or that a switch between sensory cues for each run is made is not taken into account. Also only 35 neurons out of the 400 neurons in the second hidden layer were responsive to numerosity during the training phase while 164 neurons were responsive to cumulative surface area. Next, the test to assess the model's ability to classify numerosities was performed on both the sensory and numerosity neurons (total of 199 neurons), and showed accuracy in numerosity comparison that is comparable to human adults (Piazza et al., 2010). Interestingly, the analysis on the 35 neurons that only responded to numerosity led to a significant decrease in the ANS acuity to a level comparable to that of children of 4 years of age (Halberda & Feigenson, 2008).

Another modelling study was performed by Verguts and Fias (2004). This model includes a summation stage in between the sensory processing and number selective tuning. In this summation stage, neurons fire monotonically stronger or weaker with increasing number. This stage replaces the beaker analogy of the original ANS model. The model could nicely simulate the behavioural effects often observed in numerosity studies. In a later study, Chen and Verguts (2013) demonstrated that a model based on summation coding could simulate the neural responses obtained from monkeys that were not trained on numerosity tasks and were originally believed to be numerosity-selective neurons (Bongard & Nieder, 2010; Viswanathan & Nieder, 2013). Summation coding and responses to sensory cues are both monotonic responses and it could therefore be that the summation coding stage does not monotonically respond to the number of items but instead responds to another dimension, the sensory cues present in the stimuli. Note that similar to the above-mentioned EEG and fMRI studies, the monkey physiology studies also did not properly control the sensory cues. On average the sensory cues still increased with increasing number (for a similar reasoning see Gebuis, Gevers, & Cohen Kadosh, 2014; Harvey et al., 2013).

2.2. Manipulating the sensory cues

From the research that we presented so far it is clear that controlling the sensory cues of the stimuli does not prevent participants from relying on these cues nor does it exclude a role for such cues in behavioral, imaging or computational modeling studies. Instead of controlling the sensory cues to see whether participants can still perform the task, it appears more logical to manipulate the sensory cues to study the effect of the sensory cues on performance. If the ANS works independently of sensory processes then manipulations of the sensory cues should not affect numerosity judgments. Therefore, manipulating the sensory cues, instead of controlling them can provide a more definite answer regarding the existence of an ANS. An increasing number of studies have investigated the effect of manipulating sensory cues on numerosity judgments. These studies systematically demonstrate that changes in sensory cues result in changes in performance. To name a few: density (Dakin et al., 2012), clustering (Frith & Frith, 1972), convex hull (Gebuis & Gevers, 2011; Sophian & Chu, 2008), stimulus organization (Ginsburg, 1991), and stimulus diameter (Ginsburg & Nicholls, 1988; Harvey et al., 2013; Sophian, 2007).

It should be noted though that in essence, the studies that try to control for sensory cues (prevent a role for sensory processes in numerosity judgments) are actually studies that manipulate the sensory cues (evoke sensory responses due to changes in sensory cues). Recall that studies that controlled the sensory cues created congruent and incongruent stimuli to make the sensory cues uninformative about numerosity across all trials. These studies do not make all sensory cues equal in each condition (which as discussed earlier is impossible). Instead they change them, and these sensory changes are often in two different directions, i.e., congruent versus incongruent, for the same stimuli. Therefore, a direct comparison between congruent and incongruent trials is comparable to a comparison between two sensory manipulations and would show the inevitable reliance on sensory cues for numerosity processing. The only actual difference between researchers controlling and manipulating the sensory cues is their goal. In the former case, researchers focus on numerosity and try to elucidate how numerosity is processed independent of its sensory cues while in the latter case, researchers focus on the sensory cues and try to investigate the role that sensory cues play in numerosity processes.

2.3. Saliency

Besides controlling or manipulating sensory cues, a few studies also pitted numerosity against sensory cues (Brannon, Lutz, & Cordes, 2006; Cantlon & Brannon, 2007; Gebuis & Reynvoet, 2012a, 2013; Libertus et al., 2014). The dimension to which participants respond in a forced choice-like task determines the more salient feature. In these studies it is reasoned that if numerosity is more salient than a single sensory cue, the stimuli are processed by an ANS and not by a sensory system. Studies that used such a paradigm show conflicting results. For instance, Brannon et al. (2006) showed that infants are equally sensitive for numerosity as for surface area while the data for rhesus macaques in Cantlon and Brannon (2007) were mixed. Monkeys previously trained on numerosity tasks responded to numerosity but the number naïve monkey responded initially predominantly to sensory cues. These studies included small and large numerosities but similar opposing findings are also apparent when considering studies that only included large approximate numerosities. Libertus et al. (2014) showed that infants prefer to attend numerosity to surface area while Gebuis and Reynvoet (2012a) showed that adults process various sensory cues and not numerosity even when they are instructed to focus on numerosity (Gebuis & Reynvoet, 2013). A possible explanation for these conflicting results refers back to the previously mentioned remark that often only a single or a subset of sensory cues is controlled for. More specifically, numerosity stimuli consist of different sensory cues that together represent a single numerosity. This means that pitting numerosity against a single sensory cue might not be relevant when other sensory cues remain to vary possibly in opposite directions.

To conclude, the methods used to control sensory cues are not sufficient to exclude participants' reliance on the sensory cues to perform the task. These methods assume that participants only rely on a single sensory cue at the time. However, already more than two decades ago it was suggested that multiple cues are taken into account to judge numerosity (Allik & Tuulmets, 1991), and this idea was confirmed at the behavioral (Gebuis & Reynvoet, 2012b) and physiological level (Gebuis & Reynvoet, 2012a, 2013; Soltesz & Szucs, 2014). Allik and Tuulmets (1991) theorized that *"the perceptual system is not able to abstract the number per se from all the other stimulus attributes"* (p.303). This because it was observed that the same number of dots distributed differently in space appears different in number. They derived the occupancy model that contrary to the ANS theory can explain such numerosity illusions.

3. ABSTRACT NUMEROSITY

Up to this point in our manuscript we discussed the role of sensory properties in numerosity processing. We argued that normalization as proposed in the ANS theory is unlikely to occur. Instead we argued that the different sensory cues present in the stimuli would directly contribute to the final output. Now it can be questioned: what happens next? The ANS theory suggests that the final output is an abstract number representation. Here we adopt the following definition for abstract number representation: *"Adults can be said to rely on an abstract representation of number if their behaviour depends only on the size of the number involved, not on the specific verbal or non-verbal means of denoting them"* (Dehaene, Dehaene-Lambertz, &

Cohen, 1998 p. 356; See also McCloskey, 1992, p. 497, for a similar definition). In other words: a set of visually presented items, a number of aurally presented beeps or an Arabic number, all would finally evoke the same neural response probably in the IPS, an area suggested to be dedicated to abstract number processing (Dehaene, Piazza, Pinel, & Cohen, 2003; for a review and commentaries see Cohen Kadosh & Walsh, 2009). The idea of abstract number processing has been highly discussed in the numerical cognition literature and yet no consensus has been reached (Cohen Kadosh & Walsh, 2009). Abstract number processing is possible when non-symbolic numerosity such as a set of dots or acoustic beeps are normalized after sensory processing (i.e. processed in a manner independent of the sensory cues). Thus, the normalization stage is essential for abstract number processing and would allow numerosity comparisons across sensory modalities.

A paradigm that is suggested to test abstract number processing is the cross-modal paradigm. Here, two numerosities are presented each in another sensory modality. For instance: 10 acoustic beeps and 15 visually presented dots. Following the theory of the ANS, participants will first estimate the numerosities to come to an abstract number for the two stimuli separately and next compare them. Presentation of both stimuli in the same or in two different modalities should lead to a similar performance when an abstract number representation exists (Barth, Kanwisher, & Spelke, 2003; Tokita & Ishiguchi, 2012). In contrast, poorer performance should result if participants rely on a *single* sensory cue to make the comparison. This is true because there would be no common ground to make the comparison: how to compare the pitch of acoustic beeps to the density of visual dots?

Most cross-modal comparisons studies predominantly focused on small numerosities that are part of the exact number system and not the ANS (Agrillo et al., 2007; Barth et al., 2003; Jordan & Brannon, 2006; Jordan, Brannon, Logothetis, & Ghazanfar, 2005; Nieder, Diester, & Tudusciuc, 2006). Only a few studies examined large approximate numerosities (Arrighi, Togoli, & Burr, 2014; Barth et al., 2003; Burr & Ross, 2008; Jordan, Suanda, & Brannon, 2008; Tokita, Ashitani, & Ishiguchi, 2013; Tokita & Ishiguchi, 2012). Barth et al. (2003) used the cross-modal task to rule out participants' reliance on a single sensory cue. They reasoned that if numerosity processing is based on non-numerical cues, then cross-modal judgments would yield poorer performance in comparison to intra-modal judgments. The results showed that adults did not have a performance cost when comparing visual and auditory stimuli, which excludes their reliance on a *single* perceptual representation. Barth et al. (2003) concluded *"that these findings provide strong evidence that abstract numerosity representations must be constructed from multiple perceptual cues"* (p. 218-219). It should be noted though that large differences between the to be compared numerosity stimuli were used in this task. This could have prevented a performance cost to arise (for a similar reasoning see Tokita et al. 2013). When taking a closer look, a moderate performance cost is visible for ratio 0.5 (see Figure 3 and Figure 7 in Barth et al., 2003) while for the remaining two ratios performance was at ceiling (around 90% accurate) or at chance (around 50% accurate). These results support the idea that the stimuli chosen might have prevented a performance cost to arise in the three numerosity comparisons.

In the study of Barth et al. (2003) the presentation format differed between the visual and the auditory stimulus. The visual stimuli were presented simultaneously while the auditory stimulus was presented sequentially. To account for this confound between stimulus format and modality, Tokita et al. (2013) presented the stimuli of both dimensions sequentially. The results showed that performance is better in the auditory compared to the visual domain in the uni-modal task and that performance is in between for the cross-modal task (see for similar results: Barth et al., 2003; Lechelt, 1975). The difference in performance in the intra-modal conditions suggests that differences in stimulus presentation have an influence on performance but more importantly it increases the difficulty of interpreting cross-modal results. What does a performance in the cross-modal task mean when it lies in between the performances of the two uni-modal tasks? This means that relative to the visual domain there is an improvement while relative to the auditory domain there is a cost in the cross-modal task.

Burr and Ross (2008) investigated the possible existence of abstract numerosity via the procedure of adaptation. They concluded that the presence of adaptation strongly supports the existence of perceptual mechanisms that encode numerical quantity from different senses. To test the level at which the interaction occurred Burr and colleagues performed a cross-modal adaptation study (Arrighi et al., 2014). The results showed that adaptation was selective in spatiotopic rather than retinotopic coordinates. This implies that adaptation occurs at higher levels of analysis. It was reasoned that the different streams of sensory information feed into a common sensory representation of number. It is at this level that the adaptation occurs, not at the separate streams of input. These results are in agreement with recent data showing neural adaptation of numerosity stimuli at higher order areas previously associated with both sensory and numerosity processes (Gebuis et al., 2014; Harvey et al., 2013).

Together, a limited number of cross-modal studies exist, and they do not present a clear result. This could be due to a complicating factor in cross-modal tasks that should first be resolved. Performance in these tasks is dependent on the task utilized: sequential versus simultaneous presentation. In such a case, performance already differs between sensory modalities in the uni-modal condition, which complicates drawing conclusions from the cross-modal condition. Also the effect of the different sensory cues present in both stimuli has not yet been tested. When both visual and auditory stimuli are used in a single task, the difficulty is to make both stimuli equally difficult at a sensory level. It has been shown that it is for instance easier to discriminate a three-fold change in contour length than area (Libertus et al., 2014; A. B. Starr et al., 2013), but how about a 2-fold change in pitch or frequency? And what happens when multiple sensory cues in each domain are manipulated at the same time, an inevitable situation because all sensory cues are intertwined. Clearly, more research is needed to elucidate the mechanisms underlying cross-modal comparisons and the role of methodological differences and sensory cues. Without further research in this direction no firm conclusions can be drawn with respect to the ANS or the sensory-integration theory.

4. OUTLINE OF AN ALTERNATIVE FRAMEWORK: INTEGRATING THE SENSORY CUES

The importance of numerosity judgments for day-to-day decisions cannot be disputed. The question is: what is the nature of the processing mechanism used to arrive at such an estimate? Is it a system that can extract numerosity independent of its sensory cues or rather a sensory-integration system that relies on the sensory cues present in the stimulus? In the previous parts of this manuscript, we constructed on the basis of evolutionary theory, empirical observations and logical reasoning the argument that a sensory-integration explanation appears more likely. Also personal experience or everyday examples appear in agreement with the sensory-integration theory. Take for example the situation where you have to decide which cueing line to take at the passport control. You probably decide this on the basis of the length of the line. Or in the more complex situation, at the supermarket, you probably take into account multiple cues such as the length of the line as well as the amount of groceries that people are buying (almost empty basket or a completely stuffed basket), and possibly whether there are people belonging to the same party (e.g. kids with their parents) (Gevers, Cohen Kadosh, & Gebuis, 2016). You integrate these variables, all of them including several "sensory" cues: *short* line, *little* pile of groceries, *large* groups. In contrast, it seems unlikely that you estimate the true number of active parties and multiply this with the number of groceries they are buying. To the best of our knowledge, we could not come up with a day-to-day example where a comparison between two sets of items has to be made, where it is necessary to attach a specific number to a situation instead of using sensory cues. And this is all the better as recent studies showed that humans are very poor at estimating large quantities (Gebuis & Reynvoet, 2012c; Guillaume et al., 2015; Izard & Dehaene, 2008). Izard and Dehaene (2008) showed that estimates could be off by even 50%. In contrast, humans can make fairly accurate summary statistics of different types of sensory information (Chong & Treisman, 2003, 2005; E. A. Piazza, Sweeny, Wessel, Silver, & Whitney, 2013).

Allik and Tuulmets (1991) already proposed that the reliance on various sensory cues can explain numerosity judgements. Although this idea has been acknowledged many times as a plausible alternative for the ANS theory (Barth et al., 2003; Dehaene, 1992; Izard & Dehaene, 2008; Piazza et al., 2004; Stoianov & Zorzi, 2012), it never attained real support. Only recently, an increasing number of experimental studies showed results that lend support for this idea and allow further expansion of this theory (e.g. Anobile, Cicchini, & Burr, 2014; Durgin, 2008; Gebuis & Reynvoet, 2013; Lourenco, 2015; Sophian, 2007). We therefore believe it is time to reconsider the role of sensory properties in numerosity estimation and to this end propose a first step toward an alternative viewpoint: a sensory-integration theory.

There are two important differences between the sensory-integration theory and the ANS theory. First, the sensory-integration theory *relies on the sensory properties* present in the stimuli to make a comparison between different stimuli. It does not try to remove the sensory cues from the process, as is the done in the normalization stage of the ANS. Second, the *output* of the sensory-integration theory is not an abstract number. At least not according to the definition of abstract number presented above. To recall, a number is suggested abstract when it is only influenced by the numerical value rather than other factors (for further discussion see Cohen Kadosh & Walsh, 2009). On the basis of the large biases that can be induced using different sensory cues, it is clear that two estimates of the same numerosity can highly differ. Both differences between the sensory-integration model and the ANS will be explained in detail in the following paragraphs, and a parallel will be drawn between the sensory-integration theory and the process of conservation as described by Piaget (1952).

4.1. Reliance on sensory cues

The sensory-integration theory suggests that numerosity estimates are *based on the sensory properties* present in the stimuli. We propose that different sensory cues are simultaneously taken into account to come to an integrated representation of number. Three different scenarios can be proposed. In a first scenario, a single integrator is used to make a direct comparison between the sensory properties of two stimuli (Figure 8a and 8b). Logically, direct comparison of the sensory properties only works in the case of intra-modal comparisons (Figure 8a). When cross-modal comparisons have to be made the problem arises that no common ground exists for the direct comparison of the different sensory properties (Figure 8b). In a second scenario, two different integrators are used to make a comparison. The use of different integrators essentially transforms the task from a comparison task into an estimation task. Each stimulus is estimated separately, using its own integrator, before a comparison is made (Figure 8c and 8d). Furthermore, these two separate integrators would be used, regardless of whether the comparison is intra- or cross-modal. With separate integrators, there is no reference to compare each sensory property with; only an estimate can be made about size or frequency with respect to previous stimuli. This lack of a reference will likely induce more variance in both estimates leading to poorer performance compared to a direct comparison using a single integrator only. In a third and last scenario, both mechanisms exist at the same time: a more accurate direct comparison for intra-modal comparisons (figure 8a) while separate integrators are used for cross-modal comparisons (figure 8d).

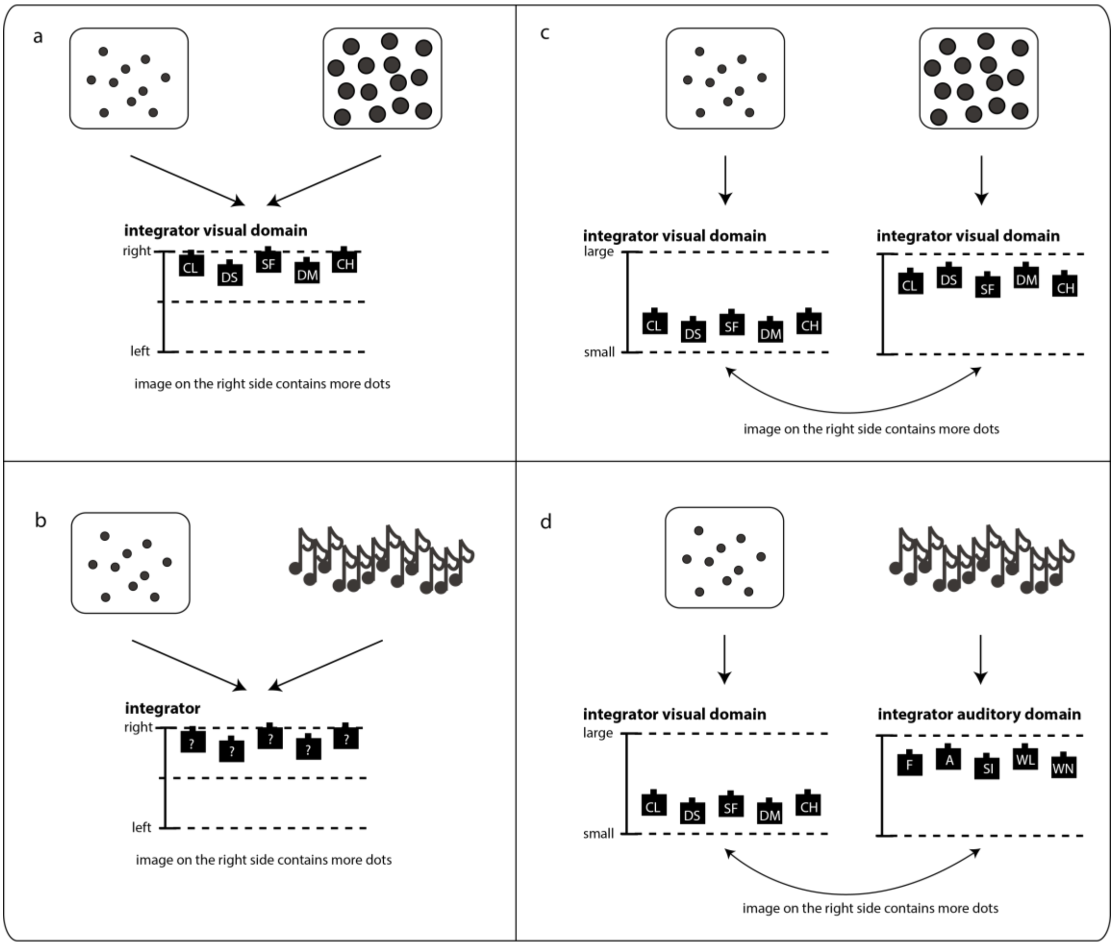

Figure 8. The sensory integrator. a) sensory information from a single modality is directly compared; b) direct comparison of sensory information from different modalities using a single integrator is not possible; c) separate estimates of each stimulus is made even when both stimuli are presented in the same modality; d) separate estimates of each stimulus is made via an integrator specific for the modality of each stimulus. This makes cross-modal comparisons feasible. The abbreviations in the image are as follows: contour length (CL), density (DS), surface (SF), diameter (DM), convex hull (CH), frequency (F), amplitude (A), sound intensity (SI), wavelength (WL), and wave number (WN). Note that the sensory properties shown in the image are the most prominent examples; likely there are more features extracted.

The third scenario fits best with data from various studies showing relatively accurate performance when intra-modal comparisons are performed and relatively poor performance when estimates have to be made. More specifically, in numerosity comparison tasks, a correct response is likely to follow when all sensory cues are correlated with numerosity (see Figure 9a) while an incorrect response is likely to follow when all sensory cues are anti-correlated with numerosity (see Figure 9b), explaining the *congruency effects* observed in the literature (Gebuis, Cohen Kadosh, de Haan, & Henik, 2009; Gebuis & Reynvoet, 2012b; Hurewitz et al., 2006; Rousselle & Noel, 2008; Soltesz et al., 2010). The difference in sensory cues between two images determines the precision with which we can decide which image contains more items. Sensory cue differences that are more prominent will get a stronger weight than those that only differ to a small extent (see Figure 9; compare a-b to c-d). This difference in sensory cues correlates again with numerical distance, when numerical distance decreases, the difference in sensory cues decreases for congruent trials and increases for incongruent trials (see Figure 9; compare a-b to e-f), making it more difficult to judge which image contains more items. This effect nicely explains the *numerical distance effects* observed in the literature (Libertus et al., 2007; Piazza et al., 2004; Sasanguie, Defever, Van den Bussche, & Reynvoet, 2012). Last, more salient sensory cues receive relatively stronger weight. This can explain opposite congruency effects (see Figure 9g and h). For example when better performance is observed for trials where the larger number consisted of smaller dots and vice versa (Gebuis & Van der Smagt, 2011; Ginsburg & Nicholls, 1988; Miller & Baker, 1968; Sophian, 2007). Finally, integration of visual cues can also explain the cancellation of congruency effects (Gebuis & Reynvoet, 2012b). In the latter case, the weights are balanced, therefore there is no clear bias towards either response, and performance will be comparable in both conditions (note that this is not visualised in the image but can be inferred from the almost balanced condition portrayed in Figure 9g and h).

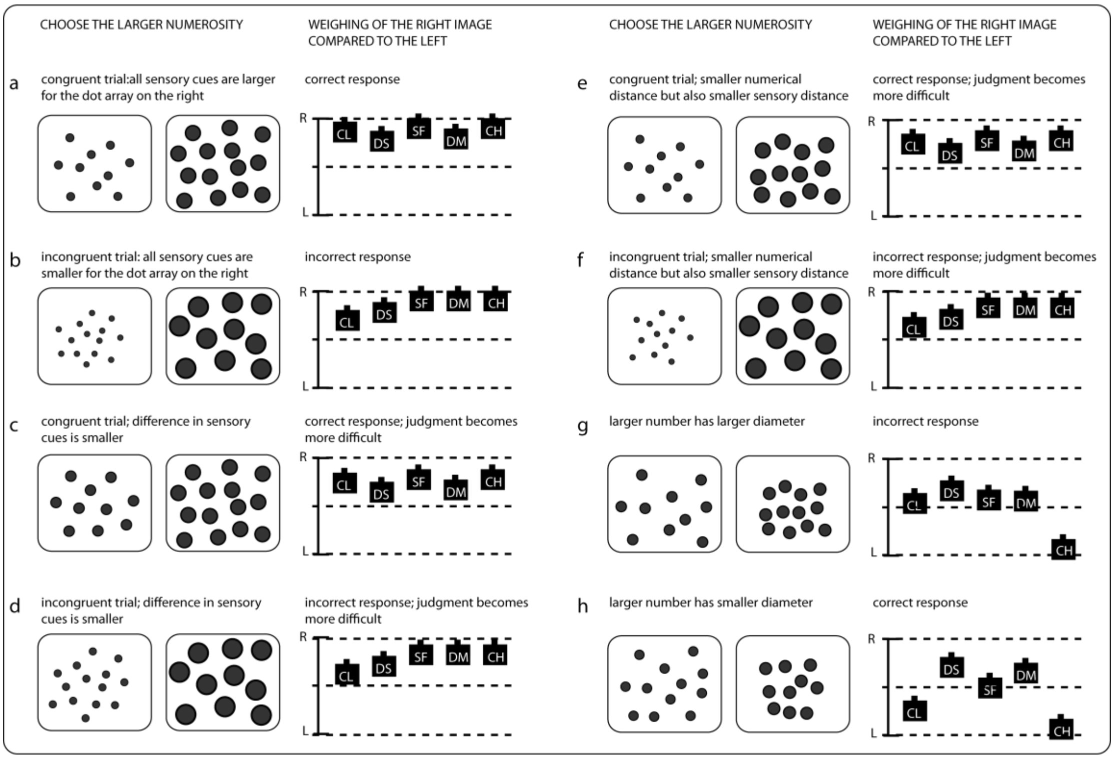

Figure 9. The integration of the sensory cues. a) In congruent trials, all sensory cues of the larger numerosity are larger, resulting in good performance; b) in incongruent trials, all sensory cues of the larger numerosity are smaller, resulting in poor performance; c-d) the difference in the size of the cues determines the weight given to each cue; e-f) this effect scales with numerical distance as numerical distance and sensory difference are correlated; g-h) weighting of sensory cues can explain better performance in partially incongruent trials than congruent trials. Last, when the weights are balanced between both response options, it can explain the cancellation of congruency effects. The abbreviations in the image are as follows: contour length (CL), density (DS), surface (SF), diameter (DM), and convex hull (CH).

Contrary to the ANS theory, the *output* of the sensory-integration theory is not the exact value presented but a sensory representation of number that allows estimating which set of items is numerically larger. This judgment is based on the different sensory cues present in both stimuli. Consequently, this model should perform well in numerosity comparison tasks where sensory cues of two stimuli can be weighted but poorly in the situation where an estimate of a single number has to be made. Contrary to the former, in the latter case the sensory cues are not helpful for estimating numerosity. For example, a large but also a small set of items can consist of a large surface area and a small density while, in the case of numerosity comparison, the stimulus with a larger surface area and larger density is likely to consist of more items. In the case of cross-modal comparisons: the stimulus with large dots and a high density and a large surface is more likely to be numerically larger than the short train of non-frequent beeps. Consistent with the expectations, the literature on numerosity estimation shows that humans are very poor at estimating the number of items presented (Gebuis & Reynvoet, 2012c; Izard & Dehaene, 2008).

The question then is: how do we come to an estimate of numerosity? Numerosity estimation studies showed that sensory cues play a role in estimation (Gebuis & Reynvoet, 2012c) and the overall very poor performance could be improved by showing participants a reference of how much items a single stimulus represents (Izard and Dehaene 2008). Presenting participants with a reference image provides them the opportunity to establish a relationship between the size of the different sensory cues and numerosity. After this relation has been established, each following stimulus can be compared on the basis of the sensory cues of the former. If the sensory cues suggest that the new image is very likely to consist of more items than the numerosity estimate should be increased accordingly. To overcome these range and inter-trial effects that can arise when images are presented sequentially, Krueger (1982) let participants only judge a single image. In line with above-mentioned studies, the results showed a strong underestimation of numerosity that increased with increasing numerosity. Interestingly, the tendency to underestimate was stronger in females than in males. They reasoned that this could be due to the better visual acuity of males under phototopic conditions, and/or the fact that males are better at visuo-spatial tasks. Together, the sensory-integration theory can explain the various results obtained in numerosity comparison and estimation studies.

An issue for future research concerns the relation between the sensory-based system as we proposed here and symbolic number. According to the ANS theory, both numerosities and symbolic number automatically lead to the activation of the same abstract number representation. For the sensory-integration account the relation between numerosities and symbolic number seems to be more task dependent. While estimation tasks force participants to label or to put a number on what is perceived (this is what the task asks you to do), this is not the case for comparison tasks. In comparison tasks it is sufficient to weigh the visual cues to know which cloud of dots is the more numerous. In other words, symbolic number and the proposed sensory based system can be differentiated depending on the type of task used. An interesting paradigm to investigate the relation between symbolic number and numerosities could be to have tasks in which participants are asked to compare a numerosity with a symbolic number (e.g. compare 39 dots with the Arabic digit 25). According to the sensory integration account, the size of the distance effect would be different depending on whether the numerosity presented would be a congruent or an incongruent one.

4.2. A possible relation with conservation

Integrating the sensory cues when comparing two clouds of dots seems to hold strong similarities with the well-known process of conservation (Piaget, 1952). In the conservation of number task, a child is typically shown 2 rows of buttons, each consisting of the same number of buttons. When the buttons of one row is spread over a larger area than the other, children are likely to say that the longer row contains more buttons. Similarly, in the liquid conservation task, a child will typically say that the smaller of two glasses contains *more* water, because the water level in that glass is *higher*. Piaget claims that these children are only able to focus on one feature of a problem at a time and are dominated by their immediate perception of things. In other words, at this stage a child is seemingly not able to inhibit the direct response to a single sensory property.

Children spontaneously rely on inefficient strategies such as a single sensory cue (i.e. the size of a glass instead of the content; for a review see Borst, Aite, & Houdé, 2015) and have to develop the ability to inhibit the incorrect response to the sensory properties of the stimulus (i.e. longer row means more items). With maturation of the prefrontal cortex develops the inhibitory control in children and consequently also the efficiency in solving conservation-like tasks (Houde et al., 2011). These patterns are also visible in numerosity studies: the size of the congruency effect observed in numerosity comparison tasks decreases with increasing age (Gebuis, Cohen Kadosh, et al., 2009; Gilmore et al., 2013; Szucs, Nobes, Devine, Gabriel, & Gebuis, 2013).

A second important factor that plays a role in a conservation task is the notion that children have to understand that an increase in one dimension is compensated for by a decrease in another dimension (Piaget, 1952). This means that in the liquid conservation task, the children should understand that an increase in height is compensated for by a decrease in width while in a numerosity task, children have to understand that a group of large dots scattered over a large area generally contains less dots than a group of small dots scattered over a smaller area at a higher density. The understanding that all these sensory variables are related to each other and all together provide information about the numerosity presented, is the basis for what we termed the ‘integration procedure’ and seems to relate to the concept of an inverse relationship between dimensions as observed in the conservation task.

The ability to judge numerosity does thus depend on the general ability to inhibit previously acquired knowledge (i.e. larger sensory cues means more) (Houde et al., 2011) but also the ability to weigh the different perceptual dimensions (i.e. conservation like processes) (for a different theoretical opinion see Leibovich & Ansari, 2016). Defever, Reynvoet and Gebuis (2013) showed that this ability gradually develops with age. Young children initially respond to a single (or a subset of correlated sensory cues), next they gradually start to inhibit their response to this single sensory cue and start to integrate different sensory cues that do not necessarily behave in the same direction (for example: an increase in total surface area and a decrease in convex hull). They also showed that participants that are better at combining the available sensory cues make more accurate estimates than those less capable of combining different sources of information. Those less proficient in combining different sources of sensory information show a stronger bias towards a single sensory cue and thus larger congruency effects (Defever, Reynvoet, & Gebuis, 2013) and performance decreases when more complex sensory manipulations are applied (Smets et al., 2015; Szucs et al., 2013).

So far we discussed the general role of two different cognitive control abilities in the development of numerosity processes. First, the child needs to inhibit an initial response to the most salient sensory cue. Without response inhibition, the child would consistently respond to the most salient feature in the stimulus. Secondly, cognitive control would be needed to learn to integrate the various sensory cues present in the stimulus. Here cognitive control would be needed to flexibly adapt the weights of the various sensory cues on a trial-by-trial basis. Both these aspects of cognitive control, inhibition and trial-by-trial adaptations are well documented in the literature on conflict tasks like the Stroop task (Stroop, 1953), the Simon task or the flanker task. These tasks, like the numerosity judgment task, reveal a congruency effect with faster responses and less errors to congruent (also termed compatible, e.g. the word red printed in the color red) than to incongruent (also termed incompatible, e.g. the word green printed in the color red) trials. If response inhibition abilities are low, larger congruency effects are expected. Conflict tasks also provide evidence for trial-by-trial adaptations. The most studied phenomenon in this respect is the congruency sequence effect (e.g. CSE, Gratton, Coles, & Donchin, 1992). This is the observation of a smaller congruency effect on trials following an incompatible compared to following a compatible trial. The typical interpretation of the CSE is that incongruent trials increase the level of inhibitory control on the irrelevant information of the next trial (Botvinick, Braver, Barch, Carter, & Cohen, 2001). As such, irrelevant information would have less influence, which translates to smaller congruency effects. Different developmental trajectories have been found for congruency effects (within trial adaptation) and for the CSE (across trial adaptations). The congruency effect seems to undergo developmental changes between pre-school ages and middle-to-late childhood (Checa, Castellanos, Abundis-Gutierrez, & Rosario Rueda, 2014). For instance, while second graders (age: 7.7) demonstrated normal inhibitory functions, this was not yet installed in first grade (age: 6.6) (Iani, Stella, & Rubichi, 2014). In contrast, the CSE is already present at the age of 5-7. This was recently demonstrated in a study comparing the CSE in 5-7 year old children in the Simon task, the Stroop task and the flanker task (Ambrosi, Lemaire, & Blaye, 2016). While task specific influences were observed, a CSE was observed in all three tasks.

The parallel with cognitive control allows developing some straightforward research questions for numerosity judgment tasks. It is clear that congruency effects are also observed when participants perform numerosity judgment tasks. It is however not clear whether trial-by-trial adaptations will be observed as well. More specifically, will congruency effects be less pronounced on trials following incongruent than on trials following congruent trials? If this would be the case, which visual feature would be the one driving the adaptation on the next trial? Would it be the most salient visual feature or would it be the combination of different features that drives the adaptation? Finally, imagine for instance that on the previous trial numerosity was incongruent with dot size, will adaptation specifically target dot size on the next trial, or would it apply to other features as well? Also developmental questions can be asked in a more direct manner. Will the observed dissociation in developmental trajectory between congruency effects and the CSE also be observed in numerosity judgment tasks?

Typically, cognitive control processes are regarded as top-down regulatory processes in which a person with volition (Bugg & Crump, 2012) and consciously (Dehaene & Naccache, 2001) performs goal directed actions to resolve action uncertainty (Norman & Shallice, 1986). However, this does not necessarily need to be the case. In a recent proposal, Abrahamse et al. (2016) provided an alternative framework by grounding cognitive control in associative learning. When performing a task, perceptual, motor and goal representations are co-activated. Because of this co-activation they become bound in a context specific associative network. This network then allows for cognitive control. Imagine the Stroop task again where a participant has to press the left key for the color red and the right key for the color green. Through activation of the appropriate network, the appearance of the color green would then activate the inhibition of the left hand response, thereby installing cognitive control. In brief, the associative learning account on cognitive control is in agreement with the currently proposed integration account for numerosity judgments. That is, different visual features would become bound with one or the other response. Assume for instance a numerosity judgment task with two numerosities presented next to each other, and the participant is asked to press to the side of the largest numerosity. In this situation, a larger convex hull on the left than on the right would activate the left is larger response while inhibiting the right hand response. As the visual cues would differ from trial to trial, the associative network would be differentially activated resulting in flexible performance. Within this view, the participant does not need to be aware of the regularities between the cues. Instead, it is sufficient that the visual cues are part of the associative network to induce cognitive control.

4.3. Conservation and future arithmetic abilities

Previous studies suggested that performance on a numerosity comparison task is related to math achievement (Gilmore et al., 2011; Halberda et al., 2008; Libertus, Feigenson, & Halberda, 2011; Mundy & Gilmore, 2009). If indeed such a relation exists, it is of importance to elucidate the mechanism supporting numerosity processing. If this is the ANS system, the ANS should be trained and improvements in mathematics should be visible at a later age. However, when the integration of sensory properties underlies the ability to strive in numerosity comparison tasks and thus later on in mathematics, training methods should be of a different nature. As mentioned above, children might continue to rely on a single sensory cue instead of integrating the full range of cues. In such a case, training on conservation like tasks with the emphasis on the cognitive control abilities underlying these tasks could be valuable (For a detailed overview of the relation between conservation and numerosity processing/arithmetic, see Gevers, Cohen Kadosh, & Gebuis, 2016). Interestingly, a strong correlation between sensory processes and mathematics has been shown (Lourenco et al., 2012; Tibber et al., 2013) as well as between conservation ability and arithmetic achievement by the end of first grade (Dimitrovsky & Almy, 1975; Dodwell, 1961; Kaufman & Kaufman, 1972). The correlation between conservation abilities and math achievement persisted even when IQ was controlled for (Taloumis, 1979).

Intervention studies are necessary to examine the causal link between arithmetic and conservation and arithmetic. Training on the conservation task did however not improve performance in arithmetic (Steffe, 1976). Children who spontaneously achieved conservation at a younger age benefited more from instructions in arithmetic but this benefit could not be induced by training the children on conservation-like tasks (Bearison, 1975). Still, the number of studies investigating this relation is limited as are the intervention studies that included some form of numerosity training and measures of math achievement (Räsänen et al., 2009; Wilson, Dehaene, et al., 2006; Wilson, Revkin, et al., 2006). In the latter case, the lack of control conditions does not allow for strong conclusions. One recent study did use language processing as a control condition, but did not find an influence of numerosity training on math achievement (Obersteiner, Reiss, & Ufer, 2013). However, caution is needed as this parallel is based on a null finding (e.g. no effect of training numerosity on math achievement). Also and more importantly these studies did not manipulate the sensory cues to stimulate sensory integration processes. If indeed, sensory integration underlies the ability to estimate numerosity a child has to learn to inhibit a response to a single prominent sensory cue such as the longer (but more widely spaced) row of buttons in Piaget’s studies. In such a case, the child has to learn that multiple cues should be taken into account, the length of the row but also the space between the individual buttons. It would therefore be of interest to develop an intervention or training study including congruent as well as incongruent trials. By manipulating a single or multiple sensory cues within each type of stimulus the difficulty of the trials can be gradually increased. In our view (see above), the possible link between numerosity and arithmetic, might be mediated by cognitive control (Houdé, 2009; Pina, Moreno, Cohen Kadosh, & Fuentes, 2015) and the ability to combine different sensory cues processes (Gebuis et al., 2014; Harvey et al., 2013). Both components should therefore be addressed in intervention studies. However, to properly address the relationship between numerosity and math abilities, more studies focusing on the specific strategies used to integrate sensory cues and the development of cognitive control in relation to math achievement are needed.

5. CONCLUSIONS

We provided evidence from comparative, developmental, cognitive, computational, and neuroimaging studies that challenge the existence of an ANS. In doing so we sincerely hope that the combination of the different arguments together will open research plans to explore or construct alternative theoretical proposals to further our understanding on numerosity processing. In our opinion, the existing evidence suggests that the integration of sensory cues present in numerosity stimuli can nicely explain the totality of observed results. This formed the basis for the construction of the sensory-integration theory here proposed. This theory can be used as a starting point to further elucidate the role of sensory cues in numerosity judgments. If validated, the impact of this theory on the field of numerical cognition can be substantial as it includes developmental, cognitive and biological theories concerning typical and atypical numerical development and possible intervention strategies.

ACKNOWLEDGMENTS

We thank Tom Verguts and Devin B. Terhune for helpful comments. RCK is supported by the Welcome Trust (WT88378).

REFERENCES

Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. *Psychol Bull, 142*(7), 693-728. doi: 10.1037/bul0000047

Agrillo, C. (2015). Numerical and Arithmetic abilities in non-primate species. . *In R. Cohen Kadosh & A. Dowker (Eds.), The Handbook of Numerical Cognition: Oxford University Press.*

Agrillo, C., Dadda, M., & Bisazza, A. (2007). Quantity discrimination in female mosquitofish. *Anim Cogn, 10*(1), 63-70. doi: 10.1007/s10071-006-0036-5

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Anim Cogn, 11*(3), 495-503. doi: 10.1007/s10071-008-0140-9

Agrillo, C., Piffer, L., & Bisazza, A. (2010). Large number discrimination by mosquitofish. *PLoS ONE, 5*(12), e15232. doi: 10.1371/journal.pone.0015232

Agrillo, C., Piffer, L., Bisazza, A., & Butterworth, B. (2012). Evidence for two numerical systems that are similar in humans and guppies. *PLoS ONE, 7*(2), e31923. doi: 10.1371/journal.pone.0031923PONE-D-11-20715 [pii]

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Percept Psychophys, 49*(4), 303-314.

Ambrosi, S., Lemaire, P., & Blaye, A. (2016). Do Young Children Modulate Their Cognitive Control? *Exp Psychol, 63*(2), 117-126. doi: 10.1027/1618-3169/a000320

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychol Sci, 25*(1), 265-270. doi: 10.1177/0956797613501520

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nat Rev Neurosci, 9*(4), 278-291. doi: nrn2334 [pii]10.1038/nrn2334

Ansari, D., & Dhital, B. (2006). Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: an event-related functional magnetic resonance imaging study. *J Cogn Neurosci, 18*(11), 1820-1828. doi: 10.1162/jocn.2006.18.11.1820

Ansari, D., Fugelsang, J. A., Dhital, B., & Venkatraman, V. (2006). Dissociating response conflict from numerical magnitude processing in the brain: an event-related fMRI study. *Neuroimage, 32*(2), 799-805. doi: S1053-8119(06)00486-1 [pii]10.1016/j.neuroimage.2006.04.184

Arrighi, R., Togoli, I., & Burr, D. (2014). A generalized sense of number. *Proc Biol Sci, 22*. doi: DOI: 10.1098/rspb.2014.1791

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition, 86*(3), 201-221.

Bearison, D. J. (1975). Induced versus spontaneous attainment of concrete operations and their relationship to school achievement. . *Journal of Educational Psychology, 67*(4), 576-580. doi: http://dx.doi.org/10.1037/h0077147

Beran, M. J. (2007). Rhesus monkeys (Macaca mulatta) enumerate large and small sequentially presented sets of items using analog numerical representations. *J Exp Psychol Anim Behav Process, 33*(1), 42-54. doi: 10.1037/0097-7403.33.1.42

Beran, M. J., Decker, S., Schwartz, A., & Schultz, N. (2011). Monkeys (macaca mulatta and cebus apella) and human adults and children (homo sapiens) compare subsets of moving stimuli based on numerosity. *Front Psychol, 2*, 61. doi: 10.3389/fpsyg.2011.00061

Beran, M. J., Perdue, B. M., & Evans, T. A. (2015). Monkey Mathematical Abilities. In R. Cohen Kadosh & A. Dowker (Eds.), The Handbook of Numerical Cognition: Oxford University Press.

Bianco, I. H., Kampff, A. R., & Engert, F. (2011). Prey capture behavior evoked by simple visual stimuli in larval zebrafish. *Front Syst Neurosci, 5*, 101. doi: 10.3389/fnsys.2011.00101

Bongard, S., & Nieder, A. (2010). Basic mathematical rules are encoded by primate prefrontal cortex neurons. *Proc Natl Acad Sci U S A, 107*(5), 2277-2282. doi: 10.1073/pnas.0909180107

Borst, G., Aite, A., & Houdé, O. (2015). Inhibition of misleading heuristics as a core mechanism for typical cognitive development: evidence from behavioural and brain-imaging studies. *Dev Med Child Neurol, 57*, 21-25. doi: DOI: 10.1111/dmcn.12688

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.

Boysen, S. T., Berntson, G. G., & Mukobi, K. L. (2001). Size matters: impact of item size and quantity on array choice by chimpanzees (Pan troglodytes). *J Comp Psychol, 115*(1), 106-110.

Brannon, E. M., Cantlon, J. F., & Terrace, H. S. (2006). The role of reference points in ordinal numerical comparisons by rhesus macaques (Macaca mulatta). *J Exp Psychol Anim Behav Process, 32*(2), 120-134. doi: 2006-04791-002 [pii]10.1037/0097-7403.32.2.120

Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Dev Sci, 9*(6), F59-64. doi: DESC530 [pii]10.1111/j.1467-7687.2006.00530.x

Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Dev Sci, 10*(6), 770-777.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science, 282*(5389), 746-749.

Bugg, J. M., & Crump, M. J. (2012). In Support of a Distinction between Voluntary and Stimulus-Driven Control: A Review of the Literature on Proportion Congruent Effects. *Front Psychol, 3*, 367. doi: 10.3389/fpsyg.2012.00367

Burr, D., & Ross, J. (2008). A visual sense of number. *Curr Biol, 18*(6), 425-428. doi: S0960-9822(08)00238-8 [pii]10.1016/j.cub.2008.02.052

Butterworth, B. (1999). The Mathematical Brain. *London: Macmillan*.

Butterworth, B. (2005). The development of arithmetical abilities. *J Child Psychol Psychiatry, 46*(1), 3-18. doi: JCPP374 [pii]10.1111/j.1469-7610.2004.00374.x

Cantlon, J. F. (2012). Math, monkeys, and the developing brain. *Proc Natl Acad Sci U S A, 109 Suppl 1*, 10725-10732. doi: 1201893109 [pii]10.1073/pnas.1201893109

Cantlon, J. F., & Brannon, E. M. (2005). Semantic congruity affects numerical judgments similarly in monkeys and humans. *Proc Natl Acad Sci U S A, 102*(45), 16507-16511. doi: 0506463102 [pii]10.1073/pnas.0506463102

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychol Sci, 17*(5), 401-406. doi: PSCI1719 [pii]10.1111/j.1467-9280.2006.01719.x

Cantlon, J. F., & Brannon, E. M. (2007). How much does number matter to a monkey (Macaca mulatta)? *J Exp Psychol Anim Behav Process, 33*(1), 32-41. doi: 2006-23338-004 [pii]10.1037/0097-7403.33.1.32

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

Cantlon, J. F., Cordes, S., Libertus, M. E., & Brannon, E. M. (2009). Comment on "Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures". *Science, 323*(5910), 38; author reply 38. doi: 323/5910/38b [pii]10.1126/science.1164773

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends Cogn Sci, 13*(2), 83-91. doi: S1364-6613(08)00259-3 [pii]10.1016/j.tics.2008.11.007

Cantrell, L., & Smith, L. B. (2013). Set size, individuation, and attention to shape. *Cognition, 126*(2), 258-267. doi: 10.1016/j.cognition.2012.10.007

Checa, P., Castellanos, M. C., Abundis-Gutierrez, A., & Rosario Rueda, M. (2014). Development of neural mechanisms of conflict and error processing during childhood: implications for self-regulation. *Front Psychol, 5*, 326. doi: 10.3389/fpsyg.2014.00326

Chen, Q., & Verguts, T. (2013). Spontaneous summation or numerosity-selective coding? *Front Hum Neurosci, 7*, 886. doi: 10.3389/fnhum.2013.00886

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Res, 43*(4), 393-404. doi: S0042698902005965 [pii]

Chong, S. C., & Treisman, A. (2005). Statistical processing: computing the average size in perceptual groups. *Vision Res, 45*(7), 891-900. doi: S0042-6989(04)00513-9 [pii]10.1016/j.visres.2004.10.004

Clayton, S., Gilmore, C., & Inglis, M. (2015). Dot comparison stimuli are not all alike: The effect of different visual controls on ANS measurement. *Acta Psychol (Amst), 161*, 177-184. doi: 10.1016/j.actpsy.2015.09.007

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychol Sci, 10*(5), 408-411. doi: doi: 10.1111/1467-9280.00177

Cleland, A. A., & Bull, R. (2015). The role of numerical and non-numerical cues in nonsymbolic number processing: Evidence from the line bisection task. *Q J Exp Psychol (Hove), 68*(9), 1844-1859. doi: 10.1080/17470218.2014.994537

Cohen Kadosh, R., Bahrami, B., Walsh, V., Butterworth, B., Popescu, T., & Price, C. J. (2011). Specialization in the human brain: the case of numbers. *Front Hum Neurosci, 5*, 62. doi: 10.3389/fnhum.2011.00062

Cohen Kadosh, R., Cohen Kadosh, K., Linden, D. E., Gevers, W., Berger, A., & Henik, A. (2007). The brain locus of interaction between number and size: a combined functional magnetic resonance imaging and event-related potential study. *J Cogn Neurosci, 19*(6), 957-970. doi: 10.1162/jocn.2007.19.6.957

Cohen Kadosh, R., Gevers, W., & Notebaert, W. (2011). Sequential analysis of the numerical Stroop effect reveals response suppression. *J Exp Psychol Learn Mem Cogn, 37*(5), 1243-1249. doi: 2011-07804-001 [pii]10.1037/a0023550

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: abstract or not abstract? *Behav Brain Sci, 32*(3-4), 313-328. doi: S0140525X09990938 [pii]10.1017/S0140525X09990938

Cohen, L., & Dehaene, S. (2001). Occam's razor is not a Swiss-army knife: A reply to Pillon and Pesenti. *Cogn Neuropsychol, 18*(3), 285-288. doi: 10.1080/02643290042000189

Cresswell, W., & Quinn, J. L. (2004). Faced with a choice, sparrowhawks more often attack the more vulnerable prey group. *OIKOS, 104*, 71-76.

Dadda, M., Piffer, L., Agrillo, C., & Bisazza, A. (2009). Spontaneous number representation in mosquitofish. *Cognition, 112*(2), 343-348. doi: S0010-0277(09)00118-8 [pii] 10.1016/j.cognition.2009.05.009

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A., & Morgan, M. J. (2012). A common visual metric for approximate number and density. *Proc Natl Acad Sci U S A, 108*(49), 19552-19557. doi: 1113195108 [pii] 10.1073/pnas.1113195108

de Hevia, M. D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition, 110*(2), 198-207. doi: S0010-0277(08)00275-8 [pii]10.1016/j.cognition.2008.11.003

De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *J Exp Child Psychol, 108*(2), 278-292. doi: S0022-0965(10)00177-3 [pii]10.1016/j.jecp.2010.09.003

Defever, E., Reynvoet, B., & Gebuis, T. (2013). Task- and age-dependent effects of visual stimulus properties on children's explicit numerosity judgments. *J Exp Child Psychol, 116*(2), 216-233. doi: S0022-0965(13)00084-2 [pii]10.1016/j.jecp.2013.04.006

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*(1-2), 1-42.

Dehaene, S. (1997). The number sense: How the mind creates mathematics. *Oxford University Press, Penguin press, New York, Cambridge (UK)*.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: a logarithmic mental number line. *Trends Cogn Sci, 7*(4), 145-147.

Dehaene, S. (2009). Origins of mathematical intuitions: the case of arithmetic. *Ann N Y Acad Sci, 1156*, 232-259. doi: NYAS04469 [pii]10.1111/j.1749-6632.2009.04469.x

Dehaene, S., & Brannon, E. M. (2010). Space, time, and number: a Kantian research program. *Trends Cogn Sci, 14*(12), 517-519. doi: DOI: http://dx.doi.org/10.1016/j.tics.2010.09.009

Dehaene, S., & Changeux, J.P. (1993). Development of elementary numerical abilities: a neuronal model. *Journal of Cognitive Neuroscience, 5*(4), 390-407.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends Neurosci, 21*(8), 355-361. doi: S0166223698012636 [pii]

Dehaene, S., Izard, V., & Piazza, M. (2005). Control over non-numerical parameters in numerosity experiments. <http://www.unicog.org/docs/DocumentationDotsGeneration.doc%3E.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*(1-2), 1-37.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487-506.

Dimitrovsky, L., & Almy, M. (1975). Early conservation as a predictor of arithmetic achievement. *The Journal of Psychology, 91*, 65-70.

Dodwell, P.C. (1961). Children's understanding of number concepts: Characteristics of an individual and of a group test. . *Can J of Psychol, 15*, 29-36.

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Curr Biol, 18*(18), R855-856; author reply R857-858. doi: S0960-9822(08)00962-7 [pii]10.1016/j.cub.2008.07.053

Emmerton, J., & Renner, J. C. (2009). Local rather than global processing of visual arrays in numerosity discrimination by pigeons (Columba livia). *Anim Cogn, 12*(3), 511-526. doi: 10.1007/s10071-009-0212-5

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends Cogn Sci, 8*(7), 307-314. doi: 10.1016/j.tics.2004.05.002 S1364661304001317 [pii]

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Dev Perspect, 7*(2), 74-79. doi: 10.1111/cdep.12019

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (Macaca mulatta) spontaneously compute addition operations over large numbers. *Cognition, 97*(3), 315-325. doi: S0010-0277(04)00182-9 [pii]10.1016/j.cognition.2004.09.004

Franks, N. R., Dornhaus, A., Metherell, B. G., Nelson, T. R., Lanfear, S. A., & Symes, W. S. (2006). Not everything that counts can be counted: ants use multiple metrics for a single nest trait. *Proc Biol Sci, 273*(1583), 165-169. doi: 10.1098/rspb.2005.3312

Frith, C.D., & Frith, U. (1972). The solitaire illusion: An illusion of numerosity. *Perception and Psychophysics, 11*(6), 409-410.

Frommen, J. G., Mehlis, M., & Bakker, T. C. (2009). Predator-inspection behaviour in female three-spined sticklebacks Gasterosteus aculeatus is associated with status of gravidity. *J Fish Biol, 75*(8), 2143-2153. doi: JFB2408 [pii] 10.1111/j.1095-8649.2009.02408.x

Gallistel, C. R., & Gelman, I. I. (2000). Non-verbal numerical cognition: from reals to integers. *Trends Cogn Sci, 4*(2), 59-65. doi: S1364661399014242 [pii]

Gebuis, T., Cohen Kadosh, R., de Haan, E., & Henik, A. (2009). Automatic quantity processing in 5-year olds and adults. *Cogn Process, 10*(2), 133-142. doi: 10.1007/s10339-008-0219-x

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; discussion 253-245. doi: S0010-0277(10)00221-0 [pii]10.1016/j.cognition.2010.09.008

Gebuis, T., Gevers, W., & Cohen Kadosh, R. (2014). Topographic representation of high-level cognition: numerosity or sensory processing? *Trends Cogn Sci, 18*(1), 1-3. doi: 10.1016/j.tics.2013.10.002

Gebuis, T., Herfs, I. K., Kenemans, J. L., de Haan, E. H., & van der Smagt, M. J. (2009). The development of automated access to symbolic and non-symbolic number knowledge in children: an ERP study. *Eur J Neurosci, 30*(10), 1999-2008. doi: EJN6994 [pii]10.1111/j.1460-9568.2009.06994.x

Gebuis, T., Kenemans, J. L., de Haan, E. H., & van der Smagt, M. J. (2010). Conflict processing of symbolic and non-symbolic numerosity. *Neuropsychologia, 48*(2), 394-401. doi: S0028-3932(09)00383-2 [pii]10.1016/j.neuropsychologia.2009.09.027

Gebuis, T., & Reynvoet, B. (2011). Generating non-symbolic number stimuli. *Behav Res Methods, 43*(4), 981-986. doi: 10.3758/s13428-011-0097-5

Gebuis, T., & Reynvoet, B. (2012a). Continuous visual properties explain neural responses to nonsymbolic number. *Psychophysiology, 49*(11), 1649-1659. doi: 10.1111/j.1469-8986.2012.01461.x

Gebuis, T., & Reynvoet, B. (2012b). The interplay between nonsymbolic number and its continuous visual properties. *J Exp Psychol Gen, 141*(4), 642-648. doi: 2011-25898-001 [pii]10.1037/a0026218

Gebuis, T., & Reynvoet, B. (2012c). The role of visual information in numerosity estimation. *PLoS ONE, 7*(5), e37426. doi: 10.1371/journal.pone.0037426PONE-D-11-22563 [pii]

Gebuis, T., & Reynvoet, B. (2013). The neural mechanisms underlying passive and active processing of numerosity. *Neuroimage, 70C*, 301-307. doi: S1053-8119(12)01234-7 [pii]10.1016/j.neuroimage.2012.12.048

Gebuis, T., & Reynvoet, B. (2015). Number representations and their relation with mathematical ability. *In: R.C. Kadosh, & A. Dowker (Eds.), Oxford handbook of numerical cognition, Oxford University Press.*

Gebuis, T., & Van der Smagt, M. J. (2011). False approximations of the approximate number system? *PLoS ONE, 6*(10), e25405.

Gevers, W., Cohen Kadosh, R., & Gebuis, T. (2016). The sensory integration theory: an alternative to the Approximate Number System. *In A. Henik (Ed.), Continuous Issues in Numerical Cognition: Elsevier.*

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. doi: 10.1371/journal.pone.0067374 PONE-D-12-39976 [pii]

Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *Q J Exp Psychol (Colchester), 64*(11), 2099-2109. doi: 10.1080/17470218.2011.574710

Ginsburg, N. (1991). Numerosity estimation as a function of stimulus organization. *Perception, 20*(5), 681-686.

Ginsburg, N., & Nicholls, A. (1988). Perceived numerosity as a function of item size. *Percept Mot Skills, 67*(2), 656-658.

Gomez-Laplaza, L. M., & Gerlai, R. (2012). Activity Counts: The Effect of Swimming Activity on Quantity Discrimination in Fish. *Front Psychol, 3*, 484. doi: doi: 10.3389/fpsyg.2012.00484

Gomez-Laplaza, L. M., & Gerlai, R. (2013). Quantification abilities in angelfish (Pterophyllum scalare): the influence of continuous variables. *Anim Cogn, 16*(3), 373-383. doi: 10.1007/s10071-012-0578-7

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J Exp Psychol Gen, 121*(4), 480-506.

Guillaume, M., Gevers, W., & Content, A. (2015). Assessing the Approximate Number System: no relation between numerical comparison and estimation tasks. *Psychol Res*. doi: 10.1007/s00426-015-0657-x

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Dev Psychol, 44*(5), 1457-1465. doi: 2008-12114-023 [pii] 10.1037/a0012682

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. *Proc Natl Acad Sci U S A, 109*(28), 11116-11120. doi: 1200196109 [pii]10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*(7213), 665-668. doi: nature07246 [pii]10.1038/nature07246

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science, 341*(6150), 1123-1126. doi: 341/6150/1123 [pii]10.1126/science.1239052

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proc Biol Sci, 267*(1445), 829-833. doi: 10.1098/rspb.2000.1078

Hauser, M. D., Tsao, F., Garcia, P., & Spelke, E. S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton-top tamarins. *Proc Biol Sci, 270*(1523), 1441-1446. doi: 10.1098/rspb.2003.2414

Henik, A., & Tzelgov, J. (1982). Is three greater than five: the relation between physical and semantic size in comparison tasks. *Mem Cognit, 10*(4), 389-395.

Hilton, G. F., Cresswell, W., & Ruxton, G. D. (1999). Intraflock variation in the speed of escape-flight response on attack by an avian predator. *Behavioral Ecology, 10*(4), 391-395. doi: doi: 10.1093/beheco/10.4.391

Hilton, G. F., Ruxton, G. D., & Cresswell, W. . (1999). Choice of Foraging Area with Respect to Predation Risk in Redshanks: The Effects of Weather and Predator Activity. *OIKOS, 87*(2), 295-302. doi: DOI: 10.2307/3546744

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: the numerical distance effect and individual differences in children's mathematics achievement. *J Exp Child Psychol, 103*(1), 17-29. doi: S0022-0965(08)00052-0 [pii]10.1016/j.jecp.2008.04.001

Holloway, I. D., Price, G. R., & Ansari, D. (2010). Common and segregated neural pathways for the processing of symbolic and nonsymbolic numerical magnitude: an fMRI study. *Neuroimage, 49*(1), 1006-1017. doi: S1053-8119(09)00861-1 [pii]10.1016/j.neuroimage.2009.07.071

Houdé, O. (2009). Abstract after all? Abstraction through inhibition in children and adults. *Behav Brain Sci, 32*(339-340).

Houde, O., Pineau, A., Leroux, G., Poirel, N., Perchey, G., Lanoe, C., . . . Mazoyer, B. (2011). Functional magnetic resonance imaging study of Piaget's conservation-of-number task in preschool and school-age children: a neo-Piagetian approach. *J Exp Child Psychol, 110*(3), 332-346. doi: doi: 10.1016/j.jecp.2011.04.008

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proc Natl Acad Sci U S A, 103*(51), 19599-19604. doi: 0609485103 [pii]10.1073/pnas.0609485103

Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: an electrophysiological investigation of small and large number representations. *J Cogn Neurosci, 21*(6), 1039-1053. doi: 10.1162/jocn.2009.21090

Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: evidence for two core systems underlying numerical cognition. *Dev Sci, 14*(2), 360-371. doi: 10.1111/j.1467-7687.2010.00987.x

Iani, C., Stella, G., & Rubichi, S. (2014). Response inhibition and adaptations to response conflict in 6- to 8-year-old children: evidence from the Simon effect. *Atten Percept Psychophys, 76*(4), 1234-1241. doi: 10.3758/s13414-014-0656-9

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychon Bull Rev, 18*(6), 1222-1229. doi: 10.3758/s13423-011-0154-1

Inglis, M., & Gilmore, C. (2013). Sampling from the mental number line: How are approximate number system representations formed? *Cognition, 129*(1), 63-69. doi: S0010-0277(13)00115-7 [pii] 10.1016/j.cognition.2013.06.003

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106*(3), 1221-1247. doi: S0010-0277(07)00156-4 [pii] 10.1016/j.cognition.2007.06.004

Izard, V., Dehaene-Lambertz, G., & Dehaene, S. (2008). Distinct cerebral pathways for object identity and number in human infants. *PLoS Biol, 6*(2), e11. doi: 07-PLBI-RA-2266 [pii] 10.1371/journal.pbio.0060011

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc Natl Acad Sci U S A, 106*(25), 10382-10385. doi: 0812142106 [pii] 10.1073/pnas.0812142106

Jones, S. M., & Brannon, E. M. (2012). Prosimian primates show ratio dependence in spontaneous quantity discriminations. *Front Psychol, 3*, 550. doi: 10.3389/fpsyg.2012.00550

Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proc Natl Acad Sci U S A, 103*(9), 3486-3489.

Jordan, K. E., Brannon, E. M., Logothetis, N. K., & Ghazanfar, A. A. (2005). Monkeys match the number of voices they hear to the number of faces they see. *Curr Biol, 15*(11), 1034-1038. doi: S0960-9822(05)00482-3 [pii] 10.1016/j.cub.2005.04.056

Jordan, K. E., Suanda, S. H., & Brannon, E. M. (2008). Intersensory redundancy accelerates preverbal numerical competence. *Cognition, 108*(1), 210-221. doi: S0010-0277(07)00306-X [pii] 10.1016/j.cognition.2007.12.001

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal Representation of Numeric Quantities by Brown Capuchin Monkeys (Cebus apella). . *Journal of Experimental Psychology: Animal Behavior Processes, 31*(1), 79-94. doi: http://dx.doi.org/10.1037/0097-7403.31.1.79

Kaufman, A.S., & Kaufman, N.L. (1972). Tests build from Piaget's and Gesell's tasks as predictiors of first grade achievement. *Child Dev, 43*, 521-535.

Krueger, L. E. (1982). Single judgments of numerosity. *Percept Psychophys, 31*(2), 175-182.

Krusche, P., Uller, C., & Dicke, U. (2010). Quantity discrimination in salamanders. *J Exp Biol, 213*(11), 1822-1828. doi: 213/11/1822 [pii] 10.1242/jeb.039297

Lansbergen, M. M., & Kenemans, J. L. (2008). Stroop interference and the timing of selective response activation. *Clin Neurophysiol, 119*(10), 2247-2254. doi: S1388-2457(08)00852-3 [pii] 10.1016/j.clinph.2008.07.218

Lechelt, E. C. (1975). Temporal numerosity discrimination: intermodal comparisons revisited. *Br J Psychol, 66*(1), 101-108.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nat Neurosci, 11*(4), 404-409. doi: nn2065 [pii] 10.1038/nn2065

Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Can J Exp Psychol, 70*(1), 12-23. doi: 10.1037/cep0000070

Leibovich, T., & Henik, A. (2013). Magnitude processing in non-symbolic stimuli. *Front Psychol, 4*, 375. doi: 10.3389/fpsyg.2013.00375

Leibovich, T., & Henik, A. (2014). Comparing performance in discrete and continuous comparison tasks. *Q J Exp Psychol (Hove), 67*(5), 899-917. doi: 10.1080/17470218.2013.837940

Lettvin, J.Y., Maturana, H.R., McCulloch, W.S., & Pitts, W.H. (1959). What the frog’s eye tells the frog’s brain. *Proc. Inst. Radio Eng., 47*, 1940-1951.

Libertus, M. E., & Brannon, E. M. (2009). Behavioral and Neural Basis of Number Sense in Infancy. *Curr Dir Psychol Sci, 18*(6), 346-351. doi: 10.1111/j.1467-8721.2009.01665.x

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Dev Sci, 13*(6), 900-906. doi: 10.1111/j.1467-7687.2009.00948.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Dev Sci, 14*(6), 1292-1300. doi: 10.1111/j.1467-7687.2011.01080.x

Libertus, M. E., Starr, A., & Brannon, E. M. (2014). Number trumps area for 7-month-old infants. *Dev Psychol, 50*(1), 108-112. doi: 2013-15560-001 [pii] 10.1037/a0032986

Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behav Brain Funct, 3*, 1. doi: 1744-9081-3-1 [pii] 10.1186/1744-9081-3-1

Lima, L., & Dill, M. (1990). Behavioral decisions made under the risk of predation: a review and prospectus. *Canadian Journal of Zoology, 68*(4), 619-640.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychol Sci, 14*(5), 396-401.

Lourenco, S. F. (2015). On the relation between numerical and non-numerical magnitudes: Evidence for a general magnitude system. *Evolutionary origins and early development of number processing, volume 1 edited by David C. Geary, Daniel B. Berch, Kathleen Mann Koepke, 01/2015: chapter On the relation between numerical and non-numerical magnitudes: Evidence for a general magnitude system: pages 145-174; Elsevier.*

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proc Natl Acad Sci U S A, 109*(46), 18737-18742. doi: 10.1073/pnas.1207212109

Mazza, V., Pagano, S., & Caramazza, A. (2013). Multiple object individuation and exact enumeration. *J Cogn Neurosci, 25*(5), 697-705. doi: 10.1162/jocn_a_00349

Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS ONE, 6*(9), e23749. doi: 10.1371/journal.pone.0023749 PONE-D-11-09239 [pii]

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: evidence from acquired dyscalculia. *Cognition, 44*(1-2), 107-157.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *J Exp Psychol Anim Behav Process, 9*(3), 320-334.

Miller, A. L., & Baker, R. A. (1968). The effects of shape, size, heterogeneity, and instructional set on the judgment of visual number. *Am J Psychol, 81*(1), 83-91.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: is number one of them? *Psychol Bull, 128*(2), 278-294.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*(109), 1519-1520.

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *J Exp Child Psychol, 103*(4), 490-502. doi: S0022-0965(09)00035-6 [pii] 10.1016/j.jecp.2009.02.003

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science, 313*(5792), 1431-1435. doi: 313/5792/1431 [pii] 10.1126/science.1130308

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297*(5587), 1708-1711. doi: 10.1126/science.1072493297/5587/1708 [pii]

Nieder, A., & Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J Neurosci, 27*(22), 5986-5993. doi: 27/22/5986 [pii] 10.1523/JNEUROSCI.1056-07.2007

Nieder, A., & Miller, E. K. (2004). Analog numerical representations in rhesus monkeys: evidence for parallel processing. *J Cogn Neurosci, 16*(5), 889-901. doi: 10.1162/089892904970807

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. *In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory (Vol. 4, pp. 1–18). New York, NY: Plenum Press.*

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance !rst-grade students' basic number processing and arithmetic skills. *Learning and instruction*.

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychol Sci, 24*(10), 2013-2019. doi: 0956797613482944 [pii]10.1177/0956797613482944

Piaget, J. (1952). The Child's Conception of Number. *Routledge & KeganPaul, London*.

Piaget, J. (1965). *The child's conception of number*. New York: W. Norton Company & Inc.

Piazza. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends Cogn Sci, 14*(12), 542-551. doi: S1364-6613(10)00215-9 [pii]10.1016/j.tics.2010.09.008

Piazza, 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*(1), 33-41. doi: S0010-0277(10)00076-4 [pii]10.1016/j.cognition.2010.03.012

Piazza, & Izard, V. (2009). How humans count: numerosity and the parietal cortex. *Neuroscientist, 15*(3), 261-273. doi: 15/3/261 [pii]10.1177/1073858409333073

Piazza, Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*(3), 547-555. doi: S0896627304006786 [pii]10.1016/j.neuron.2004.10.014

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). Humans Use Summary Statistics to Perceive Auditory Sequences. *Psychol Sci*. doi: 0956797612473759 [pii]10.1177/0956797612473759

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal approximate number system. *Psychol Sci, 24*(6), 1037-1043. doi: 0956797612464057 [pii]10.1177/0956797612464057

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron, 53*(2), 293-305. doi: S0896-6273(06)00989-5 [pii]10.1016/j.neuron.2006.11.022

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*(5695), 499-503. doi: 306/5695/499 [pii]10.1126/science.1102085

Pina, Violeta, Moreno, Alejandro Castillo, Cohen Kadosh, Roi, & Fuentes, Luis J. (2015). Intentional and Automatic Numerical Processing as Predictors of Mathematical Abilities in Primary School Children. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.00375

Preuss, S. J., Trivedi, C. A., Vom Berg-Maurer, C. M., Ryu, S., & Bollmann, J. H. (2014). Classification of Object Size in Retinotectal Microcircuits. *Curr Biol*. doi: S0960-9822(14)01132-4 [pii]10.1016/j.cub.2014.09.012

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 Psychol (Amst), 140*(1), 50-57. doi: S0001-6918(12)00034-0 [pii]10.1016/j.actpsy.2012.02.008

Räsänen, P. J., Wilson, A. J., Aunio, P. , & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development, 24*(4), 450-472.

Roggeman, C., Santens, S., Fias, W., & Verguts, T. (2011). Stages of nonsymbolic number processing in occipitoparietal cortex disentangled by fMRI adaptation. *J Neurosci, 31*(19), 7168-7173. doi: 31/19/7168 [pii]10.1523/JNEUROSCI.4503-10.2011

Roitman, J. D., Brannon, E. M., & Platt, M. L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol, 5*(8), e208. doi: 06-PLBI-RA-2389 [pii]10.1371/journal.pbio.0050208

Rousselle, L., & Noel, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition, 102*(3), 361-395.

Rousselle, L., & Noel, M. P. (2008). The development of automatic numerosity processing in preschoolers: evidence for numerosity-perceptual interference. *Dev Psychol, 44*(2), 544-560. doi: 2008-02379-021 [pii]10.1037/0012-1649.44.2.544

Rubinsten, O., & Henik, A. (2009). Developmental dyscalculia: heterogeneity might not mean different mechanisms. *Trends Cogn Sci, 13*(2), 92-99. doi: S1364-6613(08)00264-7 [pii]10.1016/j.tics.2008.11.002

Rugani, R., Cavazzana, A., Vallortigara, G., & Regolin, L. (2013). One, two, three, four, or is there something more? Numerical discrimination in day-old domestic chicks. *Anim Cogn, 16*(4), 557-564. doi: 10.1007/s10071-012-0593-8

Rugani, R., Regolin, L., & Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *J Exp Psychol Anim Behav Process, 34*(3), 388-399. doi: 10.1037/0097-7403.34.3.388

Rugani, R., Regolin, L., & Vallortigara, G. (2011). Summation of large numerousness by newborn chicks. *Front Psychol, 2*, 179. doi: 10.3389/fpsyg.2011.00179

Rugani, R., Vallortigara, G., & Regolin, L. (2013). Numerical abstraction in young domestic chicks (Gallus gallus). *PLoS ONE, 8*(6), e65262. doi: 10.1371/journal.pone.0065262

Rugani, R., Vallortigara, G., & Regolin, L. (2014). From small to large: numerical discrimination by young domestic chicks (Gallus gallus). *Journal of Comparative Psychology, 128*(2), 163-171. doi: doi: 10.1037/a0034513

Santens, S., Roggeman, C., Fias, W., & Verguts, T. (2010). Number Processing Pathways in Human Parietal Cortex. *Cereb Cortex*. doi: bhp080 [pii] 10.1093/cercor/bhp080

Santens, S., & Verguts, T. (2011). The size congruity effect: is bigger always more? *Cognition, 118*(1), 94-110. doi: S0010-0277(10)00251-9 [pii] 10.1016/j.cognition.2010.10.014

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. . *British Journal of Developmental Psychology*.

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2012). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychol (Amst), 136*(1), 73-80. doi: S0001-6918(10)00210-6 [pii] 10.1016/j.actpsy.2010.10.004

Silver, M. A., & Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. *Trends Cogn Sci, 13*(11), 488-495. doi: S1364-6613(09)00173-9 [pii] 10.1016/j.tics.2009.08.005

Smets, K., Gebuis, T., Defever, E., & Reynvoet, B. (2014). Concurrent validity of approximate number sense tasks in adults and children. *Acta Psychol (Amst), 150*, 120-128. doi: 10.1016/j.actpsy.2014.05.001

Smets, K., Gebuis, T., & Reynvoet, B. (2013). Comparing the neural distance effect derived from the non-symbolic comparison and the same-different task. *Front Hum Neurosci, 7*, 28. doi: 10.3389/fnhum.2013.00028

Smets, K., Sasanguie, D., Szucs, D., & Reynvoet, B. (2015). The effect of different methods to construct non-symbolic stimuli in numerosity estimation and comparison. *Journal of Cognitive Psychology, 27*(3), 310-325. doi: 10.1080/20445911.2014.996568

Smith, J. M. (1974). The theory of games and the evolution of animal conflicts. *J Theor Biol, 47*(1), 209-221.

Soltesz, F., & Szucs, D. (2014). Neural adaptation to non-symbolic number and visual shape: An electrophysiological study. *Biol Psychol, 103C*, 203-211. doi: S0301-0511(14)00207-5 [pii] 10.1016/j.biopsycho.2014.09.006

Soltesz, F., Szucs, D., & Szucs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: A developmental study. *Behav Brain Funct, 6*, 13. doi: 1744-9081-6-13 [pii] 10.1186/1744-9081-6-13

Sophian, C. (2007). Measuring spatial factors in comparative judgments about large numerosities. . *In D. Schmorrow & L. Reeves (Eds.), Foundations of augmented cognition: Third International Conference. Secaucus (pp. 157–165). NJ: Springer.*

Sophian, C., & Chu, Y. (2008). How do people apprehend large numerosities? *Cognition, 107*(2), 460-478. doi: S0010-0277(07)00268-5 [pii] 10.1016/j.cognition.2007.10.009

Stancher, G., Sovrano, V. A., Potrich, D., & Vallortigara, G. (2013). Discrimination of small quantities by fish (redtail splitfin, Xenotoca eiseni). *Anim Cogn, 16*(2), 307-312. doi: 10.1007/s10071-012-0590-y

Starr, A. B., Libertus, M. E., & Brannon, E. M. (2013). Infants Show Ratio-dependent Number Discrimination Regardless of Set Size. *Infancy, 18*(6). doi: 10.1111/infa.12008

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proc Natl Acad Sci U S A*. doi: 1302751110 [pii] 10.1073/pnas.1302751110

Steffe, L. P. (1976). On a model for teaching young children mathematics. . *In A.R. Osborne (Ed.), Models for learning mathematics, ERIC/SMEAC, columbus, ohio.*

Stoianov, I., & Zorzi, M. (2012). Emergence of a 'visual number sense' in hierarchical generative models. *Nat Neurosci, 15*(2), 194-196. doi: nn.2996 [pii] 10.1038/nn.2996

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental Psychology, 28*, 643-662.

Szucs, D., Nobes, A., Devine, A., Gabriel, F.C., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children *Frontiers in Developmental Psychology*.

Szucs, D., & Soltesz, F. (2008). The interaction of task-relevant and task-irrelevant stimulus features in the number/size congruency paradigm: an ERP study. *Brain Res, 1190*, 143-158. doi: S0006-8993(07)02658-3 [pii] 10.1016/j.brainres.2007.11.010

Szucs, D., Soltesz, F., Jarmi, E., & Csepe, V. (2007). The speed of magnitude processing and executive functions in controlled and automatic number comparison in children: an electro-encephalography study. *Behav Brain Funct, 3*, 23.

Taloumis, T. (1979). Scores on Piagetian area tasks as predictors of achievement in mathematics over a four-year period. *Journal of Researcher in Mathematics Education, 10*, 120-134.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proc Natl Acad Sci U S A, 95*(13), 7836-7841.

Tibber, M. S., Manasseh, G. S., Clarke, R. C., Gagin, G., Swanbeck, S. N., Butterworth, B., . . . Dakin, S. C. (2013). Sensitivity to numerosity is not a unique visuospatial psychophysical predictor of mathematical ability. *Vision Res, 89*, 1-9. doi: 10.1016/j.visres.2013.06.006

Tokita, M., Ashitani, Y., & Ishiguchi, A. (2013). Is approximate numerical judgment truly modality-independent? Visual, auditory, and cross-modal comparisons. *Atten Percept Psychophys, 75*(8), 1852-1861. doi: 10.3758/s13414-013-0526-x

Tokita, M., & Ishiguchi, A. (2012). Behavioral evidence for format-dependent processes in approximate numerosity representation. *Psychon Bull Rev, 19*(2), 285-293. doi: 10.3758/s13423-011-0206-6

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Mem Cognit, 20*(6), 727-735.

Ullas Karanth, K., & Sunquist, M.E. (1995). Prey selection by tiger, leopard and dhole in tropical forests. *Journal of Animal Ecology, 64*, 439-450.

Uller, C., Jaeger, R., Guidry, G., & Martin, C. (2003). Salamanders (Plethodon cinereus) go for more: rudiments of number in an amphibian. *Anim Cogn, 6*(2), 105-112. doi: 10.1007/s10071-003-0167-x

van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Percept Psychophys, 32*(2), 163-170.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: a neural model. *J Cogn Neurosci, 16*(9), 1493-1504. doi: 10.1162/0898929042568497

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychon Bull Rev, 12*(1), 66-80.

Viswanathan, P., & Nieder, A. (2013). Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. *Proc Natl Acad Sci U S A, 110*(27), 11187-11192. doi: 1308141110 [pii] 10.1073/pnas.1308141110

Vonk, J., & Beran, M. J. (2012). Bears "Count" Too: Quantity Estimation and Comparison in Black Bears (Ursus Americanus). *Anim Behav, 84*(1), 231-238. doi: 10.1016/j.anbehav.2012.05.001

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: the psychophysics of number representation. *Psychol Sci, 10*(2), 130-137.

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behav Brain Funct, 2*, 19.

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behav Brain Funct, 2*, 20.

Wood, J. N., Hauser, M. D., Glynn, D. D., & Barner, D. (2008). Free-ranging rhesus monkeys spontaneously individuate and enumerate small numbers of non-solid portions. *Cognition, 106*(1), 207-221. doi: S0010-0277(07)00019-4 [pii]10.1016/j.cognition.2007.01.004

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: numerical discrimination and its signature limits. *Dev Sci, 8*(2), 173-181.

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

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Dev Sci, 8*(1), 88-101.