Acta Psychologica 150 (2014) 120–128

Contents lists available at ScienceDirect

Acta Psychologica

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

Concurrent validity of approximate number sense tasks in adults and children

Karolien Smets [a,*], Titia Gebuis [a], Emmy Defever [a,b], Bert Reynvoet [a,b]

[a] Experimental Psychology, KU Leuven, 3000 Leuven, Belgium
[b] Faculty of Psychology and Educational Sciences, KU Leuven @ Kulak, 8500 Kortrijk, Belgium

ARTICLE INFO

Article history:
Received 25 October 2013
Received in revised form 30 April 2014
Accepted 5 May 2014
Available online 27 May 2014

PsycINFO codes:
23402260

Keywords:
Number processing
Change detection
Comparison
Same-different
ANS acuity
Validity

ABSTRACT

Reasoning with non-symbolic numerosities is suggested to be rooted in the Approximate Number System (ANS) and evidence pointing to a relationship between the acuity of this system and mathematics is available. In order to use the acuity of this ANS as a screening instrument to detect future math problems, it is important to model ANS acuity over development. However, whether ANS acuity and its development have been described accurately can be questioned. Namely, different tasks were used to examine the developmental trajectory of ANS acuity and studies comparing performances on these different tasks are scarce. In the present study, we examined whether different tasks designed to measure the acuity of the ANS are comparable and lead to related ANS acuity measures (i.e., the concurrent validity of these tasks). We contrasted the change detection task, which is used in infants, with tasks that are more commonly used in older children and adults (i.e., comparison and same-different tasks). Together, our results suggest that ANS acuity measures obtained with different tasks are not related. This poses serious problems for the comparison of ANS acuity measures derived from different tasks and thus for the establishment of the developmental trajectory of ANS acuity.

Humans, but also non-human species, are equipped with an Approximate Number System to estimate and compare different sets of items (ANS; Brannon, 2006; Cordes, Gelman, Gallistel, & Whalen, 2001; Feigenson, Dehaene, & Spelke, 2004; Libertus & Brannon, 2010). This system is rooted in the intraparietal area of the brain (Dehaene, Piazza, Pinel, & Cohen, 2003; Nieder, Freedman, & Miller, 2002; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Sawamura, Shima, & Tanji, 2002) and represents numerosities in an approximate manner on a mental number line from left to right (Dehaene, 1997). Because of these approximate numerical representations, numerosities which are close to each other on the mental number line will overlap. This representational overlap between numerosities that are closer to each other (e.g., four and five) makes it more difficult to discriminate between them than between numerosities that are further apart (e.g., three and nine).

The numerical representations in the ANS become noisier with increasing numerosity, since there is more representational overlap between larger numerosities, (e.g., Dehaene, 1992; Gallistel & Gelman, 1992). This is in correspondence with the adherence of the ANS to Weber's Law (Fechner, 1860), stating that a proportionally larger difference between two numerosities is required with increasing magnitude in order to maintain a constant level of discrimination performance. Hence, when people have to judge which of two numerosities is larger, their performance is determined by the relative and not the absolute difference between the numerosities. Overall, performance is more accurate and responses are faster when the relative difference between the numerosities increases (Barth, Kanwisher, & Spelke, 2003; Defever, Reynvoet, & Gebuis, 2013; Price, Palmer, Battista, & Ansari, 2012). The smallest relative difference or ratio between two numerosities that can be discriminated above chance can be held as an indicator for the acuity of the ANS and it has been demonstrated that this ANS acuity increases with age (e.g., Halberda & Feigenson, 2008; Van Oeffelen & Vos, 1982).

Several researchers demonstrated a significant positive relationship between ANS acuity and (future) math ability (e.g., Halberda, Mazzocco, & Feigenson, 2008; Piazza et al., 2010, but for alternative results, see De Smedt & Gilmore, 2011; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2012; Soltész, Szucs, & Szucs, 2010). Children that were more proficient at comparing different sets of numerosities had better scores on mathematical achievement tests. Consequently, it is suggested that ANS acuity can be used as a screening instrument to identify at an early age children who are at risk for future math problems (Gersten, Jordan, & Flojo, 2005; Piazza et al., 2010). Children that deviate in performance from what is expected of a child at that age might be at risk for developing mathematical deficiencies or dyscalculia.

However, to model the developmental trajectory of ANS acuity accurately and to detect deviations from normal development, a valid and

* Corresponding author at: KU Leuven, Experimental Psychology, Tiensestraat 102, bus 3711, 3000, Leuven, Belgium.
E-mail address: Karolien.smets@ppw.kuleuven.be (K. Smets).

http://dx.doi.org/10.1016/j.actpsy.2014.05.001

reliable measure is needed. To date, many researchers operate under the implicit assumption that different tasks measure the ANS in the same way and they therefore for instance include results from different tasks in a single graph describing ANS development (see Fig. 4 in Halberda & Feigenson, 2008; Fig. 3 in Piazza et al., 2010). However, recent research suggests otherwise. For instance, several studies showed that differences in the way that numerosities are presented (e.g., sequential versus parallel presentation) or different task instructions (e.g., “indicate the larger” in the comparison task or “detect the difference” in the same-different task) can affect estimates of ANS acuity (Inglis & Gilmore, 2013; Price et al., 2012; Sasanguie, Defever, Van den Bussche, & Reynvoet, 2011; Smets, Gebuis, & Reynvoet, 2013).

Differences in methodology between tasks designed to measure ANS acuity are even more pronounced when one compares tasks used in infant studies (i.e., preferential looking tasks; e.g., Libertus & Brannon, 2010) with techniques applied to older children and adults (i.e., comparison and same-different tasks; e.g., Defever, Sasanguie, Vandewaetere, & Reynvoet, 2012; Sasanguie et al., 2011). For instance, Gebuis and van der Smagt (2011) contrasted the comparison task, a task commonly used in older children and adults, with an explicit version of a habituation task typically used to test infants (i.e., a detection task; e.g., Xu & Arriaga, 2007; Xu & Spelke, 2000). Participants in the comparison task were presented with two numerosities and were instructed to indicate the larger of those two numerosities. The same participants in the detection task were presented with a continuous stream of 12 dots and occasionally a deviant numerosity was presented. Participants were instructed to detect the deviant numerosity by pressing a key. The results showed that participants performed worse on the detection task than on the comparison task, which suggests that a task that is more similar to infant tasks is more difficult than a task more commonly used to investigate adults' ANS acuity. Consequently, infant performance and their associated ANS acuity may have been underestimated in previous research.

The comparability of ANS acuity tasks used in different age groups (i.e., infants versus older children and adults) and thus the concurrent validity of different tasks is recently receiving more attention in the literature (e.g., Starr, Libertus, & Brannon, 2013). This is also the topic investigated in the current study. Specifically, we compared performance on a change detection task, which is used in infants (e.g., Libertus & Brannon, 2010), with performance on the comparison and the same-different task, which are frequently used in older children and adults (e.g., Defever et al., 2013). A direct comparison will provide more insight in the concurrent validity of these different tasks.

The change detection task we used was an adapted version of the task used in the study of Libertus and Brannon (2010). In this study, infants were presented with two streams of numerosities of which one remained constant in numerosity (e.g., 16-16-16…), while the other stream alternated between two different numerosities (e.g., 8-16-8-…). The ratio between the numerosities that alternated in the changing stream was manipulated. Libertus and Brannon (2010) found ratio-dependent looking-times: Infants looked longer at the changing stream if they were in fact able to discriminate between the numerosities that alternated in this changing stream, dependent on the ratio between these numerosities. In the present study, we transformed the change detection task of Libertus and Brannon (2010) into an explicit task. Participants were now instructed to indicate which of the two presented streams changed in numerosity. This was done to allow comparing performance on this task with performance on the explicit comparison and same-different tasks.

The present study consists of three different experiments. Since an overt response was not present in the original task (Libertus & Brannon, 2010), we tested both a direct and delayed response condition for the change detection task in Experiment 1. In the direct response condition, participants were free to answer at any given time within the trial: during the presentation of the streams of stimuli or at the end of the trial. Participants in the delayed response condition could only answer at the end of the trial. The presence of a ratio-dependent effect would suggest that this task is suitable to study ANS processing. In Experiment 2, we administered the change detection, the same-different and the comparison task to adults. This would give insight in the concurrent validity of the three tasks. In Experiment 3, we administered the change detection task to primary school children. The inclusion of primary school age children will show whether this task is suitable to test the ANS in children. Furthermore, the administration of this task will provide us with a measure of ANS acuity across different ages, obtained with a task that is more compatible with the task that is used in infants. Additionally, we compared performance of primary school children on the change detection task to performance of age-matched children on the comparison and same-different task. The latter results were derived from a previous study.

1. Experiment 1

1.1. Method

1.1.1. Participants

Thirty participants participated in the change detection task with a direct response (mean age = 20 years, 16 female) and fifteen participants took part in the change detection task with a delayed response (mean age = 24 years, 13 female). Participants either received course credits or were paid for their participation in the experiment. The experiment was approved by the Ethical Committee of the Faculty of Psychology and Educational Sciences of the University of Leuven. All participants gave written informed consent for their participation.

1.1.2. Stimuli and procedure

Participants in both the direct and delayed response condition were presented with two streams of non-symbolic numerosities, one on the left and one on the right side of the screen (see Fig. 1). On each trial, the same numerosity was presented in one stream (e.g., 9 dots, 9 dots, 9 dots…), while two different numerosities alternated in the other stream (e.g., 9 dots, 18 dots, 9 dots…). The stimuli were dot patterns ranging from 8 to 35 (see Table 1) and were created with an adapted version of the program developed by Gebuis and Reynvoet (2011). With this program, stimuli are created whose sensory properties are uninformative about numerosity across trials. Consequently, an increase in numerical distance is not associated with an increase in sensory properties across trials. To this end, in half of the trials, the different visual cues (dot diameter, convex hull, contour length, aggregate surface and density) are congruent with number and in the other half of the trials the visual cues are incongruent with number. We manipulated the ratio (larger numerosity/smaller numerosity) by which the numerosities within the changing stream differed, resulting in five different ratio conditions: 1.2, 1.25, 1.5, 2 and 2.5. The smallest ratio was the most difficult and the largest was the easiest.

Each trial started with a red fixation cross that was presented for 1000 ms, followed by a green fixation cross that remained on the screen for 500 ms. Next, the two streams of dot patterns were presented. Each dot pattern remained on the screen for 500 ms and a black screen was displayed for 500 ms in between the dot patterns (see also Libertus & Brannon, 2010). The stimuli in both streams changed seven times during one trial (see Fig. 1). Participants in the direct response condition could either respond during the presentation of the streams or at the end of the trial when a question mark was displayed. Participants in the delayed response condition could only respond at the end of the trial when the question mark was displayed. In both conditions, participants were instructed to indicate the stream that changed in numerosity by pressing a key at the corresponding side. For each ratio condition, there were 4 possible number pairs, which were repeated 16 times, resulting in 64 trials per ratio condition (5 ratios * 4 number pairs * 16 trials = 320 trials in total). Participants were administered